

# **SCIENTIFIC PROGRAMME**



**KEYNOTE**



# A Software Architecture for Multi-Paradigm Modelling

Juan de Lara

E.T.S. de Informática  
Universidad Autónoma de Madrid  
Ctra. Colmenar km. 15, Campus Cantoblanco  
28049 Madrid, Spain  
*Juan.Lara@ii.uam.es, jlara@cs.mcgill.ca*

Hans Vangheluwe

School of Computer Science  
McGill University  
3480 University Street  
Montréal, Québec, Canada H3A 2A7  
*hv@cs.mcgill.ca*

Ghislain Vansteenkiste

BIOMATH Department  
Ghent University  
Coupure 653  
9000 Ghent, Belgium  
*Ghislain.Vansteenkiste@rug.ac.be*

## Abstract

In this article we discuss the limitations found in regular programming language types when used in the context of *multi-paradigm* modelling. By multi-paradigm, we mean the combination of meta-modelling (*i.e.*, modelling models), multi-formalism modelling, and multiple abstraction levels. In particular, we propose the inclusion, in types, of information about *how time is represented* and ultimately simulated in each component of a system model. We also discuss the relationship between *formalisms* and *types*, and propose an architecture to implement these multi-paradigm concepts. This architecture has been implemented in a tool called AToM<sup>3</sup> [7], which allows one to model different parts of a system using different formalisms. Formalisms are modelled at a meta-level and AToM<sup>3</sup> uses the information in these meta-models to automatically *generate tools* to process (create, edit, check, and generate simulators for) the models in the described formalism. Models can be automatically converted between formalisms thanks to information found in a *Formalism Transformation Graph* (FTG). The transformations are described at a meta-level by means of models in the graph-grammar formalism. Composite types are described by constructing models in the *Types* formalism which has also been modelled at a meta-level within AToM<sup>3</sup>. *Graph grammars* are used to manipulate types. Examples of these manipulations include automatic generation of widgets to edit variables of the defined type, and determining subtype/supertype relationships.

**Keywords:** Modelling & Simulation, Multi-Paradigm Modelling, Automatic Code Generation, Graph Grammars, Types.

## 1 Introduction

AToM<sup>3</sup> is a visual Meta-Modelling tool which supports modelling of complex systems. Complex systems are characterized by components and aspects whose structure as well as behaviour cannot be described in a single formalism. Examples of commonly used modelling formalisms are Differential-Algebraic Equations (DAEs), Bond Graphs, Petri Nets, DEVS, Entity-Relationship Di-

agrams, and Statecharts.

In multi-formalism modelling, each system component and/or aspect represented is modelled using the most appropriate formalism and supporting modelling/simulation tool. To deal with this formalism heterogeneity, a single formalism needs to be identified into which each of the component models can be symbolically transformed [25]. Obviously, the system properties which we wish to investigate must be invariant under the transformations. The formalism to transform to depends on the question to be answered about the system. The Formalism Transformation Graph (see Figure 1) suggests DEVS [27] as a universal (bridging continuous-time and discrete-event worlds) common modelling formalism for simulation purposes (generating input/output trajectories). To answer symbolic questions, as when performing system verification, other target formalisms are more appropriate.

In order to make the multi-formalism approach applicable, we still have to solve the problem of interconnecting a plethora of different tools, each designed for a particular formalism. We tackle this problem by means of *meta-modelling*. Thus, using a higher layer of modelling, it is possible to model the modelling formalisms themselves. Using the information in these meta-layers, it is possible to generate customized tools for models in the described formalisms. The effort required to construct a tool for modelling in a formalism tailored to particular applications thus becomes minimal. Furthermore, explicitly modelling the formalism yields insight into it. When the generated tools use a common data structure to internally represent the models, transformation between formalisms is reduced to the transformation of these data structures. In AToM<sup>3</sup>, the basic data structures are graphs so transformations may be specified (modelled) in the graph grammar formalism.

In this article, in the context of multi-paradigm modelling, we propose extending the types commonly used in programming languages with information about how time is represented. To our knowledge, there is no modelling and simulation tool treating types in this way. Next, we point out some similarities and analogies between *formalisms* and *types*. We also show the implementation of these ideas in AToM<sup>3</sup>. AToM<sup>3</sup> has a meta-modelling layer in which dif-

ferent formalisms can be modelled. It is possible to model the graphical representation of the relevant entities in these models at the meta-level. From the meta-specification (often in the Entity Relationship formalism) of formalism  $F$ , AToM<sup>3</sup> generates a tool to process models described in  $F$ . Models are internally represented using *Abstract Syntax Graphs*. In AToM<sup>3</sup>, types are also models, and are defined by constructing a graph. The graph may contain strongly connected components in case of recursive types.

Model manipulation is specified in graph-grammar [9] models. Examples of such manipulations are:

- Given a model of a type, generating appropriate graphical widgets to edit variables of the defined type.
- Given two type models, determine if one is a sub-type of the other.
- Transform a model expressed in one formalism into a behaviourally-equivalent model, possibly expressed in another formalism.
- Optimize a model, without formalism change. Examples of such optimizations are *constant folding* and *sorting* of equations in the DAE formalism.
- Express operational semantics (specification of simulators).
- Generate code from the model. For example, the model represented in a standard simulation language such as GPSS [6].

Graph grammars (similar to string grammars) are composed of a list of rules, each of which has a left and a right hand side. A graph rewriting system tries each rule in turn. Whenever a matching is found between a left hand side and a zone in the graph, this zone of the graph is substituted by the right hand side of the rule. The graph rewriting system stops when no more rules are applicable.

Although graph grammars have been used in very diverse areas such as graphical editors, code optimization, computer architecture, etc. [11], to our knowledge, they have never been applied to formalism transformations nor to type manipulations.

## 2 Multi-paradigm M&S

Computer Automated Multi-Paradigm Modelling is an emerging field that addresses and integrates three orthogonal directions of research:

1. multi-formalism modelling, concerned with the coupling of and transformation between models described in different formalisms,
2. model abstraction, concerned with the relationship between models at different levels of abstraction, and
3. meta-modelling (models of models), concerned with the description of classes of models, which allows for formalism specification.

Multi-paradigm modelling explores the possible combinations of these notions. It combines and relates formalisms, generates maximally constrained domain- and problem-specific formalisms, methods and tools, and verifies consistency between multiple views. Because of the heterogeneous nature of embedded systems and the many implementation technologies, multi-paradigm modelling is a critical enabler for holistic design approaches (such as mechatronics), to avoid overdesign, and to support system integration. Multi-paradigm techniques have been successfully applied in the field of software architectures, control system design, model integrated computing, and tool interoperability.

Table 2 depicts the levels considered in our meta-modelling approach.

Formalisms such as Entity-Relationship Diagrams are often used for meta-modelling. To be able to fully specify modelling formalisms, the meta-level formalism may have to be extended with the ability to express constraints (limiting the number of meaningful models). For example, when modelling a Deterministic Finite Automaton, different transitions leaving a given state must have different labels. This cannot be expressed within Entity-Relationship diagrams alone. Expressing constraints is most elegantly done by adding a constraint language to the meta-modelling formalism. Whereas the meta-modelling formalism frequently uses a graphical notation, constraints are concisely expressed in textual form. For this purpose, some systems [16] (including ours) use the Object Constraint Language OCL [21] used in UML. As AToM<sup>3</sup> is implemented in the scripting language Python [23], arbitrary Python code can also be used.

Our use of graph transformations allows to express model *behaviour* and thus formalism semantics. These graph transformations allow us to transform models between formalisms, optimize models, or describe basic simulators. For example, we have implemented a simulator for block diagrams using graph grammars [4]. Another advantage of our approach, is that we consider meta-levels; we don't need different tools to process different formalisms, as we can model them at the meta-level.

Other approaches to interconnecting formalisms are Category Theory [13], in which formalisms are cast as categories and their relationships as functors. See also [26] and [20] for other approaches.

There are other visual tools to describe formalisms using meta-modelling, among them DOME [8], Multigraph [24], MetaEdit+ [19] or KOGGE [10]. Some of these allow one to express formalism semantics by means of a textual language (KOGGE for example uses a language similar to Modula-2). Our approach is quite different, as we express such semantics by means of graph grammar models. We believe that graph-grammars are a natural, declarative, and general way to express transformations. As graph grammars are highly amenable to graphical representation, they

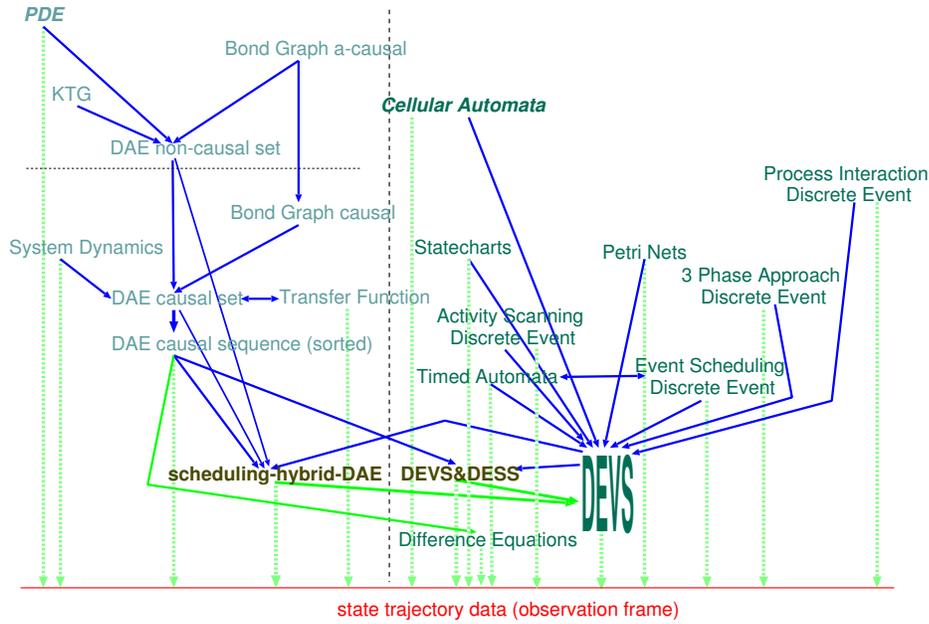


Figure 1: Formalism Transformation Graph.

Level	Description	Example
Meta-Meta-Model	Model that describes a formalism that will be used to describe other formalisms.	Description of Entity-Relationship Diagrams, UML class Diagrams
Meta-Model	Model that describes a simulation formalism. Specified under the rules of a certain Meta-Meta-Model	Description of Deterministic Finite Automata, Ordinary differential equations (ODE)
Model	Description of an object. Specified under the rules of a certain Meta-Model	$f'(x) = -\sin x, f(0) = 0$ (in the ODEs formalism)

Table 1: Meta-modelling levels.

are superior to a purely textual language. Also, none of the tools consider the possibility of “translating” models between different formalisms.

Also, there are some languages and systems for graph-grammar manipulation, such as PROGRES [22], GRACE [15] and AGG [2]. All of them lack a meta-modelling layer.

Our approach is original in the sense that we combine the advantages of meta-modelling (to avoid explicit programming of customized tools) and graph transformation systems (to express a tool’s behaviour and formalism transformation). Our main contribution is in the field of multi-paradigm modelling [25], as we have a general means to transform models between different formalisms.

### 3 Programming vs. Modelling Types

Programming languages usually do not include the notion of time in types. However in simulation, state variables are

functions of time. For example, an *Integer* variable  $a$  is really  $a : TimeBase \rightarrow \mathbb{N}$ . In programming languages (as opposed to modelling languages), one is not interested in the detailed evolution over time of variables, only in the sequence of changes. In a multi-formalism modelling environment, having information about how a certain variable may change with respect to time can be useful. In complex systems, there may be components which are best modelled using a discrete time formalism (the time base is isomorphic to  $\mathbb{N}$ ), whereas for others, a continuous time formalism (the time base is isomorphic to  $\mathbb{R}$ ) is preferred. In the first case, *Integer* variables should be expressed as the composite type  $\mathbb{N} \rightarrow \mathbb{N}$  whereas in the second case they should be expressed as  $\mathbb{R} \rightarrow \mathbb{N}$ . This has important implications, mostly with respect to subtype relationships, as we will see below.

In object-oriented programming languages, a type  $O'$  is a subtype of type  $O$  if  $O'$  has the same components as  $O$

and possibly more [1]. We denote such a relationship as  $O' <: O$ . The existence of such a relationship enables us to perform operations on variables, such as replacement: if  $O' <: O$ , then it is allowed to replace any variable of type  $O$  by a variable of type  $O'$ . For atomic types, this relationship is postulated. As an example, usually one considers  $Int <: Float$ , and allows assignments such as  $f := i$  with  $typeof(f) = Float$  and  $typeof(i) = Int$ . The type system must then convert the  $Int$  variable into a  $Float$ . When such conversion is automatically performed by the system, it is called *coercion* [3]. For some other cases, an explicit conversion mechanism must be given.

Composite types are built using *type constructors*. Examples of such constructors are “ $\times$ ” to build tuples, “ $\cup$ ” for unions, and “ $\rightarrow$ ” for functions.

Using the above, suppose we want to model Integers in Discrete and Continuous Time components. Using the  $\rightarrow$  type constructor, we end up with:  $\langle CT\_Int \rangle ::= \langle Float \rangle \rightarrow \langle Int \rangle$  and  $\langle DT\_Int \rangle ::= \langle Int \rangle \rightarrow \langle Int \rangle$ . Where  $CT$  stands for Continuous Time and  $DT$  stands for Discrete Time.

The  $\rightarrow$  operator is contravariant [1], that is,  $A \rightarrow B <: A' \rightarrow B'$  iff  $A' <: A$  and  $B <: B'$ . In our example, that means that  $\langle CT\_Int \rangle <: \langle DT\_Int \rangle$ . This relationship is very useful, as it allows us to discover incorrect assignment between variables whose time-dependence differs. Using the derived sub-type relationship, the assignment  $\langle DT\_Int \rangle ::= \langle CT\_Int \rangle$  would be allowed (coercion), whereas  $\langle CT\_Int \rangle ::= \langle DT\_Int \rangle$  would not, an explicit type casting is needed. This makes sense, as a Continuous Time component will need information at time values which a Discrete Time component does not even consider. Using the programming language type  $\langle Int \rangle$  for both continuous and discrete components would not catch such a wrong assignment. This is shown graphically in Figure 2, where we have tried to connect different models (continuous and discrete time) via  $\langle DT\_Int \rangle$  and  $\langle CT\_Int \rangle$  ports. Here, the semantics of connections is given by replacing them by assignments of the port variables. It can be seen that a causal connection from a Discrete Time System Specification component (DTSS) to a Continuous Time component (an Ordinary Differential Equation model, ODE) is not allowed. Therefore an intermediate module such as an interpolator would be needed to provide information about the variable at instants of time which the DTSS module does not provide. In the example in Figure 2 the interpolator would be a zero order hold. In this way, this interpolator would provide the mechanism for performing an “*explicit casting*” between those types.

The  $\langle Int \rangle$  type could also be used to store values for time independent variables (constants). The type must automatically be promoted to  $\langle CT\_Int \rangle$  or  $\langle DT\_Int \rangle$ . The reason is easily seen in the following example: consider the expression (in a continuous time formalism)  $a(t) := 2 + b(t)$ . The types are  $\langle CT\_Int \rangle ::= \langle Int \rangle + \langle CT\_Int \rangle$ , therefore we need to coerce  $\langle Int \rangle$  into  $\langle CT\_Int \rangle$  to be able to sum them (using the Continuous Time add operator  $+_{CT} : \langle CT\_Int \rangle \times$

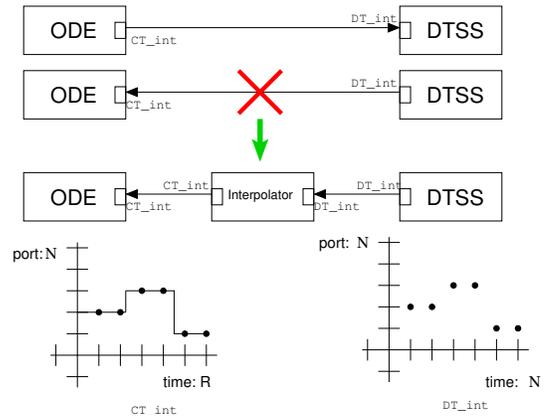


Figure 2: Valid and invalid model connections

$\langle CT\_Int \rangle \rightarrow \langle CT\_Int \rangle$ ). Similarly for a discrete time formalism.

### 3.1 Formalisms vs. Types

We use composite types as a means to structure data. One interpretation of a type is the set of all possible values a variable of that type can take. On the other hand, using a formalism to build a model imposes certain syntactic rules, in a way similar to a type. As such, a formalism describes the set of all legal models in the formalism. Note also the close relationship between (the syntactic part of) a formalism and a modelling language. This indicates how input (syntax analysis) and output of models in a standard syntax such as XML could be automated on a formalism meta-model. A formalism, as opposed to a type, also includes semantics (in particular would be desirable execution semantics).

In multi-paradigm modelling, the equivalent of *coercion* and *explicit type conversions* in Types would be desirable. For this purpose, we need structural comparison between formalisms. In a multi-paradigm environment, this is possible, as the paradigms themselves have been modelled using a meta-meta-language (Entity-Relationship in our case).

If we define formalism inheritance in the usual way, as syntactic concatenation with latest-overrides semantics<sup>1</sup>, then it is easy to find sub-formalism relationships. For example, consider the Non-Deterministic Finite Automaton (NFA) formalism. For models in this formalism, we allow definition of states and transitions, and connections between them. We could define the Deterministic Finite Automata (DFA) as a refinement (through inheritance) of the NFA formalism, in which we add the constraint that transitions departing from the same state must have different conditions.

Figure 3 shows this refinement. On the upper part, the NFA meta-model is expressed as an Entity-Relationship diagram extended with OCL constraints. The latter are represented

<sup>1</sup>Note how the semantics of inheritance within a formalism can be modelled in ATOM<sup>3</sup> by means of a graph-grammar model

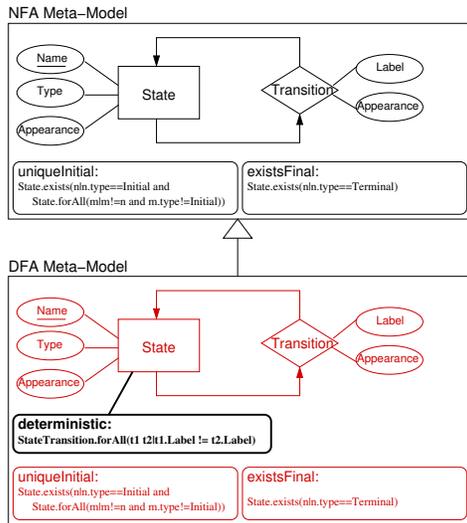


Figure 3: Defining the DFA formalism through extension of the NFA formalism

as rounded rectangles. Local constraints are connected to the entity which they constrain. Global constraints appear unconnected. In the lower part of the figure, the DFA meta-model is defined by inheriting from the NFA meta-model. Inherited elements are shown in lighter color. We should not allow removing inherited elements, but we allow its refinement. We also permit adding new elements to the meta-model.

Adding constraints to meta-models reduces the set of allowed models: exactly the meaning of sub-type. Defining NFAs and DFAs in this way, we have that  $DFA <: NFA$  and we automatically allow assignments such as  $\langle NFA \rangle := \langle DFA \rangle$ , whereas we reject  $\langle DFA \rangle := \langle NFA \rangle$ . For the latter case we need an explicit type conversion, or more properly, a formalism transformation. We can find out whether such a transformation exists from the Formalism Transformation Graph. We express the actual transformation by means of a graph-grammar model.

## 4 AToM<sup>3</sup>: an overview

AToM<sup>3</sup> is a tool written in Python [23] which uses and implements the concepts presented above. Its architecture is shown in Figures 4 and 5. In both figures, models are represented as white boxes, having on their upper-right corner an indication of the meta-...model they were specified with. It can be seen that in the case of a graph-grammar model, to convert a model in formalism  $F_{source}$  to  $F_{dest}$ , it is necessary to use the meta-models of both  $F_{source}$  and  $F_{dest}$  together with the meta-model of graph-grammars.

The main component of AToM<sup>3</sup> is the kernel, which is responsible for loading, saving, creating and manipulating models (at any meta-level, via the Graph-Rewriting Module), as well as for generating code for customized tools.

Both Meta-models and meta-meta-models can be loaded into AToM<sup>3</sup> as shown in Figure 4. The first kind of models allows constructing valid models in a certain formalism, the second are used to describe the formalisms themselves.

The Entity-Relationship formalism extended with constraints is available at the meta-meta-level. Constraints can be specified as OCL or Python expressions, and the designer must specify when (pre- or post- and on which event) the condition must be evaluated. Events can be *semantic* (such as editing an attribute, connecting two entities, etc.) or *graphical* (such as dragging, dropping, etc.)

When modelling at the meta-meta-level, the entities which may appear in a model must be specified together with their attributes. We will refer to this as the semantic information. For example, to define the Petri Net Formalism, it is necessary to define both Places and Transitions. Furthermore, for Places we need to add the attribute name and number of tokens. For Transitions, we need to specify the name.

In general, in AToM<sup>3</sup> we have two kinds of attributes: *regular* and *generative*. Regular attributes are used to identify characteristics of the current entity. Generative attributes are used to generate new attributes at a lower meta-level. The generated attributes may be generative in their own right. Both types of attributes may contain data or code for pre- and post-conditions. Only meta-meta-level entities are provided with generative attributes.

The meta-meta-information is used by the kernel to generate a meta-model, which, when loaded by the kernel, allows the processing of models in the defined formalism.

In the meta-model, it is also possible to specify the graphical appearance of each entity in the lower meta-level. This appearance is, in fact, a special kind of generative attribute. For example, for Petri Nets, we can choose to represent Places as circles with the number of tokens inside the circle and the name beside it, and Transitions as thin rectangles with the name beside them. That is, we can specify how some semantic attributes are displayed graphically. Constraints can also be associated with the graphical entities. Each graphical form, part of the graphical entity, can be referenced by an automatically generated name which has methods to change its color, or hide it.

The left side of Figure 6 shows AToM<sup>3</sup> being used to describe the Petri Net formalism in the Entity-Relationship formalism. Entity and relationships can be edited. This includes editing semantic attributes as well as graphical appearance (and the link between both).

The right side of Figure 6 shows AToM<sup>3</sup> loaded with the meta-model for processing Petri-Net models, generated automatically from the previous description and in use for editing a Petri Net model of a producer-consumer process.

For the implementation of the Graph Rewriting Processor, we have used an improvement of the algorithm given in [9], in which we allow non-connected graphs in LHS's in rules. It is also possible to define a sequence of graph-grammars that have to be applied to the model. This is useful, for ex-

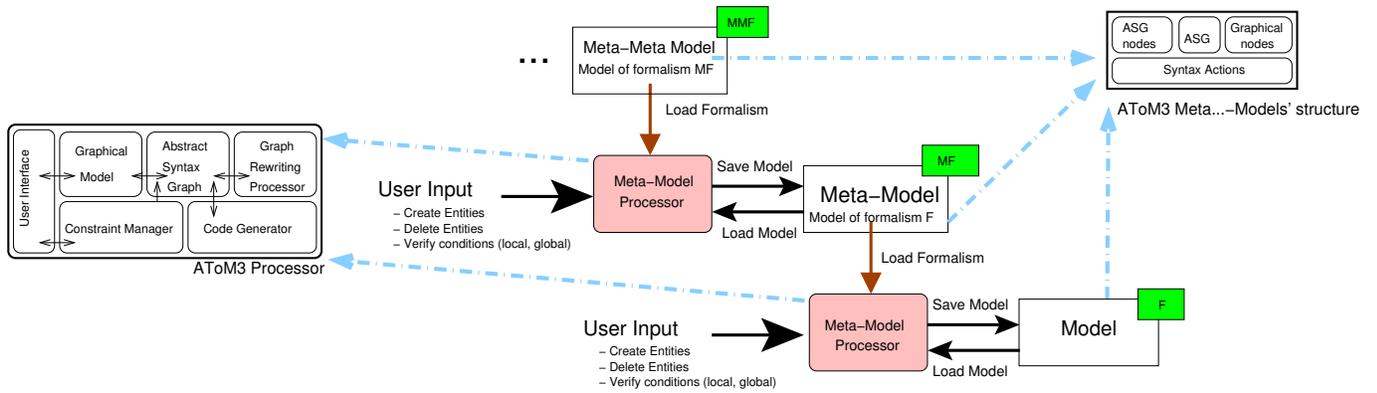


Figure 4: Meta-... Modelling in ATOM<sup>3</sup>.

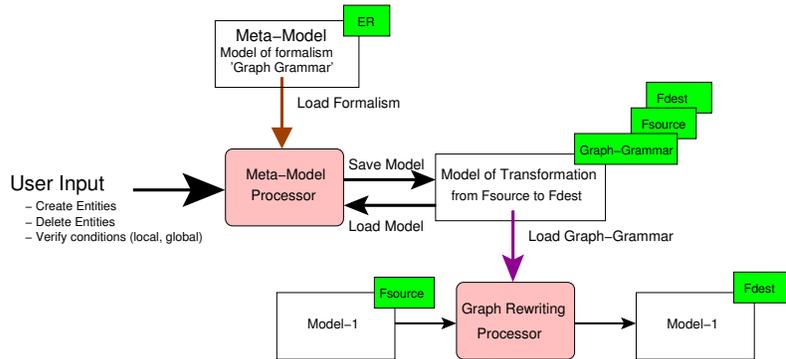


Figure 5: Model Transformation in ATOM<sup>3</sup>.

ample, to couple grammars to convert a model into another formalism, and to then apply an optimization. Controlling the execution of the rules (stopping after each rule execution or continuous execution) is also possible. As the LHS of a rule can match different subgraphs of the host graph, we can also control whether the rule must be applied in all the subgraphs (if disjoint), if the user can choose one of the matching subgraphs interactively, or the system chooses a random one. As in grammars for formalism transformations we have a mixing of entities belonging to different formalisms, it must be possible to open several meta-models at the same time (see Figure 5). Obviously, the constraints of the individual formalism meta-models are meaningless when entities in different formalisms are present in a single model. Such a model may come to exist during the intermediate stages of graph grammar evaluation when transforming a model from one formalism into another. It is thus necessary to disable evaluation of constraints during graph grammar processing (i.e., all models are reduced to Abstract Syntax Graphs).

## 5 Defining and manipulating Types

ATOM<sup>3</sup> has a number of basic types, such as Strings, Integers, and Floats. Each type in ATOM<sup>3</sup> has an associated

Python class, which is responsible for creating a widget to edit its value, checking the validity of its value, making the variable persistent, etc.

It is also possible to define composite types. Following the tool's philosophy, composite types are models, which have a meta-model in their own right (the "Types" formalism). Thus, types are defined as graphs, as described in [3]. Type models have a *Root* node, which is labeled with the type's name. From the *Root* node, *Operator* nodes can be connected. The *Operator* type can be either *Product* (to build tuples), *Union* or *Function*. *Operator* nodes can be connected to other *Operator* nodes, to the *Root* node or to *LeafType* nodes. We allow recursive types, with backwards connections. The *LeafType* node type can be any valid type (basic or composite) in the current ATOM<sup>3</sup> session. *LeafType* nodes cannot have any outgoing connection, thus, these kind of nodes are the graph leaves. *LeafType* nodes may have associated a constraint to restrict the type value. This is useful for example if we try to define subranges of a type.

A tool for processing types was generated automatically from this meta-description and then incorporated into the ATOM<sup>3</sup> core. Consequently, types may be loaded, saved and manipulated as any other model. The meta-level has been

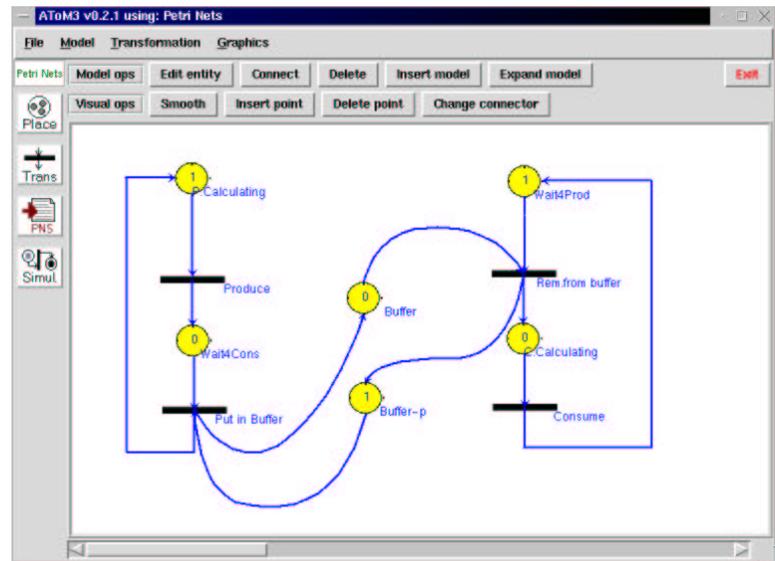
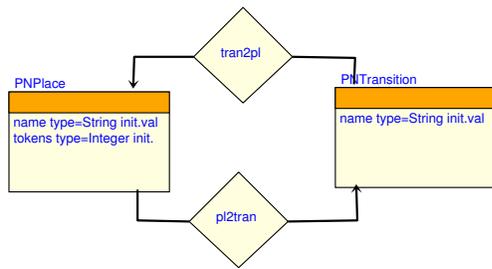


Figure 6: ER Modelling the Petri-Nets formalism (left) and Generated tool for processing Petri-Nets models (right)

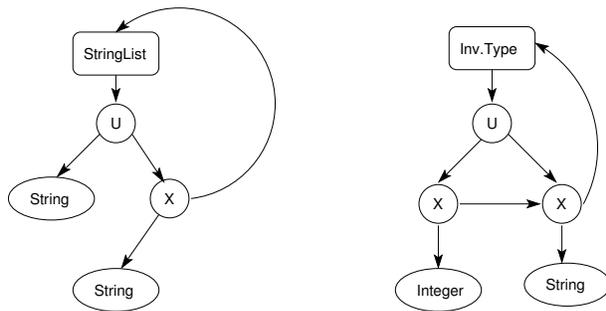


Figure 7: A model of a string list type and an invalid recursive type

constrained to avoid closed, infinite recursion in the type definition. A cyclic type graph is valid if there exists at least one Union Operator node in a cycle and at least one Union Operator node in the cycle reaches terminal nodes. When checking terminal nodes' reachability from a Product Operator node, one has to check all the outgoing connections (all must end up in a terminal node).

Figure 7 shows a valid type graph (a list of Strings with at least one element) and an invalid recursive type. The latter is invalid, as from the Union node one cannot reach a terminal node (from both Product nodes one should reach terminal nodes from all the outgoing connections, which is not the case).

Once a model for the type has been built, ATOM<sup>3</sup> generates a widget to edit variables of that type. Furthermore, the type's model is used for type-checking. As has been mentioned, model manipulation is expressed by means of graph grammars. Thus, code generation for type models has been implemented as a graph grammar. Another graph grammar has been defined to find subtype relationships between type

models.

The graph grammar for Python code generation for types is composed of an initial action and three rules. The initial action adds a 2-element tuple to each node. Its first element is a flag which indicates whether the code for the node has been generated yet, the second one holds the name of the generated Python class.

*LeafType* nodes do not need their code to be generated, as they pre-exist in the system or they are composite types whose code has been generated before. For this kind of nodes this tuple is set to (1, < name of class >). For *Operator* nodes, the slot is set to (0, < given name >), where < given name > is a unique name assigned to each *Operator* node by the initialization routine. For the *root* Node, the tuple value is set to (1, < type name >).

Each one of the three rules recognizes different situations in the type graph, the order in which they are applied does not matter in our case. They are:

- The first rule is applied if a *Product Operator* is found, and the first element of the auxiliary tuple has a value of 0. When the rule is executed, this value is changed to 1, and the code is actually generated.
- The second and third rules are similar to the first, but search for *Union Operator* and *Root* nodes, and the code generation routines are different.

## 6 Conclusions and future work

In this paper we have discussed the advantages of including time information in types in multi-formalism modelling. When such information is not present (such as when using regular programming languages types), some wrong assignments may remain undiscovered during syntactic checks of the model. Note how this also indicates that multi-

formalism modelling languages should allow explicit representation of the time base. In Modelica [12], model annotations are currently used, which is not a general solution. We have also discussed the similarities between types and formalisms, and have drawn some analogies, with respect to coercion, and between explicit type conversions and formalism transformations.

We have also presented AToM<sup>3</sup>, a meta-modelling tool that is able to generate customized, formalism-specific tools. As models are stored in the form of graphs, AToM<sup>3</sup> can manipulate them using graph-grammars. Types are also treated as models, and have their own meta-model (“Types” formalism). This adds flexibility, as type manipulation can be expressed by means of graph grammar models.

The use of a model (in the form of a graph grammar) of graph transformations has some advantages over an implicit representation (embedding the transformation computation in a program) [5]:

- It is an abstract, declarative, high level representation. It can be used to reason about the transformation in a declarative fashion. It is expected inheritance will simplify management of many similar transformations.
- The theoretical foundations of graph rewriting systems can assist in proving correctness and convergence properties of the transformation tool.

On the other hand, the use of graph grammars is constrained by efficiency. In the most general case, subgraph isomorphism testing is NP-complete. However, the use of small subgraphs on the left hand side of graph grammar rules, as well as using node labels and edge labels greatly reduce the search space.

The advantages of our code-generating approach are also clear: instead of building a whole application from scratch, it is only necessary to specify –in a graphical manner– the kinds of models we will deal with. Our approach is also highly applicable if we want to work with a slight variation of some formalism, where we only have to specify the meta-model for the new formalism and a transformation (or a sequence of transformations found as a path in the Formalism Transformation Graph) into a “known” formalism (one that already has a simulator available, for example). We then obtain a modelling tool for the new formalism, and are able to convert models in this formalism into the other for further processing. A side effect of this code-generating approach is that some parts of the tool have been bootstrapped. An example of this is the dialog to specify composite types: the meta-model for this graph has been specified with AToM<sup>3</sup>, and subsequently Python code was automatically generated.

We are currently able to describe the dynamic semantics of some formalisms using graph grammars, and thus generate simulators for these formalisms.

In the future, we plan to extend the tool in several ways:

- Describing another meta-meta-model in terms of

the current one (the Entity-Relationship meta-meta-model) is also possible. In particular, we are currently describing UML class diagrams. For this purpose, relationships between classes such as *inheritance* are modelled described. Thanks to our meta-modelling approach, we will be able to describe different subclassing semantics and their relationship with subtyping [1]. Furthermore, as the semantics of inheritance will be described at the meta-level, code can be generated in non-object-oriented languages.

- Extending the tool to allow collaborative modelling: for this purpose, we are working on putting the APIs for constructing graphical interfaces in Java (Swing) and Python (Tkinter) at the same level. These developments, together with the possibility of using Python on top of the Java Virtual Machine (e.g., by means of Jython [18]), will allow us to make our tool in applet form accessible through a web browser. This possibility as well as the need to exchange and re-use (meta...) models raises the issue of formats for model exchange. A viable candidate format is XML.

## Acknowledgements

Juan de Lara’s work has been partially sponsored by the Spanish Interdepartmental Commission of Science and Technology (CICYT), project number TEL1999-0181. Prof. Vangheluwe gratefully acknowledges partial support for this work by a National Sciences and Engineering Research Council of Canada (NSERC) Individual Research Grant.

## References

- [1] Abadi, M., Cardelli, L. 1996. *A Theory of Objects*. Monographs in Computer Science. Springer
- [2] AGG Home page: <http://tfs.cs.tu-berlin.de/agg/>
- [3] Aho, A.V., Sethi, R., Ullman, J.D. 1986. *Compilers, principles, techniques and tools*. Chapter 6, *Type Checking*. Addison-Wesley.
- [4] AToM<sup>3</sup> home page: <http://moncs.cs.mcgill.ca/MSDL/research/projects/ATOM3.html>
- [5] Blonstein, D., Fahmy, H., Grbavec, A.. 1996. *Issues in the Practical Use of Graph Rewriting*. Lecture Notes in Computer Science, Vol. 1073, Springer-Verlag, pp.38-55.
- [6] de Lara, J. and Vangheluwe, H. Using meta-modelling and graph grammars to process GPSS models. In Hermann Meuth, editor, *16<sup>th</sup> European Simulation Multi-conference (ESM)*, pages 100–107. Society for Computer Simulation International (SCS), June 2002. Darmstadt, Germany.
- [7] de Lara, J. and Vangheluwe, H. AToM<sup>3</sup>: A tool for multi-formalism and meta-modelling. In *European*

- Joint Conference on Theory And Practice of Software (ETAPS), Fundamental Approaches to Software Engineering (FASE)*, LNCS 2306, pages 174 – 188. Springer-Verlag, April 2002. Grenoble, France.
- [8] DOME guide. <http://www.htc.honeywell.com/dome/>, Honeywell Technology Center. Honeywell, 1999, version 5.2.1
- [9] Dorr, H. 1995. *Efficient Graph Rewriting and its implementation*. Lecture Notes in Computer Science, 922. Springer.
- [10] Ebert, J., Sttenbach, R., Uhe, I. *Meta-CASE in Practice: a Case for KOGGE* In A. Olive, J. A. Pastor: Advanced Information Systems Engineering, Proceedings of the 9th International Conference, CAiSE'97, Barcelona, Catalonia, Spain, June 16-20, 1997 LNCS 1250, 203-216, Berlin, 1997. See KOGGE home page at: <http://www.uni-koblenz.de/ist/kogge.en.html>
- [11] Ehrig, H., Kreowski, H.-J., Rozenberg, G. (eds.) 1991. *Graph Grammars and their application to Computer Science: 4th International Workshop, Bremen, Germany, March 5-9, 1990, Proceedings*. Lecture Notes in Computer Science, Vol. 532, Springer.
- [12] Elmqvist, H., Mattson, S.E. 1997 *An Introduction to the Physical Modeling Language Modelica*. Proceedings of the 9th European Simulation Symposium ESS'97. SCS Int., Erlangen, pp.:110-114. See also: <http://www.Modelica.org>
- [13] Fiadeiro, J.L., Maibaum, T. 1995. *Interconnecting Formalisms: Supporting Modularity, Reuse and Incrementality* Proc.3rd Symposium on the Foundations of Software Engineering, G.E.Kaiser(ed),pp.: 72-80, ACM Press.
- [14] Finkelstein, A., Kramer, J., Goedickie, M. *ViewPoint Oriented Software Development* Proc. of the Third Int. Workshop on Software Engineering and its Applications, Toulouse, December 1990.
- [15] GRACE Home page: <http://www.informatik.uni-bremen.de/theorie/GRACEland/GRACEland.html>
- [16] Gray J., Bapty T., Neema S. 2000. *Aspectifying Constraints in Model-Integrated Computing*, OOPSLA 2000: Workshop on Advanced Separation of Concerns, Minneapolis, MN, October, 2000.
- [17] Harel, D. On visual formalisms. *Communications of the ACM*, 31(5):514–530, May 1988.
- [18] Jython Home Page: <http://www.jython.org>
- [19] Kelly, S., Lyytinen, K., Rossi, M. *MetaEdit+: A fully configurable Multi-User and Multi-Tool CASE and CAME Environment* In Constantopoulos, P., Mylopoulos, J., Vassiliou, Y: Advanced Information System Engineering; LNCS 1080. Berlin: Springer 1996. See MetaEdit+ Home page at: <http://www.MetaCase.com/>
- [20] Niskier, C., Maibaum, T., Schwabe, D. 1989 *A pluralistic Knowledge Based Approach to Software Specification* 2nd European Software Engineering Conference, LNCS 387, Springer Verlag 1989, pp.:411-423
- [21] OMG Home Page: <http://www.omg.org>
- [22] PROGRES home page: <http://www-i3.informatik.rwth-aachen.de/research/projects/progres/main.html>
- [23] Python home page: <http://www.python.org>
- [24] Sztipanovits, J., et al. 1995. "MULTIGRAPH: An architecture for model-integrated computing". In ICECCS'95, pp. 361-368, Ft. Lauderdale, Florida, Nov. 1995.
- [25] Vangheluwe, H. *DEVS as a common denominator for multi-formalism hybrid systems modelling*. In Andras Varga, editor, *IEEE International Symposium on Computer-Aided Control System Design*, pages 129–134. IEEE Computer Society Press, September 2000. Anchorage, Alaska.
- [26] Zave, P., Jackson, M. 1993. *Conjunction as Composition* ACM Transactions on Software Engineering and Methodology 2(4), 1993, 371-411.
- [27] Zeigler, B., Praehofer, H. and Kim, T.G. *Theory of Modelling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems*. Academic Press, second edition, 2000.



# **SOFTWARE RE-USE AND ANALYSIS**



# INDICATORS SYSTEM FOR REACTIVITY EVALUATION OF A PUBLIC KEY INFRASTRUCTURE

Ioana Filipaş, Maxime Wack, Bernard Mignot and Abdellah El Moudni  
Laboratory of Systems and Transports  
University of Technology at Belfort-Montbéliard  
90010 Belfort cedex, France  
{ioana.filipas,maxime.wack,bernard.mignot,abdellah.elmoudni}@utbm.fr

## KEYWORDS

Performance Indicator System, Public Key, Reactivity.

## ABSTRACT

We are working on a computer science environment achieving four functionalities: one certification of electronic transactions through a third certificate party, the secured storage through a third archive party, a toolset for information researching and tracking of the different transactions.

A such storage system sets performance problems, notably in term of answer time (reactivity). In this article we propose a system for reactivity analysis of an architecture based on Public Key Infrastructure.

To evaluate and improve the reactivity, it is therefore important to elaborate an indicators performance system adapted to reactivity.

## 1. INTRODUCTION

Since the French law n°2000-230 in date of march 13<sup>th</sup>, 2000 relative to electronic signature (JO of march 14<sup>th</sup>, 2000, p. 3968), the storage support of the proof is not necessarily a paper support, but also an electronic support. This support answers to fidelity and perenity features explained by the Civil Code, as well as future requirements of proof integrity and attribution. It is why, in response to storage documentation problem and economic profitability, the professionals will now have access to other methods as Electronic Document Storage (EDS) (Comité Ialta 1998, Sénat 1999, CSO experts comptables 2000, Piette-Coudol 2000).

The substitution to traditional physical supports or the use of electronic origin documents in a such solution involves, principally when the projects have a vocation for a legal storage, the respect of some recommendations as those contained in the electronic storage guide and AFNOR norm NF Z42-013.

We are working in the context of a computer science environment achieving four functionalities: one certification of electronic transaction through a third certificate party, the secured storage through a third archive party, a tool for information researching and tracking of the different transactions.

According to the exchanges complexity in the context of a such system, an analyze tool for real time performances has to be developed. It is what we propose in the purpose of this article.

## 2. The storage and certification system

We are working on a document storage and certification system, based on the principles of secured electronic communication. This system answers particularly to recommendations published in the French decree no 2001-272 in date of march 30<sup>th</sup>, 2001, in application to the Civil Code article 1316-4 relative to electronic signature.

The legal document storage differs from classical electronic document storage (EDS). Indeed, the difficulty come from the fact that the system has to answer to both legal constraints and professional needs in terms of system response time, and absolute protection of documents integrity and confidentiality.

The third storage party will include in its services the elementary concepts of signed documents creation (including cryptography, signature appending, and creation of a unique print), and storage of them. The EDS has to intrinsically offer a solution to secured data storage.

The combination of cryptography and physic protection of storage documents does not offer sufficient integrity guarantees regarding to the law. This storage system allows to guarantee both integrity of legal documents, and required Quality of Service (QoS), in terms of storage speed and documents consultation.

The general structure of this system is the following one :

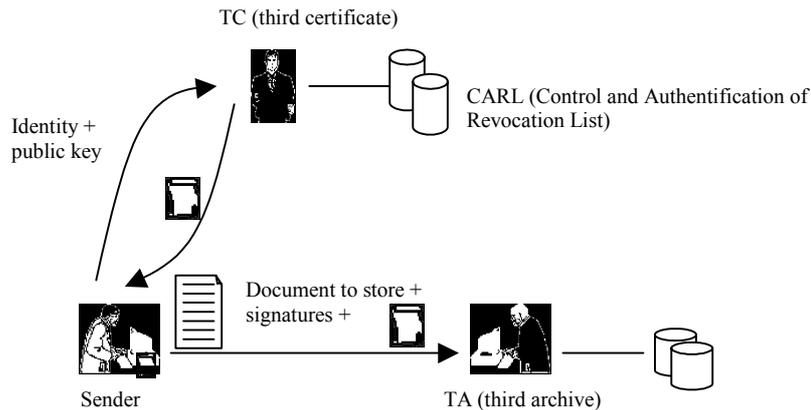


Figure 1: Scheme of storage system

This scheme involves three main physical entities: the customer, or user of the system ; the third certificate part, in charge of delivery and guarantee a certificate to the user ; the third storage party, in charge of conservation, and eventually delivery of storage documents.

The third archive party, according to the law, has to store in a secured manner all storage documents. A complex computer system is involved at this level. Different technological points have in that case to be solved, related to the arrangement, storage security, availability and working order guarantee of the system.

### 3. PERFORMANCE

A chain of storage as we described previously sets performance problems, notably in term of answer time.

In order to evaluate and improve reactivity of this system, it is therefore important to elaborate a performance indicators system fit to reactivity.

The development of a such system comes up against four difficulties:

- find some applicable indicators for reactivity is not a trivial task, because it is global and qualitative;
- use a formalism able to manipulate qualitative data, bounded to an expert appraisalment;
- use a graphic representation that allows the person responsible for the system to understand easily, to visualize globally and quickly the reactivity of the system on which he acts (criteria of cognitive ergonomics);
- to permit comparisons in time and space.

#### 3.1. Performance indicators

The performance indicator “measures the efficiency of all or a part [...] of a system (real or simulated), in relation with a norm, a workplan or an objective, determined and accepted in the setting of a enterprise strategy ” (AFGI 1992, Haurat 1998).

On a more general way, we can define the indicator of performance like all synthetic information (that underwent a treatment) on results of the target system transmitted to the level of management system. A performance indicator is designed and measured in order to inform the action model and allows him to start adapted processes: stabilization, improvement, etc.

#### 3.2. Typologies

There are at least two different typologies for performance indicators. The first concerns the nature of the performance, the second the actions caused by these indicators.

##### 3.2.1. Typology according to the nature of the performance

Used by nearly all enterprise actors, indicators of performance appears under very varied shapes. Simplifying, we can say that they can be quantitative, if they are expressed with the help of physical units (number of pieces produced per year, for example) or accountants (yearly business number, for example), or qualitative (quality of the aspect and the touch of a cloth, for example). They can be cardinal, that to be-to-say expressed by a number (x kilograms, for example.), or ordinal, expressed by a pre-order relation (more that, less that, at least, as much that). They can be connected to measurements of capacity, efficiency result or efficiency.

##### 3.2.2 Typology according to the action caused by the indicator

The indicator of performance can also be an indicator of result or follow-up indicator.

In the first case, the indicator measures an effect for a given action. It is a measure *a posteriori*, a report. If this measure is intended for an upper decision level that exercises a routine hierarchical control (the indicator is compared to a defined objective as a norm, and the decision is achieved without dialogue with the one that produces the measure), we speaks then of reporting indicator. If there is dialogue, that is common interpretation of the indicator, we speak then of management indicator (Lorino 1997; Berah 2000).

The indicator of follow-up (or process indicator), follows all the actions sequence associated to an objective realization. It allows to a local decision-maker to react (corrective action) before the final result is achieved (Lorino1997).

The only setting up of independent performance indicators doesn't necessarily bring to measure a global performance of the system. For that, it is necessary to design and run a performance indicators system.

### 3.3. Performance indicators system

#### 3.3.1. Definition

Indicators of performance have be established coherency with the global objective of the system. This objective is often expressed in very general way, using a linguistic, qualitative expression (the case of reactivity, for example). For that, in every enterprise, it is necessary to find the significant indicators contributing to the global performance, and then determine relations between global objective and significant indicators. This action materializes by setting up a system of performance indicators, in which one first defines an indicator that corresponds to the global performance expected for the target system, then of relations putting in consistency several meaningful indicators.

A performance indicators system can be defined as being an applicable indicator whole in relation to an objective, an

action and coherent between them. This system evolves during the time, when change objectives and actions.

In a performance indicators system, the different indicatory are joined between them to help:

- for vertical relations, active of most global to most detailed;
- for horizontal relations, describing causalities or interrelationships in the same way between indicators level.

#### 3.3.2. Developent

To elaborate a system of indicators, several stages are necessary:

- to define a global objective of reference in local objectives;
- to associate indicators of performance to objectives, global and local;
- to measure these local indicators with the help of the system of information;
- to synthesize the local indicators to measure the global performance of the system.

In the same way to the third stage, the choice of a formalism to express indicators and the indicators system must be made.

#### 3.3.3. Indicators system for reactivity

The indicators system for the global measure of the system reactivity is described by the figure 2.

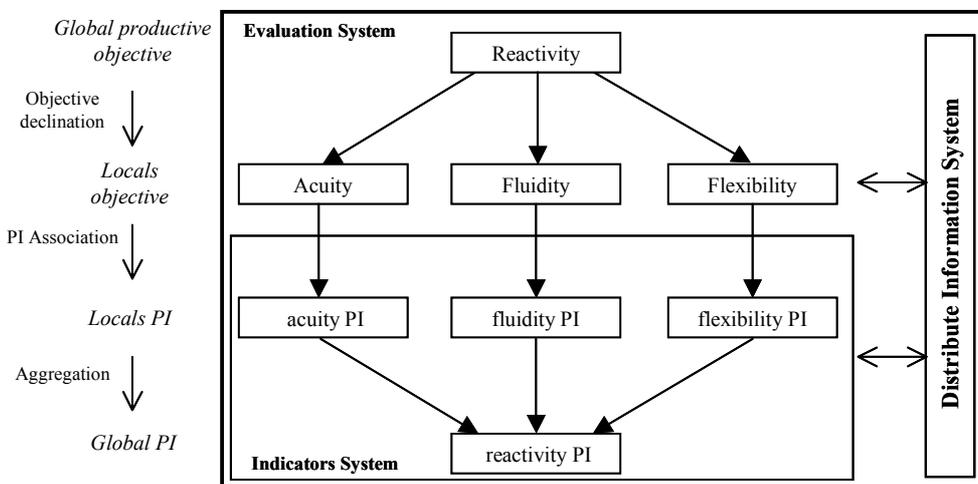


Figure 2: Reactivity: objective – indicators (Filipas 2001)

This figure presents the arborescent shape of the indicator system and its stages of development.

### 4. REACTIVITY DEFINITION

The reactivity of a system represents its capacity to perceive events, signs of the change of its external or internal environment (acuity), “to act in useful time, while using to

best flexibility of which it arranges” (Neubert 1997), and to assure the fluidity of its processes. Reactivity represents a shape of system behaviour efficiency.

The acuity concerns the quantitative and qualitative variations of the demand; flexibility, resources of the system and fluidity, the configuration and the speed of out-flow of information.

#### 4.1. Properties of reactivity

We retail the way of which this performance must fear himself of way more detailed and more concrete.

##### 4.1.1. A qualitative performance

We propose a particular approach of reactivity like a qualitative performance that is appreciate on relative way.

It is appreciated in relation to a temporal referential given by the system external environment. A measure of reactivity in absolute time, while using units of measure of the physical time, don't mean big thing if it is about comparing the respective reactivity of several systems. It is necessary to always appreciate reactivity to the look of the system that produces the referential: this system is component of the external environment of the system (for example the customer).

The reactivity is appreciated as relatively to the degree of flexibility of the component of the considered system, therefore of a structural property.

##### 4.1.2. A performance associated to system environment uncertainty

Reactivity is a property that becomes crucial when a system must face the uncertain events, of which at least the content, the date or effects cannot be anticipated. These events can have some negative consequences, if it is about "of risks" susceptible "to make a degraded working pass the system of a good working state" (Neubert 1997). They can have some positive consequences, if it is about opportunities to seize, for example, a new body of profession wanting to archive documents, a new idea of improvement for a system responsible.

According to us, three performances contribute to reactivity: the acuity, flexibility, fluidity.

#### 4.2. The acuity

The acuity designates behaviour capacity of a piloting system to discern the least variations, quantitative or qualitative, of its external environment.

We can approach this performance while measuring the number of relative measures to variations of state of the external environment of a system so that this one detects a tendency and acted. Concerning the infrastructure to public key, these variations concern the fluctuations of the demand, but as its qualitative changes (for example, interest of consumers for a new product launched by the competition, introducing a strong authentication).

The acuity is measured in entrance of the system (figure 3). It designates the delay between the beginning of perception of a need and the beginning of the card production (equation (1)). An enterprise must answer in the briefest delay to a customer's needs. It is as the time of reaction that exists between the moment or the demand has felt, and the

moment where the enterprise takes conscience of the existence of this demand.

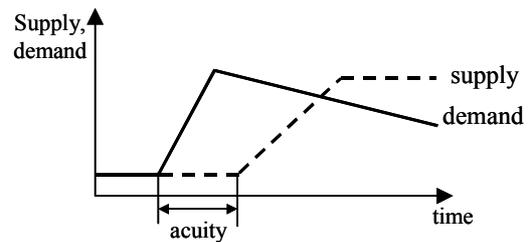


Figure 3: The acuity

$$A = T_s - T_p \quad (1)$$

with:

- $A$ , system acuity;
- $T_p$ , perception time of demand changing;
- $T_s$ , card production start time;

#### 4.3. Flexibility

Very numerous works, emanating varied disciplines, tried to define what is flexibility. We define the flexibility of the following way however:

Flexibility designates the capacity of components of the system to absorb some quantitative or qualitative changes. Components took in account can be resources or processes.

Most authors agree to distinguish the static flexibility of the dynamic flexibility. In the first case, resources are over-dimensioning. The static flexibility is a matter for the existence of a potential of resources and concern the capacity of reaction of a piloting system facing an evolution unforeseeable of the environment. It is a flexibility "instantaneous, potential and present at all times" (Neubert 1997). The dynamic flexibility represents the capacity of the system to react continuously in the time to the uncertain variations of the environment. It supposes a qualitative adaptation capacity: to reorganize the structure of the system, to facilitate cooperation, etc.

Authors also agree to consider the flexibility of processes, and no only of resources. A supple process can either be shortened reversible. In the last case, It is possible to reiterate to least cost on the previous stages when a problem appears (reversibility in relation to the past, for example: development of a software). One can generate a big number of possible future states (reversibility in relation to the future, for example: re-write a software).

The content of the first two indicators of reactivity (the acuity and flexibility) having been specified, it now agrees to specify the last of them: fluidity.

#### 4.4. Fluidity

Of all contributively factors of reactivity, fluidity is simplest to define and to measure. It is about measuring, within the system, the speed of out-flow of the informational flow.

One designates by fluidity the capacity of a system to assure the out-flow of information flow without stop or slowing.

When we try to measure the fluidity of a computer system, it is about, for an  $i$  card given, to use a ratio of type:

$$f_i = \frac{T_{ALE} + T_{AE} + T_{AC}}{TC_i}, \quad i \in \{1, n\}$$

with:

- $f_i$  fluidity associated to an  $i$  map ;
- $T_{ALE}$  necessary time for local authority of registration to collect the information of demand of obtaining of an  $i$  certificate;
- $T_{AE}$  necessary time for the authority of registration to collect the demand of an  $i$  map from the agency of registration, to analyse it, to adjust it and to take a decision;
- $T_{AC}$ , necessary time for the authority of certification to take a decision (acceptance or refusal) relative to the request made by a customer;
- $TC_i$  total cycle time (or delay between the demand of the map and the end of his/her/its manufacture) of  $i$ . (knowing that for a legal problem, the third certificate cannot keep the key deprived more than one hour).

#### 5. SUMMARY AND CONCLUSIONS

In this article we proposed a system of reactivity analysis for an architecture to public key.

We first of all defined an performance indicator like all synthetic information on results of the system targets ascents at the level of the piloting system. It was necessary to find indicators then to the three performances that contribute to reactivity: the acuity, flexibility and fluidity;

- the acuity has been measured from the delay that flows out before the system of piloting doesn't modify his/her/its model of demand;
- flexibility is the capacity of the system to react continuously in the time to the uncertain variations of the environment. It is about reorganizing the structure of the system, to facilitate cooperation, etc.;
- fluidity has been measured from necessary times for every authority of certification to take one decision with regard to the electronic map.

We also presented a methodology of development of an indicator system for reactivity. This system must be integrated in the system of assessment of the architecture to public key.

#### 6. REFERENCES

- AFGI Association Française de Gestion Industrielle. 1992. Evaluer pour évoluer, les indicateurs de performance au service du pilotage industriel. Editions AFGI, Paris.
- Comité Ialta. novembre 1998. "Livre blanc : les nouveaux tiers de confiance impliqués dans les échanges électroniques".
- Conseil Supérieur de l'Ordre des Experts Comptables. juillet 2000. "Guide de l'archivage électronique sécurisé", version 5.
- Berrah, L ; G. Mauris; A. Haurat and L. Foulloy. 2000. "Global vision and performance indicators for an industrial improvement approach", Computers in Industry, Vol. 43, No. 3, pp. 211-226.
- Filipaş I., "Définition et mise en œuvre d'un pilotage adapté à la réactivité industrielle : Application à un atelier de fabrication de coiffes de sièges d'automobile", Thèse de doctorat en Automatique et Informatique de l'Université de Franche-Comté, 2001.
- Haurat A. 1998. " Approche systémique de la performance", *Université d'été du Pole productique Rhône-Alpes (PPRA), La performance industrielle*, 1-4 septembre, Annecy.
- Lorino P. 1997 *Méthodes et pratiques de la performance. Le guide de pilotage*, Éditions d'Organisation, Paris.
- Neubert G. 1997. "Contribution à la spécification d'un pilotage proactif et réactif pour la gestion des aléas", thèse de doctorat en Productique, Institut National de Sciences Appliquées de Lyon.
- Piette-Coudol Thierry. 2000. Président du comité Ialta-France, "Droit de l'écrit électronique".
- Sénat, service des affaires européennes. décembre 1999 "La signature électronique", n° LC 67.

#### 7. BIOGRAPHY

Ioana Filipas received her Ph.D in Automatic and Computer Science at the University of Franche-Comté (September, 2001) and is an engineer from the Timisoara Polytechnic University (Romania). Her work deal with the definition and implementation of an adapted piloting system for enhancing industrial reactivity.



**SIMULATION  
IN  
NEURAL  
NETWORKS  
AND  
ARTIFICIAL  
INTELLIGENCE**



# OBJECT ORIENTED AND NEURAL NETWORKS SIMULATION : AN APPLICATION TO THE STUDY OF WATERSHED

Frederic Chiari, Marielle Delhom and Jean-François Santucci  
University of Corsica - UMR CNRS 6134  
BP 52 – F-20250 Corte - Corsica  
E-mail : [chiari@univ-corse.fr](mailto:chiari@univ-corse.fr)

## KEYWORDS

Hydrology, Forecasting, Hybrid Simulation, Neural networks, Object-oriented Simulation, Rainfall - Runoff

## ABSTRACT

In order to study the hydrologic behavior of a watershed we have performed different kinds of simulation.

In a first step, we have developed an original approach based on object oriented modeling and simulation. So we have proposed a simulation architecture, a software has been realized and good results has been obtained. This approach is deterministic and so we need to know with accuracy the behavior of the watershed, that is to say the physical parameters of the system. Unfortunately, in the case of study of natural systems these parameters are not always available.

So, we have chosen to use Artificial Neural Networks (ANNs) for their learning features that seem to be particularly adapted to our problem. We have studied different types of networks in order to find the more suitable for the study of natural systems and we have selected the following : back propagation, RCC, Radial Basis. We have performed experiments on real data and some results are presented in this paper.

## INTRODUCTION

The paper is organized as follows:

In Section 1, we present some general notions concerning watersheds and the problems arisen from the modeling and the simulation of the hydrologic behavior of a watershed. Section 2 presents the two facets of our approach, the Object Oriented Modeling and Simulation concepts and the Artificial Neural Networks concepts. The section 3 explains our experimentation and we give in the section 4 the results we obtained. Finally, in section 5, we present the current status of our work as well as its limitations and our future investigations.

## BEHAVIOR OF A WATERSHED

We present, in this section, some notions concerning watersheds and problems which are set up by the modeling and the simulation of the hydrologic behavior of a watershed.

Figure 1 gives a representation of a watershed : a watershed is a geographic space which receives precipitation and restores a part of these precipitation in the form of discharges at a single point (for instance a river).

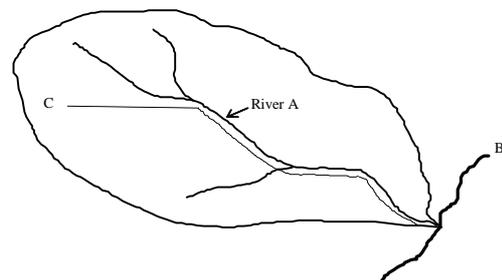


Figure 1 : Watershed Representation

The study of the hydrologic behavior of a watershed can be realized for the following reasons:

- understanding the behavior of the system in order to forecast a river discharge according to the observed precipitation on the watershed,
- obtaining estimates of average performance measures in order to assist the management of an hydraulic barrage built on the river belonging to the watershed.

Generally, watersheds have important surfaces. This often explains that they have diversified landscapes (such as mountains, plains) with different altitudes, vegetation. This diversity involves, depending on the watershed places, different hydrologic behaviors.

The Spectral Analysis (Dautray and Lions 1993) has allowed to model the hydrologic behavior of a watershed. This approach basically consists in representing the behavior of a natural system using mathematical and physical laws. The watershed is considered globally, without integrating its complexity. Besides, the simulation model realized from the Spectral Analysis, does not take into account some kinds of phenomena, such as the influence of the snow; this one is merged with the rain, and this is not justified because the snow is transformed in water according to the temperature (this transformation is not instantaneous).

In order to study the hydrologic behavior of a watershed, and faced with this type of limits, we have chosen to develop an original approach based on an hybrid model integrating object oriented simulation and neural networks simulation.

## SIMULATION METHODOLOGIES

In this section we present the simulation methodologies used to define our models of the watershed. We have in a first time developed an original approach based on object oriented modeling and simulation ; but this deterministic approach needs to know with accuracy the behavior of the system, and this behavior is not always available. Because of this limits, we have chosen to direct our researches in the ANNs domain.

### Object Oriented Modeling and Simulation

In this paragraph, we briefly present our object oriented modeling and simulation approach (Delhom et al. 1995, Delhom et al. 1997). The originality of our approach is :

- to use three types of hierarchies (abstraction hierarchy, description hierarchy and time hierarchy) in order to model the studied system; this kind of modeling allows to take into account the complexity of the real system;
- to use the Object Oriented Programming concepts (Meyer 1988) in order to implement and simulate the studied system, and then to easily take into account any modifications.

In order to take into account the complexity of a system, it is often necessary to use different types of hierarchies. We have defined three important types of hierarchies : the description hierarchy, the abstraction hierarchy and the time hierarchy. The definitions of these hierarchies are based on different works:

- The C. Oussalah's work (Oussalah 1988) which introduces the abstraction hierarchy and views notions.
- The B.P. Zeigler's work which proposes the DEVS formalism (Zeigler 1976, Zeigler 1984) have been used to define the description hierarchy.
- The J. Euzena's work which proposes the notion of granularity (Euzenat 1994) has allowed us to define a time hierarchy.

The integration of these concepts has allowed us to define a simulator architecture. Besides, a simulation software has been realized. Particularly, it allows to process independently the modeling and the simulation parts. So, this makes easier the definition and the modification of the model. Owing to this explicit separation between the modeling and the simulation aspects, several models can be implemented and studied without any modification of the simulation software.

### Neural Networks Paradigm

This section concerns a presentation of the Artificial Neural Networks Paradigm. Actually, in the case of the study of natural systems, our modeling and simulation approach presents some limits. So we have chosen to

complete our approach with the use of the Neural Networks Paradigm.

In the first part we present basics of Neural Networks. Then, we present networks which are particularly adapted to the simulation of natural systems.

#### Basic Idea

A neural network is a computational method inspired by studies of the brain and nervous systems in biological organisms (Rosenblatt 1962).

A neural network is composed of interconnected processing elements (neurons) working in unison to solve a specific problem. In (Hecht-Nielsen 1989), R. Hecht-Nielsen gives the following definition of a neural network : a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external input.

Neural networks are organized in layers; these layers are made up of a number of interconnected nodes which contain an activation function. Patterns are presented to the network via the input layer, which communicates to hidden layers where the processing is done using weighted connections. Then, the hidden layers transmit the answer to the output layer (see figure 2).

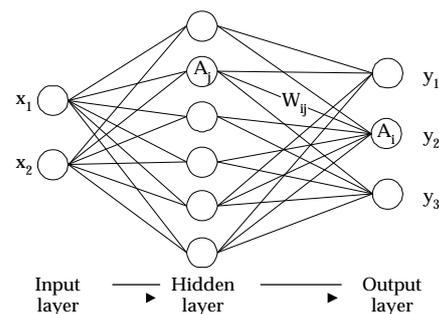


Figure 2 : Organization of a Neural Network

#### Algorithms

We present here the basics of the algorithm we used, the famous Back-Propagation algorithm (Jodouin 1994), the Recurrent Cascade Correlation algorithm (Fahlman and Lebiere 1990) and Radial Basis Functions networks (Orr 1996).

##### ◆ Back propagation

A pattern ( a couple of input values and output values) from the training data set is chosen randomly. The input pattern is presented to the network at the input layer by assigning each value of the pattern to one input neuron. Inputs are then propagated by the network until they reach output layer. I.e., for each neuron an activation  $a_i$  is computed :

$$a_i = F(\sum_j O_j W_{ij}) \text{ where}$$

$O_j$  is the output of neuron  $j$  on the precedent layer,

$W_{ij}$  is the weight of the connection from neuron  $j$  to neuron  $i$ .

F is the transfer function (activation function) of the neuron i (usually the Sigmoid function  $f(x) = 1/(1+e^{-x})$ ).

The output pattern that the network produces for each input pattern is compared to the expected value pattern. An error value is computed as follows :

$$E = \sum_i (O_i - T_i)^2 \text{ where}$$

E : is the error value that corresponds to the pattern presented to the network,

$O_i$  is the output value of neuron i on the output layer,

$T_i$  is the i'th value on the target output.

If the error value is not near zero, weights of the connections have to be changed in order to reduce this error. Each weight is either increased by some fraction or decreased back-propagating the computed error. The mathematical formula used by this algorithm is known as the Delta rule :

$$\Delta W_{ij} = \eta \delta_i O_j \text{ where}$$

$\Delta W_{ij}$  is the amount by which the weight  $W_{ij}$  should change correspondingly to training pattern pair

$\eta$  is the learning rate

$\delta_i$  is the error on the output of unit i on layer.

The computation of it's value depends on the type of the neuron. If the neuron is an output unit then the error is :  $\delta_i = F'(a_i) (T_i - O_i)$  else (the neuron is an hidden neuron ),  $\delta_i = F'(a_i) \sum_k \delta_k W_{ik}$  where neurons k are the successors of neuron i.

#### ◆ Recurrent Cascade Correlation

Recurrent Cascade-Correlation (RCC) is a recurrent version of Cascade-Correlation and can be used to train recurrent neural nets .

Recurrent nets have some features that distinguish them from normal neural networks. For example they can be used to represent time implicitly rather than explicitly.

One of the most commonly known architectures of recurrent neural nets is the Elman (Elman 1990) model, which assumes that the network operates in discrete time steps. The outputs of the network's hidden units at a t time are fed back for use as additional network inputs at time t+1. To store the outputs of the hidden units Elman introduced context units, which represent a kind of short-term memory. To integrate the Elman model into the cascade architecture some changes are necessary: The hidden units' values are no longer fed back to all other hidden units. Instead every hidden unit has only one self recurrent link. This self recurrent link is trained along with the candidate unit's other input weights to maximize the correlation. When the candidate unit is added to the active network as hidden unit, the recurrent link is frozen with all other links.

#### ◆ Radial Basis Functions

The principle of radial basis functions derives from the theory of functional interpolation. A function f is approximated with a linear combination of radial functions.

$$f(\vec{x}) = \sum_{i=1}^K c_i h_i(|\vec{x} - \vec{t}_i|)$$

The method of radial basis functions can easily be represented by a three layers feed-forward neural network : an input layer, an hidden layer which is composed by kernel functions (generally gaussian) and a single output neuron. The input layer consists of n units which represent the elements of the vector  $\vec{x}$ . The links between input and hidden layers contain the elements of the vectors  $\vec{t}_i$ . The hidden units compute the Euclidian distance between the input pattern and the vector which is represented by the links leading to this unit. The activation of the hidden units is computed by applying the Euclidian distance to the function h. Figure 3 shows the architecture of the special form of hidden units.

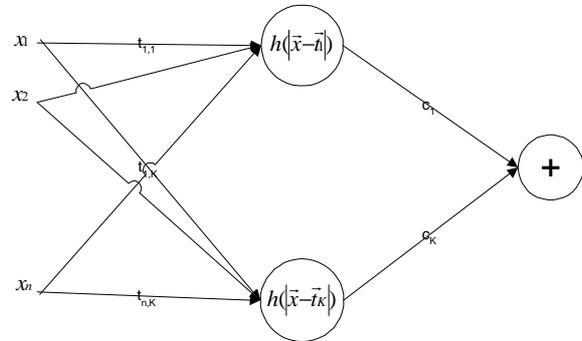


Figure 3: The RBF network topology

## IMPLEMENTATION AND EXPERIMENTATION

In this section, we present how we modeled the hydrologic behaviour of a watershed (Chiari et al. 1998). These models are described using concepts presented in section II.

### Object Oriented Modeling and Simulation

We have first defined a basic model of the watershed, integrating its main characteristics. Since our approach is evolutionary, we have been able to easily improve this model by taking into account the effects of the seasons and the altitude on the behaviour of a watershed. Finally, in order to take into account the snow, we have defined a third model of a watershed (Delhom et al. 1995, Delhom et al. 1997).

### Artificial Neural Networks

The main interest of ANNs is that one can work with any a priori knowledge of the studied system (Labbi 1993, Szilas 1995). However the structure used for our experiments is based on several past works. The general topology of the experimented ANNs follows :

- The inputs of the networks are pluviometric data and two sets of temperatures of the day and over the last 29 days (collected at two points of the

watershed) and discharges over the last 29 days. So we have 119 inputs which are 30 rainfall data (current day to d - 29), 60 temperatures (current day to d - 29) and 29 discharges (d-1 to d-29)

- There is only one output of the ANNs, the expected discharge of the current day.

This structure has been chosen thanks to the work we have made on physical and DEVS modeling of watersheds.

For our experiments, we use a learning set of 730 patterns ( from 01/01/1984 to 12/31/1985), a validation set (365 patterns for the year 1983) and a test set (365 patterns for the year 1986).

We performed all the experiments using the Stuttgart Neural Network Simulator (SNNS) (Zell et al. 1995).

## RESULTS AND COMPARISONS

We present in this section four graphics. On every graph, we plot the real discharge of the watershed and the computed one for the year 1986, the test data pattern. Then the correlation between the real and the computed discharge is studied through two statistics (see Figure 8), the Pearson's correlation coefficient and the root mean square error.

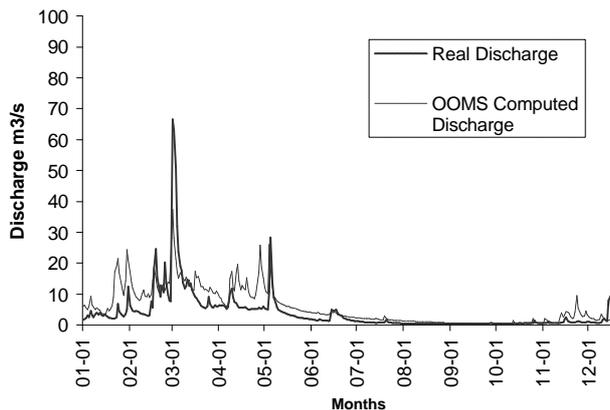


Figure 4 : Real and O O Computed Discharges

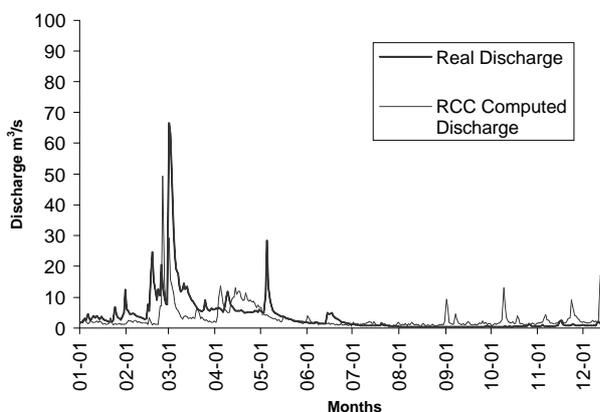


Figure 5 : Real and RCC Computed Discharges

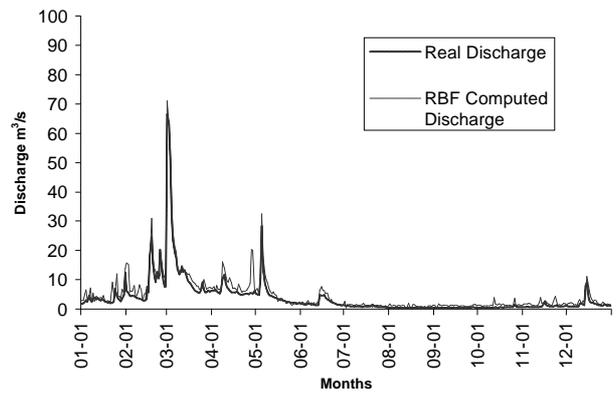


Figure 6 : Real and RBF Computed Discharges

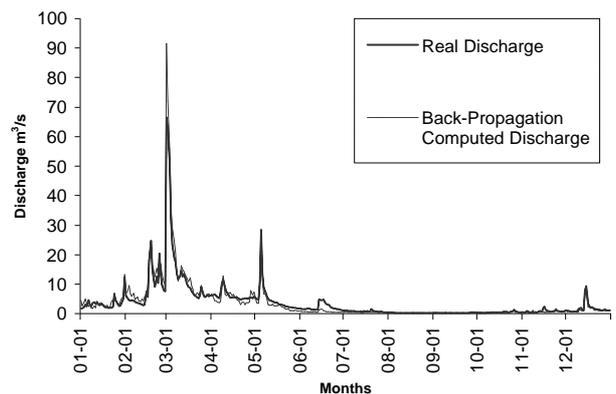


Figure 7 : Real and Back-Prop Computed Discharges

These different results show the efficiency of our approach. Indeed, the main problem involved by the Object Oriented Modeling and Simulation was the representation in the beginning of the year of the snow (see Figure 4). Neural Networks, particularly the back-propagation network, allow to obtain better results, the curves follow the same variations and the distance between real discharges and estimated discharges are small (see Figure 7).

Table 1 presents, for each models the correlation between expected (real) results and computed ones. We can see on this table that the smallest root mean square errors are obtained with the back-propagation and radial basis function networks. The Pearson's correlation coefficient point out the correlation between the two curves (estimated and real discharges), the nearer to 1 the coefficient is, stronger the correlation is.

Nevertheless, some improvements are necessary : particularly, we have to optimize the pre-process of the input patterns.

	OOMS	RCC	RBF	BackProp
<b>Pearson correl.</b>	0,71 moderate	0,51 moderate	0,97 strong	0,98 strong
<b>RMSE</b>	5,26	5,87	2,18	1,90

Table 1 : Pearson's Correlation and root mean square error

## CONCLUSION AND PERSPECTIVES

After the conception of an Object Oriented simulator, we investigated applications of neural networks to non-linear series prediction. Results we obtained are quite encouraging but our aim isn't to make an Artificial Neural Network based simulator. In a near future we plan to use connectionism techniques that we have experimented to improve our object oriented simulator. For now, the outputs of the OO simulation don't produce any change, neither on the model part nor the simulation algorithm. Indeed we intend to apply error correction rules to the simulator using the DEVS formalism. On an other hand, as we are faced to important amount of data, we are working on the integration of Geographical Information Systems (GIS) to our simulator so as to make our software user-friendly.

Accordingly, the purpose of our future work will concern the definition of a hybrid framework mainly based on Object Oriented concepts integrating ANNs features and GIS enhancements.

Then, after have tested our simulator architecture on watersheds, we will generalize it in order to model and simulate other kinds of systems.

## REFERENCES

- Amy B. 1996. Recherches et perspectives dans le domaine des réseaux connexionistes, Rapport interne du laboratoire Leibniz-Imag.
- Carling A. 1992. Introducing neural networks, Sigma Press, New York.
- Chiari F., Delhom M., Ciccoli F.X., Santucci J.F. 1998. Object oriented and neural networks modelling for the simulation of natural systems, Fourth International Conference on Computer Science & Informatics, Durham, North Carolina, October 23-28, pp. 170-173.
- Chiari F., Delhom M., Santucci J.F., Filippi JB 2000. Prediction of the hydrologic behavior of a watershed using artificial neural networks and geographic information systems. 2000 IEEE International Conference on Systems, Man and Cybernetics, pp. 382, 386.
- Chiari F., Delhom M., Santucci J.F. 2002. An hybrid methodology for the modeling and the simulation of natural systems. SAMS, Vol 42(2), pp. 269-287.
- Dautray R, Lions JL 1990. Mathematical Analysis and Numerical Methods for Science and Technology: Functional and Variational Methods, Springer Verlag.
- Davalo E., Naim P. 1993. Des réseaux de neurones, Eyrolles.
- Delhom M., Bisgambiglia P.A., Santucci J.F. 1995. Modeling and Simulation of Discrete Event Systems : the study of the hydrologic behavior of a catchment basin. IEEE international conference on Systems, Man and Cybernetics, Vancouver, Canada, October 22-25, Vol. 5, pp. 4191-4195.
- Delhom M., Santucci J.F., Bisgambiglia P.A., Aiello A. 1997. Object Oriented Modelling and Simulation : application to natural systems. SCS international Western Multiconference on Computer Simulation, Phoenix, Arizona, January 11-15, pp. 115-120.
- Dorffner G. 1996. Neural networks for time series processing, Neural Network World, 6(4)447-468.
- Elman, J.L. 1990. Finding structure in time. Cognitive Science 14: pp. 179-211.
- Euzenat J. 1994. Granularité dans les représentations spatio-temporelles, INRIA research report n°2242, 62 p.
- Fahlman S.E., Lebiere C. 1990. The Cascade-Correlation Learning Architecture, Neural information processing systems, D.S. Touretzky ed, Vol. 2, pp. 524-532.
- Hebb D. 1949. The organisation of behavior, Wiley.
- Hecht-Nielsen R. 1989. Neural Network primer: part I, AI Expert, Feb.
- Jodouin J.F. 1994. Les réseaux neuromimétiques, Hermes.
- Labbi A. 1993. Sur l'approximation et les systèmes dynamiques dans les réseaux neuronaux, Thesis, INPG, Grenoble, France.
- Lorrai M., Sechi G.M. 1995. Neural Nets for Modelling Rainfall-Runoff. Water Resources Management, vol. 9, 4: pp. 299-313.
- Meyer B. 1988. Object Oriented Software Construction, Prentice Hall.
- Orr M. 1996. Introduction to Radial Basis Function Networks, technical report.
- Oussalah C. 1988. Modèles hiérarchisés/multi-vues pour le support de raisonnement dans les domaines techniques, PhD thesis, Univ. Aix-Marseille III.
- Rosenblatt F. 1962. Principles of neurodynamics, Academic Press, New York.
- Szilas N. 1995. Apprentissage dans les réseaux récurrents pour la modélisation mécanique et étude de leurs interactions avec l'environnement, Thesis, INPG, Grenoble, France.
- Zeigler B.P. 1976. Theory of Modelling and Simulation. Wiley, New-York.
- Zeigler B.P. 1984. Multifaceted Modelling and Discrete Event Simulation. Academic Press.
- Zell A. et al. 1995. SNNs User Manual, Version 4.1 University of Stuttgart Institute For Parallel and Distributed High Performance Systems.

## AUTHORS BIOGRAPHIES

**Frederic Chiari** is 32 years old. He is computer engineer since 1996, he's also PhD student at the University of Corsica. His interests are TCP/IP networks, Geographic Information Systems and Natural Systems Simulation.

**Marielle Delhom** is associated professor in computer science at the University of Corsica. She is researcher at the UMR CNRS 6134. She is interested in the Objected Oriented Modelling and Simulation, Neural Nets Modelling applied to the study of natural systems.

**Jean-François Santucci** is Professor in computer science at the University of Corsica. He is co-director of the UMR CNRS 6134. His research activities concern the Object Oriented Modelling, Simulation techniques applied to complexes systems and Digital Circuits Testing techniques.

# NEURAL NETWORK TECHNIQUES IN ATM CONGESTION CONTROL

Arif Al-Hammadi.  
Etisalat College of Engineering  
P.O.Box 980, Sharjah, UAE  
E-mail:arif@ece.ac.ae,

John Schormans  
Department of Electronics Engineering,  
Queen Mary University of London,  
London E1 4NS, United Kingdom  
j.a.schormans@qmul.ac.uk,

## KEYWORDS

Neural Networks, ATM Simulation, Congestion Control.

## ABSTRACT

In this paper we show the simulation results of Neural-Network techniques for congestion control in a three buffers ATM switch with time priorities. Tests are carried to compare different NN modules. This paper assesses the importance of the higher priority buffers in deciding the congestion in the lower priority ones and shows that in a prioritised switch it is necessary to monitor the buffer to be controlled as well as buffers with higher priorities. It also shows that NN scheme has a better CLR, delay and utilisation when compared to the conventional CB scheme.

## INTRODUCTION

Multimedia information services are growing rapidly, increasing the demand for broadband integrated services digital networks (B-ISDN) that will provide high-speed communication channels suitable for transporting multimedia signals efficiently. Asynchronous Transfer Mode (ATM) is the transfer mode of choice for B-ISDN (Onvural. 1994). Performance prediction and quantitative analysis of these networks have become extremely important in view of their ever expanding usage, and the complexity of their functioning.

Machine Intelligence has been widely used in solving engineering problems (Habib 1995.). The strength of Machine Intelligence comes from the adaptive, learning and inference features that can alleviate the shortcomings of conventional algorithmic approaches. Machine intelligence techniques can be classified as computational and artificial intelligence techniques. NN and fuzzy logic are examples of computational intelligence, whereas expert systems are examples of artificial intelligence (Habib 1995). Computational intelligence is also referred to as "soft computing" in (Zadeh 1996), where it is defined as being a consortium of computing methodologies which collectively provide a foundation for the conception, design and deployment of intelligent systems.

It is clear from the ATM forum (ATM forum 1994) that future ATM traffic will be classified into at least 4 categories (may be more), Constant Bit Rate (CBR), Variable Bit Rate (VBR), Unspecified Bit Rate (UBR) and Available Bit Rate (ABR): These traffic classes were proposed to guarantee different Quality of Service (QoS)

for different traffic types. The ABR and UBR service categories will not have any bandwidth guarantees (except a Minimum Allowed Bandwidth allocated for ABR, which can be zero). This means that ABR and UBR as well as non-real time VBR categories will act as low priority traffic (or best effort). The choice of implementing a certain scheme was left by the ATM forum to the manufacturer. In the schemes proposed in (ATM forum 1995) such as EPRCA, the ER value was set by the switch and represents the fair share allocated to the source out of the total link bandwidth available. However, in this scheme the reduction factor RDF and the increase factor AIR were constant or static. The aim in this paper is to show that by using an explicit scheme, such as NN, in an ATM switch with time priorities, better network performance can be achieved when compared with the static scheme (referred to as the CB scheme in this paper).

Neural Networks (NN) has been recommended by many researchers to provide an alternative approach to the conventional traffic control approaches for ATM networks. The advantage is in their learning and adaptive capabilities that can be used to construct adaptive control algorithms for optimal allocation of resources. A limitation of the previous work on the NN techniques in (Liu and Douligeris. 1997) is that they were implemented in a single buffer with a fixed service rate scenario. However, most of the feedback mechanisms are implemented on best effort traffic, which has zero, or minimum allocated bandwidth (such as in Available Bit Rate traffic). In (Al-Hammadi and Schormans, 1998) we have proposed a NN scheme in a prioritised ATM switch with two priorities, high priority and low priority. In this paper we test the NN performance in an ATM switch with three priorities, High Priority-1 (HP-1), High Priority-2 (HP-2) and Low Priority (LP), the same procedures of creating the training file and training the NN in the two priorities case were used. This paper also proposes a modification to the NN scheme in (Al-Hammadi and Schormans, 1998) and tests the effect of monitoring the upper priority buffers on controlling congestion in the lower priority ones when using three buffers. This is a more realistic case since the number of classification of traffic in ATM is at least 4.

## TRAFFIC SOURCES

Two MPEG video traces were used in testing our proposed scheme. The MPEG trace-1 data set, used for training and testing, can be found at the Bellcore ftp site (Garrett and Fernandez 1994). Trace-1 represents a variable bit rate video trace from the movie "Star Wars". The MPEG

trace-2, which is used for testing can be found in (Uniwuerzburg). Trace-2 represents a variable bit rate video trace from a football match.

## CB SCHEME

Comparing with congestion control schemes that are based on detecting violation of a threshold has been used to test previous work on NN techniques in congestion control (Liu and Douligeris. 1997). In this paper we compare the CB scheme proposed in (Newman 1993) to our proposed scheme. In the CB scheme, when the buffer occupancy reaches a certain predefined threshold  $Q_{th}$  (e.g. 50% or 60% of the total buffer size  $B$ ), the sources feeding the buffer are reduced to 50% of their current transmission rate. If further reductions are required the transmitter will ignore reduction requests until one cell at least is transmitted. Successive reductions will cause a source to reduce its cell transmission rate to: 50%, 25%, 12% and 6% after which the source stops transmission. A transmission recovery mechanism is built into each source. If no backward control cells are received, the source will be restored to its previous level, until it restores its original peak value.

## NN SYSTEM DESCRIPTION AND NN TRAINING

In order to create the NN training files an MPEG video trace-1 was used. The training file was created using a trace-1 source feeding the HP-1, HP-2 and LP with different starting times, and a testing file using trace-2 feeding the LP while the HP-1 and HP-2 were fed with trace-1. The three buffers share the same server, however buffer-1 has full priority over buffer-2 and buffer-2 has full priority over buffer-3. This means that any cells found in buffer-1 will be served before cells in buffer-2 and buffer-3. Time is divided into intervals with the same length  $T$  ( $T$  was taken to be 8.3 ms as in (Liu and Douligeris. 1997) and (Al-Hammadi and Schormans, 1998)). Three NN models were tested. In Model 1, a three layers NN (6-8-1) was fed with 5 inputs representing the number of cell arrivals  $A(i-1)$ ,  $A(i-2)$ ,  $A(i-3)$ ,  $A(i-4)$ ,  $A(i-5)$ ,  $A(i-6)$  measured in the previous time periods  $T(i-1)$ ,  $T(i-2)$ ,  $T(i-3)$ ,  $T(i-4)$ ,  $T(i-5)$ ,  $T(i-6)$  respectively, from the source feeding the LP buffer. The inputs were normalised between 0-1 by dividing them by the peak number of arrivals to the LP buffer. In Model 2, three layers NN (12-8-1) is fed with the 6 inputs from the LP buffer (as in Model 1), as well as 6 inputs from the HP-2 buffer. In Model 3, a three layers NN (18-8-1) is fed with the 6 inputs from the LP buffer (as in Model 1), as well as 6 inputs from the HP-1 and 6 inputs from HP-2 buffer. All models were trained to predict  $O$ , where  $O$  is defined as the

ratio of the number of cells discarded at the LP buffer to the number of cells arrived to the LP buffer in the time period  $T(i+2)$ . The NN was trained to predict 2-cycle time periods ahead to include the propagation time period ( $\Delta$ ) from the buffer to the source which is equal to the sampling time period  $T$  for simplicity. In (Liu and Douligeris. 1997) the NN output  $O$  was fed back to regulate each source from its old rate ( $R$ ) to its new rate ( $r$ ), as given in equation (1):

$$r = (1 - O) \times R \quad (1)$$

Training files representing the inputs and outputs described above were created when no regulation of the sources was considered. Each file contains 2700 samples representing 0.31% of the whole trace-1 file. The NN used in this letter was the three layer Feed-forward NN with Error Back-Propagation learning algorithm. The Mean Square Error (MSE) against the number of passes through the training file for both NN models is shown in graph 1. The results in graph 1 show that Model-3 gives better MSE for both testing and training files when compared to Model-2 and Model-1. This result shows that as the number of priority levels increases the need to monitor the higher priority buffers becomes more important if the NN scheme is to be used in a prioritised ATM switch.

The NN scheme uses one NN to detect the occurrence of the congestion status as well as finding the explicit amount of reduction required. This scheme is referred to as, the NN-1 scheme. In order to achieve better results, the task was split between two NN's as shown in figure 1. The first NN-1 is trained to predict the congestion status (congestion or no congestion), and the second, NN-2, is trained to predict the explicit amount of reduction required. The later scheme is referred to as, the 2-NN scheme. In graph 1, the 2-NN scheme shows slightly better learning performance when compared to the 1-NN scheme.

Since monitoring the arrival rate might increase the complexity in the switch, in the next test, the buffer changes (i.e. amount of increase or decrease in buffer length in the time period  $T$ ) were used instead of the arrival rate as an input to the NN. From graph 2, a number of observations were noted. For the training error curves, Model-3 shows a clear advantage over Model-1 however; Model-3 has only a small advantage over Model-2. This is because most of the cells arriving to the HP-1 get served and hence no fluctuation in the buffer length is observed. So monitoring the HP-1 buffer does not add any advantage to the NN learning. It was also noted that the generalisation (i.e. the testing error curve) for Model-2 was better than that of Model-3. Graph 2 also shows that the NN-2 scheme gives better generalisation performance when compared to NN-1 scheme

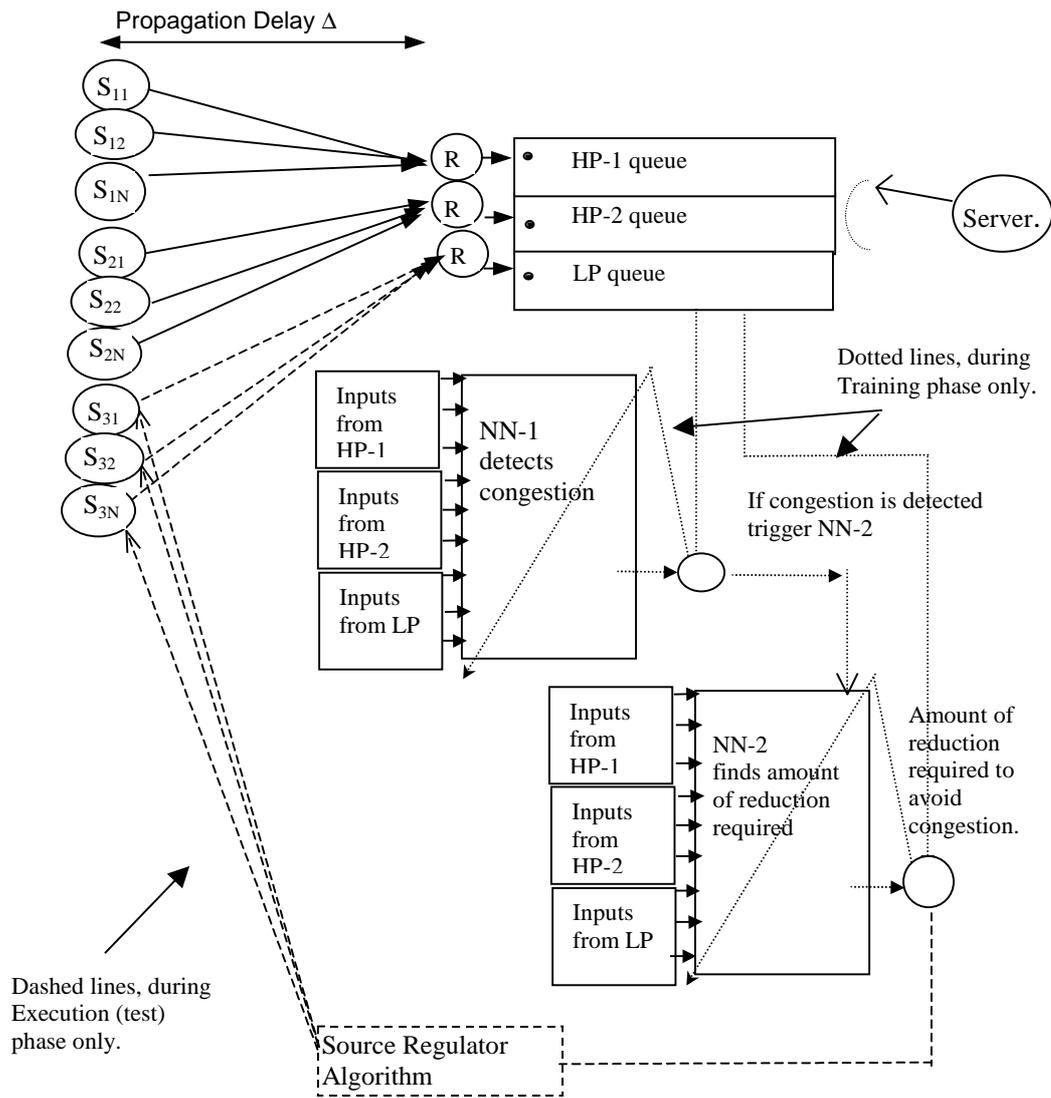
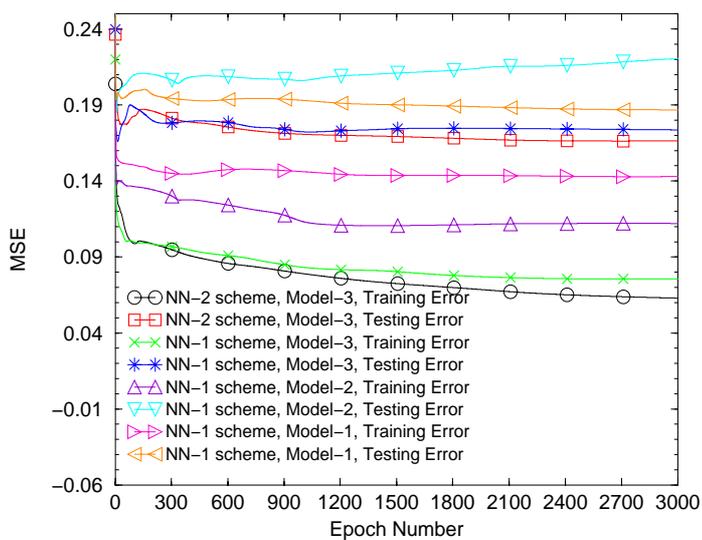
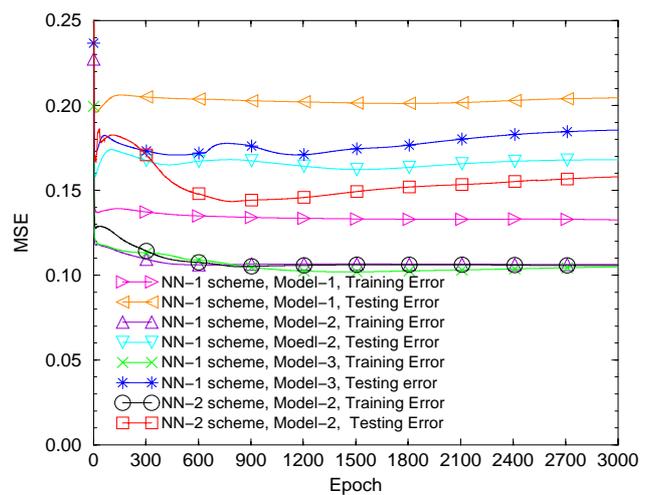


Figure 1, Feedback Congestion Control using Two NNs Scheme.



Graph 1, MSE Comparison for Different NN Models (monitoring cell arrival).



Graph 2, MSE Comparison for different NN Models. (Monitoring the Queue Changes)

## RESULTS

After training the NN on the data files, the weights of the NN were fixed. We compare both NN and CB schemes with the No-feedback scenario where no source regulation is considered. In our simulations we used test using trace-1 feeding HP-1 and HP-2 and trace-2 feeding LP buffer. In all simulations, the HP buffers size was fixed to 200 cell slots and the bandwidth was set to 5520 cells/s giving a cell loss ratio of  $10^{-4}$  in the HP-2 buffer and zero CLR for HP-1 buffer. Simulations were ended when the sources feeding the controlled LP buffer completed transmission. The performance metrics used to compare both NN and CB controllers were the Cell Loss Ratio (CLR), server utilisation and the transmission delay. In this paper, we define CLR as the total number of cells discarded ( $C_{d,LP}$ ) at the controlled LP buffer divided by the total number of cells generated ( $C_{g,LP}$ ) at the sources feeding the LP buffer. For delay, we define the time to complete the transmission without feedback control to be  $TD_k$  seconds, and the time to complete the transmission with feedback control as  $TD_h$  second. Delay is then defined as:

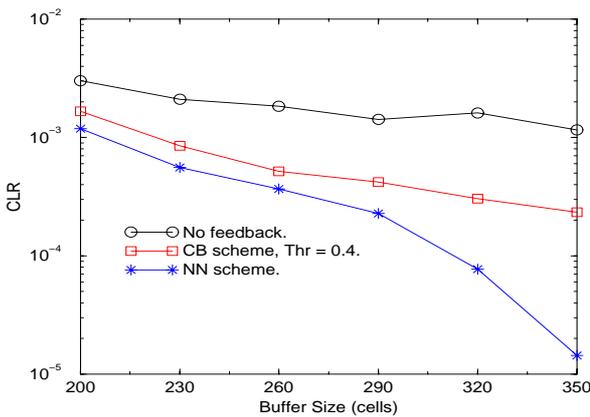
$$Delay = 100\% \times \frac{(TD_h - TD_k)}{TD_k} \quad (2)$$

Utilisation in the LP buffer is defined as:

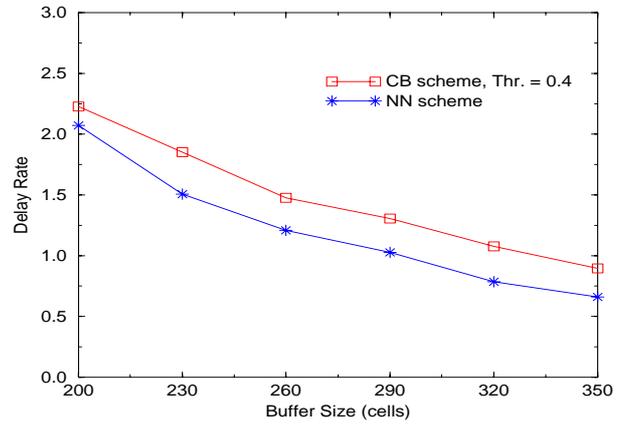
$$LP\_Utilisation = \frac{(C_{g,LP} - C_{d,LP})}{C_{g,LP,Max}} \quad (3)$$

where  $C_{g,LP,Max}$ , is the maximum number of cells that could be transmitted from the LP buffer.

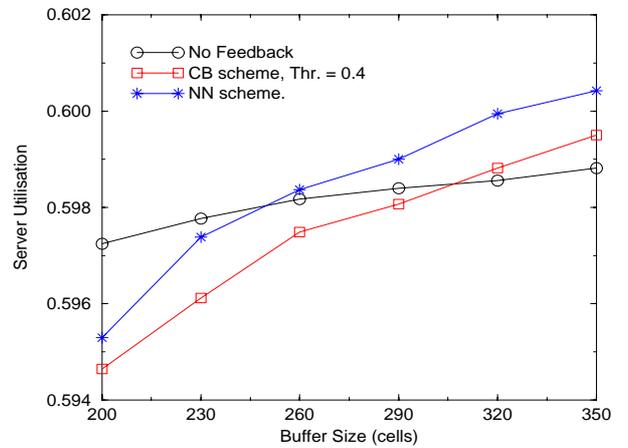
The results in graphs 3, 4, and 5 show that the performance of the NN scheme has 1-10 times better CLR, lower delay and higher utilisation when compared to the CB scheme.



Graph 3, CLR Comparison, One Trace1 in HP-1, 1 Trace-1 in HP-2, 1 Trace-2 in LP.



Graph 4, Delay Comparison, 1 Trace1 in HP-1, 1 Trace-1 in HP-2, 1 Trace-2 in LP.



Graph 5, Utilisation Comparison, 1 Trace1 in HP-1, 1 Trace-1 in HP-2, 1 Trace-2 in LP

## CONCLUSION

In this paper we have proposed a new NN congestion controller for a prioritised ATM switch. The test carried out in this paper is for a switch with three buffer priorities. The NN learning curves have shown that to control congestion in the low priority buffer using the NN scheme it is necessary to monitor the low priority buffer as well as higher priority buffers. It was also found that monitoring highest priority buffers has a little effect when controlling the lowest priority buffer. In this paper task of NN was split into two NNs, the first NN-1 to predict the congestion status (congestion or no congestion), and the second NN-2 to predict the explicit amount of reduction required. This approach has shown better NN generalisation compared to the one NN approach. We have also shown that the NN congestion controller outperforms the CB congestion controller in CLR and delay and utilisation.

## REFERENCE

- Al-Hammadi A, Schormans J, 1998, "Neural Network Congestion Controller in a Prioritised ATM switch", *Electronics Letters*, 29-Oct. 1998, Vol. 34, No. 22, pp2097-2098.
- ATM forum 1995, Dec. 1995. technical committee, traffic management specification, Version 4.0. ATM Forum/95-0013R9, Straw Vote,
- Garrett M and Fernandez A, (1994). Variable bit rate video bandwidth trace using MPEG code. Available FTP: *thumper.bellcore.com*, Directory: pub/vbr.video.trace.
- Habib, I. 1995 "Neurocomputing in High-Speed Networks", *IEEE communication Magazine*, , pp. 38-40
- Liu Y-C and Douligeris, c. 1997. "Rate Regulator with Feedback Controller in ATM Networks- A Neural Network Approach" *IEEE Journal on Selected Areas in Communications*, Vol.15, No.2, Feb. 1997, pp. 200-208.
- Newman P 1993, "Backward Explicit Congestion Notification for ATM local Area Network" *IEEE, Globecom'93*, pp.719-723, Houston, TX, 1993.
- Onvural., R. O. 1994. "Asynchronous Transfer Mode Networks, Performance Issues", by Artech House.
- Uniwuerzburg, Available FTP: *ftp-info3.informatik.uni-wuerzburg.de*.
- Zadeh L, Oct. 1996. "The rule off fuzzy logic and soft computing in the concetion, design and deployment of intelligent systems", *BT Technolgy Journal*, Vol.14 No.4 Oct. 1996, pp.32-36.

## BIOGRAPHY

ARIF SULTAN ALHAMMADI was born in Abu Dhabi, UAE. Graduated in 1995 from Etisalat College of Engineering in Sharjah. Received his PhD from Queen Mary University of London in 2000 after which he joined Etisalat College as an Assistant Professor where he teaches Artificial Intelligence and Computer Networks

# MODELLING FIXTURE PLANNING AND DESIGN USING ARTIFICIAL INTELLIGENCE

Zaid H. Yakoub  
Mithal A. Al-Bassam  
Mazin B. Adel

Department of Industrial Engineering  
University of Technology  
Baghdad, Iraq  
*e-mail: [m.bassam@uruklink.net](mailto:m.bassam@uruklink.net)*

Ibrahim A. Rawabdeh

Department of Industrial Engineering  
University of Jordan  
Amman, Jordan  
*e-mail: [rawabdeh@fet.ju.edu.jo](mailto:rawabdeh@fet.ju.edu.jo)*

## KEYWORDS

Computer Aided Tooling (CAT), CIM, Fixture Design, Knowledge-Based Systems, Artificial Intelligence.

## ABSTRACT

Continual advances in computer-aided design (CAD), computer-aided manufacturing (CAM) are enabling more multidisciplinary design. However, computer-aided tooling (CAT) which is a critical part of process design and bridge between CAD and CAM has been least addressed and remains a missing link. In manufacturing, tooling includes selection, design, and fabrication of cutter tool, fixturing system, inspection devices and molds.

The aim of this paper is to present a developed intelligent knowledge-based system for planning and designing fixture systems for rotational parts. The developed system helps designers in increasing the flexibility of automatically designing a fixturing system by building and using a database system that gives the user the ability to enter, remove, change and expand all related data, by building a knowledge-based system that helps in giving decisions concerning work piece setup, constraints of work piece geometry and positions of contact points of the work piece, by providing the ability to select the collection of different fixturing elements to assist the user in building the fixturing system configuration and by generating a drawing file of these elements selected for easy fixturing system construction. The system has been tested using a rotational part that needs machining a hole at an inclined position.

## INTRODUCTION

In a real CIM environment, it is expected that the entrenched fixture design system should be able to deal with any common fixture design process so that a complete automation of the overall manufacturing process is possible. Proper fixture design is essential to the operation of advanced manufacturing system and product quality in terms of precision, accuracy, and finish of the machined part (Bijan 2002; YU 2002). Fixture design is a complex task since many variants must be satisfied simultaneously to achieve the optimum result; therefore, the manufacturing research community has focused on the developing and improving technology such as CAD/CAM, Knowledge-Based (Expert) System to develop automated fixture design system (Pham and Lazaro 1990).

In the hierarchy of computer aided design and manufacture (CAD/CAM) a fixture design system is the interface between product design (CAD) and process planning on one hand and computer aided manufacturing (CAM) on the other. Such integration could increase design efficiency via important on-line information about manufacture ability or design analysis result. Generally, these characteristics reduce the time required for the designing process and aid in producing a better product design (Chang et al. 1982; Kale and Pande 2000).

Fixture design is an important activity in manufacturing which has significance upon productivity and product quality. It is a complicated, experience based process which needs comprehensive qualitative knowledge about a number of design issues (Utal and Jianmin 1998). Traditional fixture design activity requires heavily human effort, skill, heuristic knowledge and link between design, development and production and therefore, requires longer lead times, also an extra cost arises in manufacture. Thus, when applying CAD techniques to automatically generate fixture configurations and using CNC techniques to fabricate the fixture component, the design time, cost and the manufacturing lead-time involved in producing jigs and fixtures are minimized (Miller and Hannam 1985; Kale and Pande 2000).

Developing a design procedure which can be computerized either in full or in part, requires that a systematic structured design procedure exists. It is also extremely helpful if a large amount of the design is carried out by quantitative analysis. Such design procedures generally have the form of the object known from previous practice. Only its detailed dimensions are unknown and these are determined by calculation. Jig and fixture design does not result from a quantitative analysis but comes within a range of design procedures which are partly creative and requires existing items to be brought together as assemblies. Thus, it is required to provide an appropriate fixture design environment which fosters communication between various experts involved in the field.

A Computer Aided Fixture System Design (CAFSD) will put emphasis on the cooperation with the other design systems; the connection should be bi-direction. The computer-based fixture design environment is therefore, expected to accommodate different modeling paradigms within a single integrated framework. This framework would be accessed by designers through interactive programs so that their intelligence and experience could also be incorporated within the total CAD/CAM based design procedure. The results of

CAFSD should be sent to CAPP system directly and CAFSD processed information should also be feedback to the other systems as early as possible (Zhang and BI 2001).

## FIXTURING SYSTEM DESIGN (FSD)

There are four general requirements of a fixturing system (Yu 2002; Sunder 1999; Tao et al. 1999): Accurate Location (stable resting and deterministic localization), Total Restraint (stable clamping), Limited Deformation (support reinforcement) and No-Machine Interference. Practical consideration that indicate a good design are as follows (Chang et al. 1982):

- Shape, material and state of work piece.
- Pre-machined surfaces and tolerance.
- Type of machining operations and machine tools.
- Workpiece handling, ergonomic and safety factors.
- Economic consideration in type of fixtures selected.
- Reducing idle time of the machine tool to minimum.
- Other considerations such as chip removal, coolant flow, setup, work piece inspection, fixture storage, avoiding vibration effect, design foolproof, adequate clearance, etc.

The design of jigs and fixture is a very specialized activity that has to be completed quickly because time spent in the design and subsequent manufacture of jig and fixture for a part usually causes on equivalent delay in the initial manufacturing of the part. It is also a critical activity because jigs and fixtures are an extra cost arising in manufacture. Thus, it is important that both the design time and cost involved in producing jig and fixture are minimized (Miller and Hannam 1985).

### Phases of Fixturing System Design

**First Phase:** Examination of all information pertinent to the product (Workpiece) as given by engineering and/or manufacturing drawing or process sheets (Wilson 1962).

**Second Phase:** Setup planning to determine the number of setup, the position and orientation of the workpiece in each setup and machining surface in each setup (Huang et al. 1999).

**Third Phase:** Fixture planning to layout a set of locating and clamping points on workpiece surfaces such that the work is completely restrained (Chou et al. 1989).

**Fourth Phase:** Fixture configuration design to select or generate components and place them into a final configuration to locate and clamp the workpieces. A fixture configuration is made up of three types of components:

- a- Functional Component: The components which are directly in contact with the workpiece to perform locating and holding function. This component consists of three units: Locating Unit, Clamping Units, Locating and Clamping Units, and Guiding Tool Units.
- b- Fixture Base: Are components, which hold all the functional components and support together to make one integrated structure.
- c- Mounting Component: are components, which connect the locator/clamps down to the fixture base (Rong et al. 1997).

For most operations no kinematics freedom is permitted (Sayeed and Meter 1994). The most common method of restricting the movements of the part is the 3-2-1 or the 4-2-1

locating principle (Sunder 1999, 30]. In the 4-2-1 method of location four points are positioned on the primary locating face (Sunder 1999). For simple cylindrical or conical shape, just five locaters are needed, because one rotation is stopped purely by friction. Also, force analysis is concerned with checking that the forces applied by the fixtures are sufficient to maintain static equilibrium in the presence of cutting force. Clamp selection and placement is critical, if too few clamps are chosen or improperly placed, the workpiece will be displaced during machining. On the other hand, too many clamps applied to the workpiece can interfere with the desired tool path as well as drive up the cost of fixture (Sayeed and De meter 1994).

## LITERATURE REVIEW

Fixture planning is building the relation among the design variables, constraints, and evaluation objectives. The optimal formula of fixture configuration design is not completed until all of objective and constraint models are established (Zhang and BI 2001). The feasible methodologies of fixture planning are discussed.

### *Geometric (Workspace) Representation*

The principle of fixturing system requires knowledge of the workpiece, which includes specifications such as tolerance and surface finish. The goal is to represent this information and knowledge relevant to fixture design for analysis and assembly of the fixture. The use of solid modeling (geometric modeling) has been a focal point of research because it offers a realistic view of the workpiece. Du and Lin (1998) developed a CNC three-fingered Fixturing system able to configure automatically, with flexibility to deal with, planer objects of different shapes and sizes. The approach used can be extended to 3D object. Wang and Pelinescu (2002) described an approach to automatically design optimal fixtures for 3D workpiece. The approach applies to part with arbitrary complex geometry.

Wu et. al. (1998) presented an analysis on fixturing accuracy, clamp planning, fixturing accessibility, and clamping stability. Geometric analysis has been conducted to determine the fixturing surfaces and points, which provide feasible geometric constraint and assembly relationships with modular fixtures. Youcef et. al. (1989) presented the overall requirements and guideline for automatic setup of a reconfiguration of modular fixtures. These include requirements for a proper design, and guidelines for determining the layout requirements by examining task, workpiece system information. Don et. al. (1999) provided a reference model for fixturing planning in which Hopfiled ANN algorithm is adopted to identify a work part to be fixed.

### *Kinematics Analysis*

Kinematic Analysis provides an accurate medium for representing cutting forces as well as locating and clamping of workpieces. Some of the researches referring to this methodology use mathematical kinematics analysis. Chou et. al. (1989) applied screw theory to study the kinematics constraints of fixtures workpiece. Sayeed and De Meter (1994) created a software, which provides analysis capabilities to validate kinematics and total restraint. Brost

and Goldberg (1996) presented an implementation algorithm that accepts a polygonal description of the part silhouette, and efficiently constructs the set of all feasible fixture designs that kinematically constraint the part in the plane.

### Finite Element Analysis (FEA)

Due to the strong capabilities of deriving compressive force and displacement, the FEA approach has been used in many deflection-related research (Zhang and BI 2001). Lee and Huynes (1987) used finite element analysis to minimize fixturing force and work piece deflection.

### Artificial Intelligence

The fixture design process is a highly intuitive process which is based on heuristic knowledge and the experience of tool designer. Artificial intelligence (AI) has gained wide acceptance as a tool for solving manufacturing problems. The field of AI has advanced to the point where AI techniques are beginning to produce promising result both in academic and industry (Sluga et al.1988). Because of the gained advantages of representing experience-based knowledge, expert systems assume a logical role for automating the fixture design process.

## AI-BASED FIXTURE PLANNING AND CONFIGURATION

In fixturing system, design and development algorithms supported by computer software and AI system to increase automatic fixturing system design have become very important. Because, fixturing system design is of a high degree of complexity, it requires the use of computer software and AI techniques to support decision-making appropriate for problem solving. Also success in getting on the best workpiece setup and fixturing design planning lead to success or improvement in the design of sequence operations in process planning and prevent change or re-design in these sequences. The techniques based on artificial intelligent and using fixturing system design offers most promising results or solutions as shown in following section:

Markus et. al. (1984) developed an expert system using PROLOG to generate the fixture configuration using rest buttons, towers, and clamps. Kale and Pande (2000) presented the algorithm for automatic fixture layout planning for machining setups; focusing upon determining the most suitable and clamping position in accordance with the 4-2-1 configurations, by geometric reasoning of the CAD part model. Pham and Lazaro (1990) described Auto Fix (Fully automated Jig and Fixture design system) which is integration of knowledge-based reasoning with a whole range of traditional CAD techniques including solid modeling, parametric design and finite-element analysis. Miler and Hannam (1985) described how a knowledge-base can be set up and exploited interactively without a full-scale expert system needing to be developed as well. Ngoi and Leow (1994) presented a software to assist the tool designer of modular fixturing assembly; the software comprises a knowledge-based advisor, a fixturing assembly program and a CAD system. Joneja and Chang (1999) presented the development of a knowledge-based setup planner and an interacting fixture planner for the Quick Turnaround Cell (QTC) system. Several AI techniques are used in the fixture

configuration area of research such as Rule-based reasoning (Miller and Hannam 1985), Genetic algorithm, Case-Based Method and Neural Network Algorithms (Zhang and BI 2001, Sunder 1999, Huang et al. 1999, Mark et al. 2000).

### ALGORITHM DEVELOPMENT

The following demonstrates the use of AI in the design and development of an algorithm that serves in increasing the flexibility of fixturing system design for rotational work pieces. This Automated Fixture System Design (AFSD) requires developing and using a database system to present inputs (Workpiece, Production plan, fixturing system existing and fixturing element) to this algorithm. Also a knowledge-based system was developed to find the best fixturing analysis to the workpiece and find fixturing configuration. Finally, an AutoCAD package is used for creating drawing file of selected elements.

#### System Architecture

The system architecture of AFSD is depicted in Figure 1.

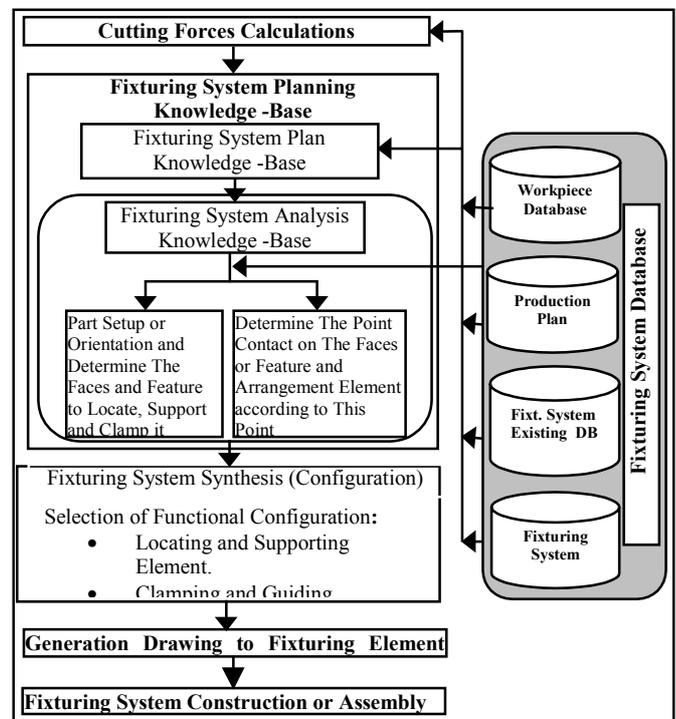


Figure 1: The (AFSD) System

The AFSD system consists of the following:

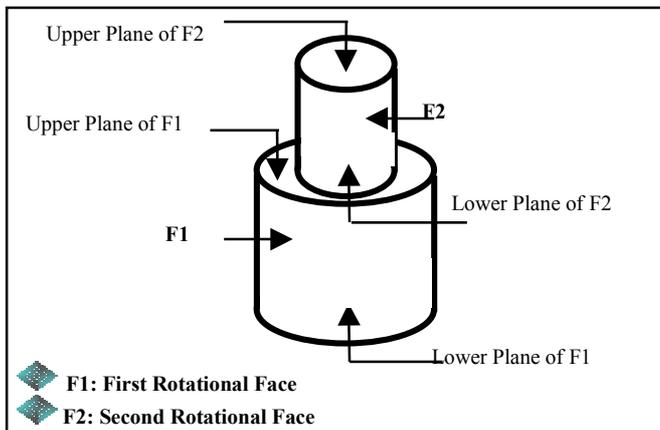
#### Fixturing System Database

The used database is classified into four types:

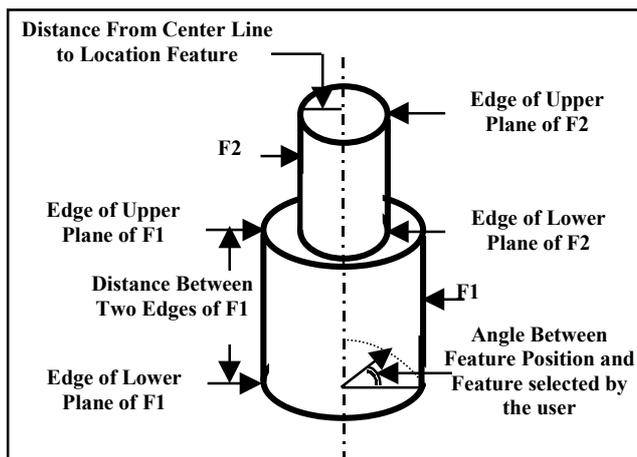
##### 1) Workpiece Database

This describes in detail the workpiece to be machined and is central to a fixturing system design. It contains the overall physical properties information (Part Name or Classification, Type and Kind of Material and Properties of Part Material), technological information (Number Of Pieces to be Made; Degree of Accuracy (Initial Surface Quality); Spindle Direction Specification rectangular Cartesian coordinate system; Required Cutting Direction), and configuration (geometry) information of the workpiece (Overall shape, Overall Size, Relative Size, Number and Name of Faces, Number of Machined Feature, Status and Faces Quality, List or Name of Machined Feature, Dimension of Machined Feature, Number of Machined Feature in Each Face,

Location of Machined Feature and Position of Machined Feature). As an example, for rotational workpiece each change of diameter on this work piece is represented as a face consisting of rotational face, upper plane and lower plane as shown in Figure 2. Also to determine the location of machined feature on outer faces the user uses the distance between the upper and lower edge of the same face and use rule specified above to find the angle to specify the position on this face. This rule can be shown in Figure 3.



**Figure 2: Determining Number and Name of Faces on the Rotational Workpiece**



**Figure 3: Determining Position of Machined**

## 2) Production Plan Database

This information applied by the user in this database can be classified into two main types:

- **Machine Representation Information:** -Contains the following fields: Machine Name or Classification; Type of Operation; Dimension of Machine Table; T-Slot Representation; Table Swivel; Max or Min Distance Between Table and Spindle; Range of Table.
- **Tool and Cutting Condition Information:** Contains Number and Rang of Speed, Number and Rang of Feed and Type and Size of Cutter (Specified the type, size and material of cutter or tool used) fields.

## 3) Existing Fixturing System Database:

contains classifications of types of fixturing system that can be used to produce different workpieces.

## 4) Fixturing System Element Database:

can be classified into five basic fields that contain the overall type of fixturing system elements.

### Cutting Force Calculations Module

Determining values and directions of cutting forces for operation required is a very important stage in building a successful fixturing system design since it can specify location of fixturing system element, in addition to the amount of force required from the element to provide equilibrium to the work piece. This equilibrium leads to get the full constraint (resistances deflection and movement by cutting force) to the work pieces.

### Fixturing System Planning Knowledge-Based Modules

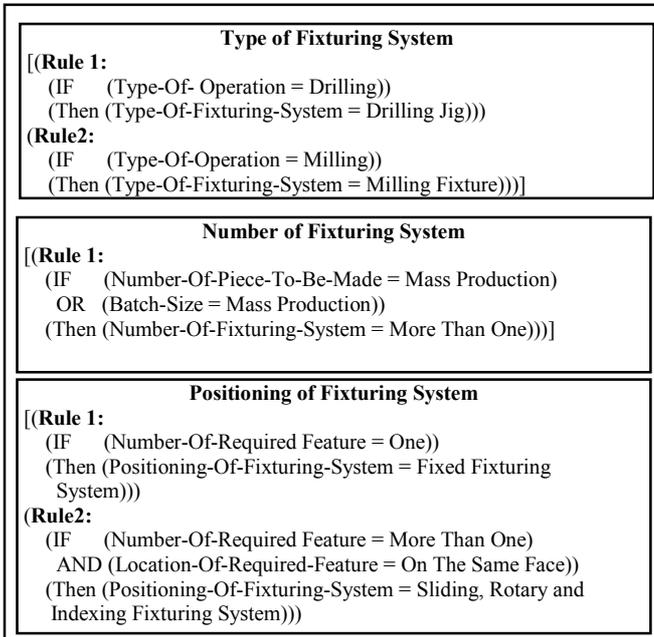
The objective of fixturing planning is to determine the number of setups needed, the position and orientation of workpiece in each setup and to determine the locating and clamping surfaces and contacts on the workpiece. Generally, determining the fixturing system planning depends on the representation or description of the details of geometric workpiece. A rule-based system is being used to represent and process these details to get the efficient results to this stage. The condition of part of a rule is checked every time a rule is selected for firing. Rules (Procedural knowledge) are best represented as rule base. These rules are the real corner stone of building the knowledge base of system (Fixturing Plan and Fixturing System Analysis).

The inference strategy used in this system is forward chaining (data driven) that starts with known goal and tries to use the expert programs rule to get to the goal. The rules used in this research may be categorized into rules for finding the type of fixturing, number of fixturing system and positioning of fixturing system, rules for finding the number of setup operations, rules for specifying the faces that are used for locating, supporting and clamping, and rules for specifying the number, type and arrangement of contact on these faces.

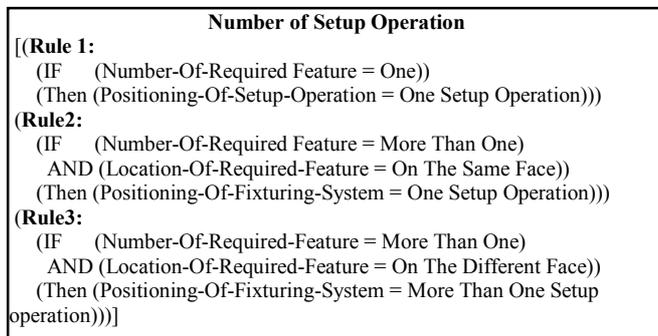
**Fixturing System Plan Knowledge-Base (FPKB):-** This Knowledge Base contains specification type, number and position of fixturing system depending on the type of operation required, number of workpieces to be made or batch size and number of required feature and location of this feature on the work piece. These stages of fixture planning are performed through three steps:

1. **Determine Type of Fixturing:** Where the user enters the type of operation required which the system processes this entered data to select the types of fixturing system appropriate to that type of operation.
2. **Identify Number of Fixturing Needed:** When the user enters the number of pieces to be made or batch size the system informs the user to use one fixturing.
3. **Identify Positioning of Fixturing:** In this phase, by entering the number of required features then the system select a fixed fixturing for a single feature if not the user have to specify location of required features so that the system would select a sliding, rotary and indexing fixturing system.

These types of fixturing systems are selected when features to be machined are located on same faces of workpiece. Otherwise if these feature are located on more that one face of workpiece, the system will advise the user to use a fixture for each group of features that are located on the same face. Examples of rules used for finding Type, number and Positioning of fixturing system shown in Figure 4.



**Figure 4: Example of Rules Used For Fixturing Plan**



**Figure 5: Rules Used For Determining No. of Setup Operation**

**Fixturing System Analysis Knowledge-Base (FAKB):**

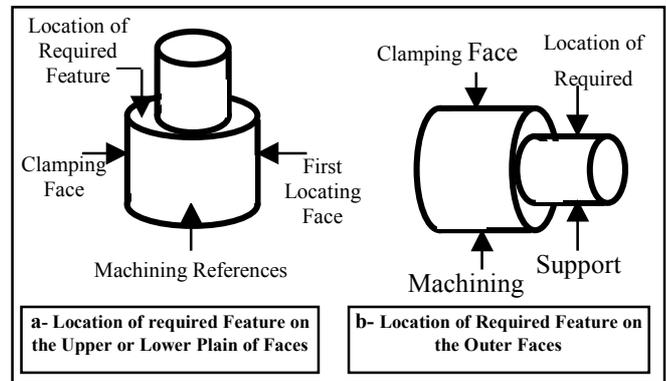
This stage can be classified into two main steps:

**1- Workpiece Setup (Machining References, Locating, Supporting and Clamping Face):** For each setup, one must ensure the fixturing configuration that will provide a force closure on the workpiece and enable reaching the feature to be machined without interfering with tool path. Generally, determining the workpiece setup depends on the representation or description of the details of geometric workpiece. In this research, specifying the number of setups depends on the number of required features and location of these feature. For example, if the number of required feature equals one, the system specifies one setup operation. If the number of required features is more than one, the user has to select features to be machined that are located on one face in order to be machined in one setup operation, and also to repeat for

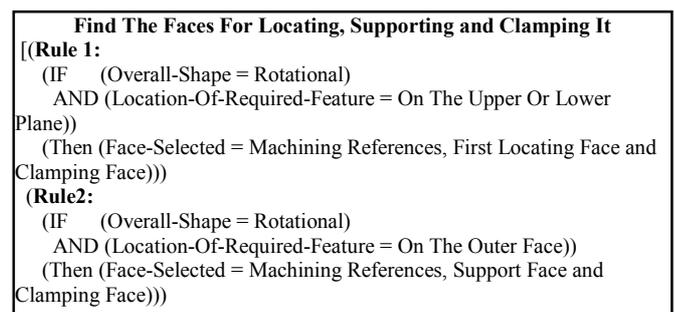
other features located on the other faces. Figure 5 shows examples of rules used for determining number of setup operation.

After specifying the required the number of setups, the machining references, first locating face, supporting face and clamping face are determined. In the case of rotational workpiece, if the required feature (to be machined) exists on the lower or upper plane of faces, the system selects from upper or lower plane of workpiece as machining references and select the face that has maximum height for locating, supporting and clamping according to the rule used in constraining the large rotational workpiece.

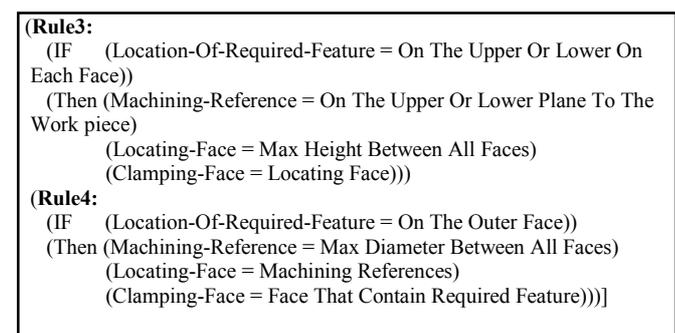
On the other hand, if required feature exists on the rotational face (outer face), the system select the outer face that has maximum diameter as machining reference and clamping face, also uses supporting face to support the face that contains required feature (to be machined). This rule used in this part is shown in Figure 6-a,b and Figure 7. Some of the rules used for determining the faces used locating, supporting and clamping are shown in Figure 8.



**Figure 6: The Rule Used in Constraint Rotational**



**Figure 7: Example of Rules Used For determining the Faces**



**Figure 8: Example of Rules Used For Determining The Faces**

**2- Determining Contact On The Faces Selected:**

After specifying the faces for locating and clamping, number, type and arrangement of contact on these faces must be found. This part begins by:

- i) Defining and determination of the machined feature existing on the faces selected.
- ii) Specifying the number and type of clamping contact on the clamping face. This contact is formatted according to the feature existing on the clamping face.
- iii) Selecting the locating and supporting point contact on machining references. In the case of rotational work piece use two-support point contact if no hole feature is selected.
- iv) Selecting the locating and supporting contact on the first and support faces. In the case of rotational workpiece use locating face contact that is formatted according to the feature existing on first locating face and use supporting face in the second case of rotational workpiece setup (if required feature existing on the each outer face) and selecting this support face depends on the length of selected face.

Intelligent fixturing analysis is being carried out by the system through two steps:

1. **Identify Number of Setup Required:** By entering the number of features to be machined, the system would suggest number of setups.
2. **Determine Contacts On The Selected Face:** During this phase contacts are selected depending on the location of the feature to be machined (RF) whether it is located on the upper or lower plane. This system will specify the machining reference, clamping face and supporting face. Thus, the number and type of contact is determined. Finally, these points are arranged on the selected faces and their location is determined. Some of rules used in this step are shown in Figure 9-a, b, c.

#### Fixturing System Synthesis Modules

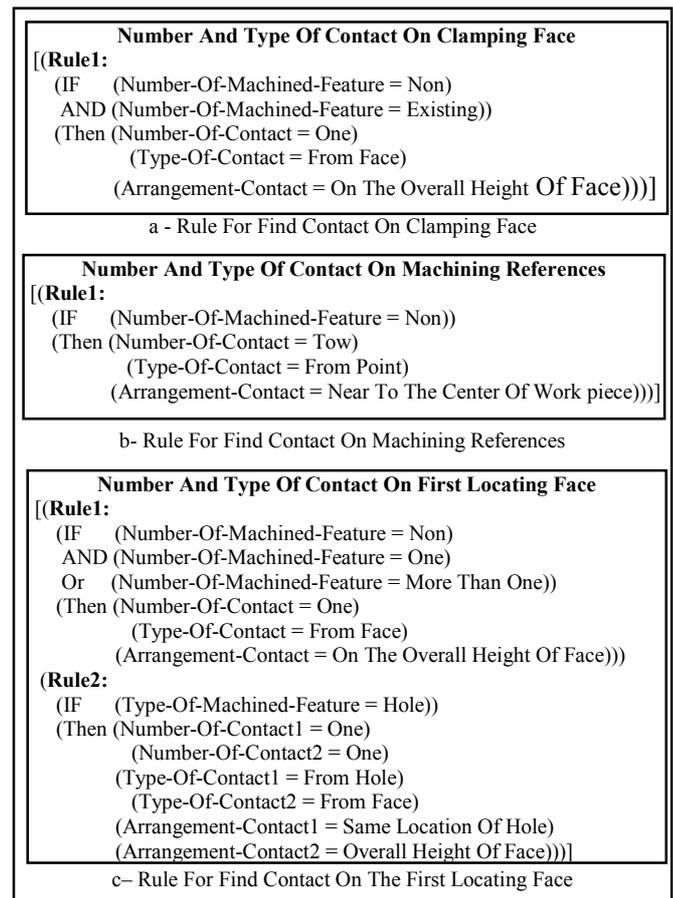
This stage contains the following steps:

1. **Selection of Fixturing Element:** including selecting the type and number of fixturing element appropriate to the type and number of contacts. This element is applicable to rotational workpiece.
2. **Design Consideration:** This step is applicable to the clamping, locating and supporting element:

##### *a) Clamping Element:*

- **Direction of Clamping Force:** - in rotational workpiece, clamping direction is vertical to the workpiece to prevent the movement resulted from cutting force and loading and unloading workpiece are easy.
- **Clamping Force:** - Forces applied to the workpiece from each clamping element selected depends on the screw force. The clamping force must be greater than the cutting force and moment to the machining operation.
- **Clamping Operating Method:** - each clamp must be manual operating to easy use and flexibility in movement of this element.

*b) Locating and Supporting Element:* This pinpoints relation element with clamping and cutting force and relation between them that is referred to in the arrangement contact on the faces using for locating and clamping.



**Figure 9. Examples of rules used for determining fixturing analysis.**

- **Element dimension:** The dimension of element selected is specified according to dimension of workpiece, faces and feature used for clamping and locating,
- **Material used to make element:** Carbon steel is the primary material of fixturing system. The ease of fabrication, low cost, availability, and versatility of this material have made it popular for fixturing construction. Connector and base plate are usually made from low carbon steel due to resistance to stress or wear, low-production and are easily to be welded and jointed. For supporting, clamping, locating and guiding tools are usually made from high carbon steel or tool steel because of its resistances the wear.
- **Direction of Fixturing Element:** - These include pinpointing the position of Fixturing element (Vertical or Inclined) depending on the direction of cutting force.

#### **AFSD SYSTEM TESTING**

AFSD system can be applied to rotational workpiece (X) to find fixturing system design for this application. Implementation of the developed system gives an efficient result in fixturing system design appropriate to this application. The system was developed using Visual Basic version (5) on a Pentium (2) PC. Also the developed system was linked automatically with Microsoft Access (2000) to build the database system and AUTOCAD version (2000) to help in generation of the required drawing.

**Description of the Application:** This involves applying the full description for a practical workpiece and production plan such as:

**Status of Material:** Machining  
**Type of Material:** Steel35φ14  
**Strength of Material:** 54 Kg/mm<sup>2</sup>  
**Hardness and Ductility of Material:** 207BHN and 32 Kg/mm<sup>2</sup>  
**Overall Shape:** Rotational  
**Overall weight:** 195.374 g  
**Relative Size:** Medium  
**Overall Size:** φ40<sup>-0.34</sup> mm and length 90.8<sup>+0.2</sup> mm  
**Pieces to be Made:** 500  
**Accuracy of Fix. Element:** 20-50% from each dimension  
**Spindle direction:** Vertical  
**Required Cutting direction:** Inclined in Angle 22<sup>±20</sup>° on the workpiece Axis (X-Axis).

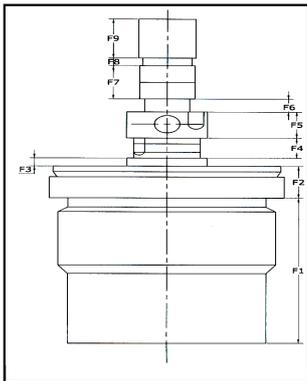


Figure 10. The Faces Exiting on the Practical Workpiece

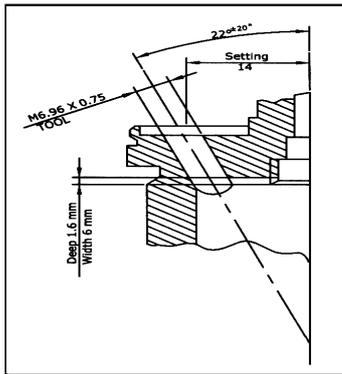


Figure 11. The Required Feature to be machined

**Workpiece Configuration:** This contains description of all faces and features machined previously, and to be machined according to the company's route sheets. This configuration is shown in Figures 10 and 11. The first stage, a User of AFSD System can enter the data referred to previously in description of practical application and production plan within database system. The database system considers the input stage to the AFSD system and contains all fixturing system and fixturing elements existing. These database windows are shown in Figures 12 to 16.

## CONCLUSIONS

Through this paper it has been observed that computer aids the designers in providing intelligent and automated tools for working faster and more accurately. Applying knowledge – based (rule –based) system to fixturing system analysis has proved to be a powerful tool and very useful to get the efficient results for the number of workpiece setups and finding faces or feature(s) for clamping, locating and supporting it. Also the arrangement of contact on these faces that provides the best equilibrium to the workpiece.

The (AFSD) system gives fixturing system designers the ability to take quick decisions at site when using it during any of the phases or stages of fixturing system. (AFSD) system gives the ability to change or increase the element used in fixturing system configuration depending on the fixturing element database that contains all elements existing in fixturing system. Finally, the (AFSD) system can respond quickly to any change in workpiece or production plan

design. Further work is required to improve the work so that it can be applied to other shapes of workpieces

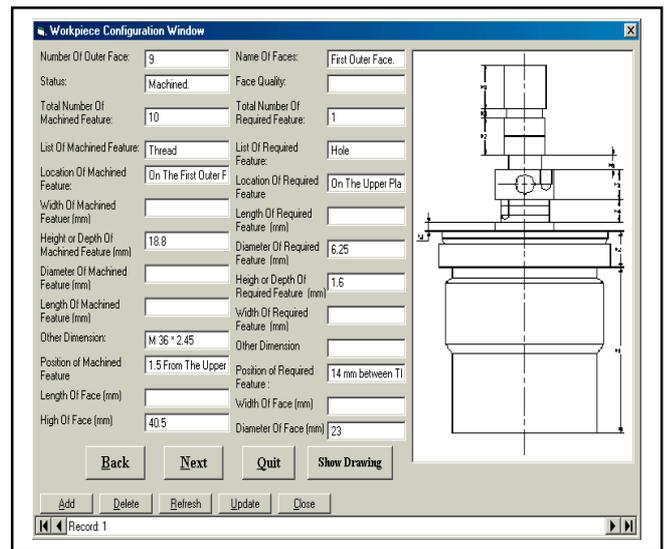


Figure 12. Workpiece Configuration Window

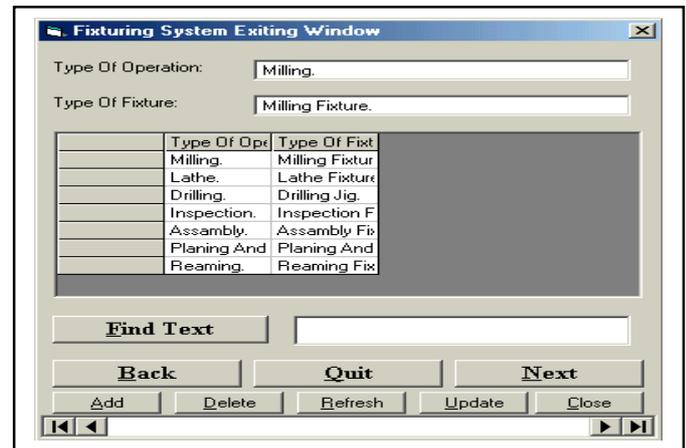


Figure 13. Existing Fixturing System

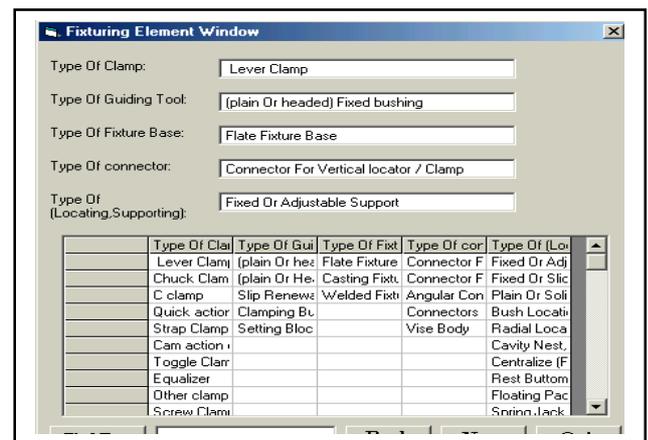


Figure 14. Fixturing Element

## References

- Ajay J. and Chang Tien-Chien. 1999. "Setup and Fixture Planning in Automated Process Planning System"; IIE Transaction; 31, 653-665.

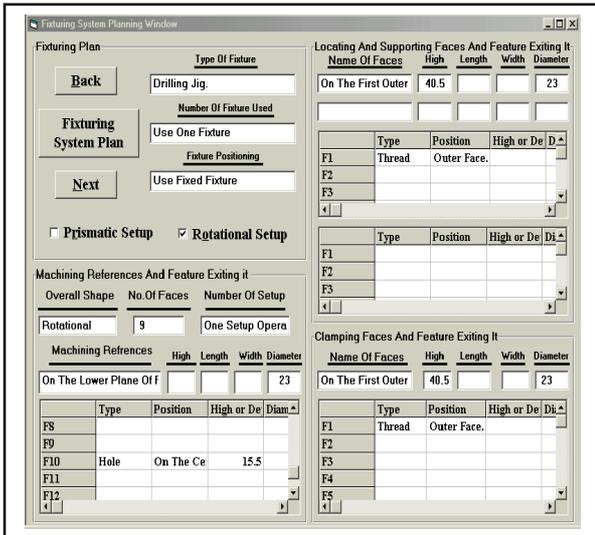


Figure 15. Fixturing Planning

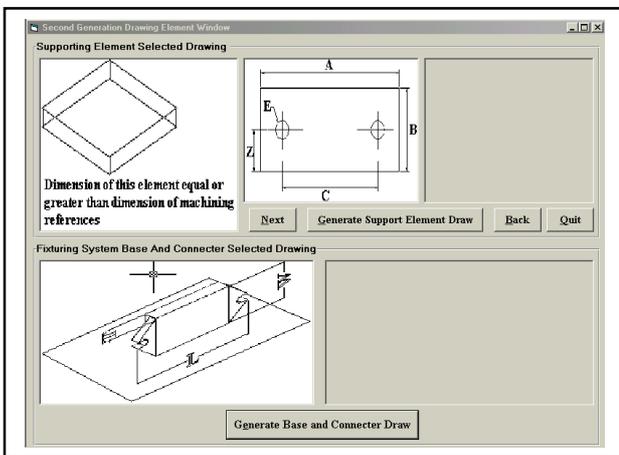


Figure 16. Element Generation Window.

of CAD Part Model”; Accepted – Flexible Automation and Intelligent Manufacturing Conference, FAIM 200, Univ. of Maryland, Maryland, USA; June.

- Lee D.L. and S.L. Haynes. 1987 “Finite-Element Analysis of Flexible Fixturing System”; Transaction of the ASME; Vol.109; May.
- Markus A., Z. Markusz, J. Farkas and J. Filemon. 1984. “Fixture Design Using Prolog: An Expert System”; Robotics and Computer-Integrated Manufacturing; Vol.1, No.2, PP.167-172.
- Miller S.A. and G.R. Hannam. 1985. “Computer Aided Design Using a Knowledge Base Approach and its Application to the Design of Jigs and Fixtures ”; Proc. Instn. Mech. Eng.; Vol.199, No. B4.
- Ngoi A.K.B. and L.G. Leow. 1994. “Modular Fixture Design: a Designers assistance” Int. J. Prod. Res.; Vol.32, No.9, PP.2083-2104.
- Pham T.D. and D. A Lazaro. 1990. ”AUTOFIX- An Expert CAD System For Jig And Fixtures”; Int. J. Mach. Tools Manufact.; Vol.30, No.3, PP. 403-411.
- Randy B.C. and G. Y. Kenneth. 1996. “A Complete Algorithm for Designing Planner Fixtures Using Modular Component”; IEEE Transaction on Robotics and Automation; Vol.12, No.1, PP.31-46.
- Rong Y., T.C. Cha and Y. Wu Y. 1997. “Automated Generation Of Dedicated Fixture Design”; Int. J. of Computer Applications In Technology; Vol.10, No.3-4, pp. 213-235.
- Roy U. and J. Liao 1998. “Application of a Blackboard Framework to a Cooperative Fixture Design System”; Computers in Industry, 37, 67-81.
- Sayeed A.Q. and C.E. DeMeter. 1994. “Machining Fixture Design and Analysis Software”; Int. J. Prod. Res.; Vol.32, No.7, PP.1655-1674.
- Sayeed, Q.A. and E.C. DeMeter. 1999. “Cmoliance based MIP model and Heuristic for Support Layout Optimization” Int. J. Prod. Res.; Vol.37, No.6, pp.1283-1301.
- Sluga A., P. Butat and M. Gams. 1988. “An Attempt to Implement Expert System Techniques In CAPP”; Robotic and Computer – Integrated Manufacturing; Vol.4, No.1, PP.77-82.
- Sundar B. 1999 “Fixture-Based Design similarity Measures For Variant Fixture Planning”; Institute For Systems Research (ISR).
- Sundar B., E. Alexei, W. J. Herrmann and Nau S. Dana. 1998. “Fixture-Based Usefulness Measures For Hybrid Process planning”; ASME Design Engineering Technical Conference; September 13-16; Atlanta, GA.
- Tao Z.J., S. Kumar, and A.Y.C. Nee. 1999. “Automatic Generation of Dynamic Forces For Machining Fixtures”; Int. J. Prod. Res. Vol. 37, No. 12, . 2755-2776.
- Wang M. Y. 2002. “Tolerance Analysis For Fixture Design”; Assembly Automation, Vol. 22, No. 2, pp. 153-162.
- Wu Y., Y. Rong, W. Ma and R.S. Leclair. 1998. “Automated Modular Fixture Planning: Accuracy, Clamping, and Accessibility analysis”; Robotic and Computer Integrated Manufacturing; Vol.14, No.1, PP.17-26; February.
- Youcef-Toumi K., J.J. Bausch and J.S. Blacker. 1989. “Automated Setup and Reconfiguration for Modular Fixturing”; Robotics and Computer-Integrated Manufacturing; Vol.5, No.4, PP.357-370.
- Zhang J.W. and M.Z. BI. 2001. “Review on Flexible Fixture Design and Automation”; Int. J. of Production Research; Vol.39, No.13, p.2867-2894.

- Balci, O. and R.G. Sargent. 1981. “A Methodology for Cost-Risk Analysis in the Statistical Validation of Simulation Models.” *Communications of the ACM* 24, No.4 (Apr), 190-197.
- Chang Tien-Chien, A. Wysk A. R. and D.P. Robert. 1982. “Interfacing CAD and CAM- A Study in Hole Design”; Computer and Indus. Eng.; Vol.6, No.2, pp. 91-102.
- Chou Y.C., V. Chandru and M.M0. Barash. 1989. “A Mathematical Approach to Automatic Configuration of Machining Fixtures: Analysis and Synthesis”; J. of Engineering for Industry; Vol. 111; November.
- Dong Tang, Ye Cheng, Fuzhi CAI and Quansheng LI. 1999. “Automatic Fixture Planning”; TSINGHUA SCIENCE AND TECHNOLOGY; Vol.4, No.2.
- Du Hong and I.C. Grier. 1998. “Development of an Automated Flexible Fixture for Planar Objects”; Robotics and Computer-Integrated Manufacturing; Vol.14, No.3, PP.173-183; Jun.
- Frank W.W. and H.M. John. 1962. “Handbook of Fixture Design”; McGraw-Hill Book Company.
- Huang S., Z. An , Y. Rong and S. Jayaram. 1999. “Development of Automated Dedicated Fixture Configuration Design System With Predefined Fixture Component Types: Part1, Basic Design”; Int. J. of Flexible Automation and Integrated Manufacturing; Vol.7, No.3/4, PP. 321- 341.
- Kale S.M. and S.S. Pande. 2000. “Automatic Planning of Machining Fixture Layouts Through Geometric Reasoning

# Neural Network Approach for the Optimal Control of a Container Crane

*Chokri REKIK, Mohamed DJEMEL and Nabil DERBEL,*

Research unit on Intelligent Control, design and Optimisation of complex Systems (ICOS)

University of Sfax, ENIS, BP. W, 3038, Sfax, Tunisia.

Phone: (216-4) 274.088, Fax. (216-4) 275.595.

Emails: [chokri.rekik@enis.rnu.tn](mailto:chokri.rekik@enis.rnu.tn) , [mohamed.Djemel@enis.rnu.tn](mailto:mohamed.Djemel@enis.rnu.tn) , [nabil.derbel@ieee.org](mailto:nabil.derbel@ieee.org)

**Abstract:** The aim of this paper is the determination of optimal control trajectories of a rotary crane. The proposed method is based on the decomposition of the system into two interconnected subsystems. Firstly, we assume many approximations for computing the optimal control. This is done by identifying optimal gains by neural networks. Secondly, the optimal controller is applied to the full system (without approximations). Simulation results show that the approximations yield satisfactory performances.

## 1- Introduction:

Artificial neural networks have been the focus of a great deal of attention during last decades [1,2,3,4], due to their capabilities in solving nonlinear problems by learning. Such networks provide a parallel structure with very simple processing elements. This paradigm has been applied to solve different technological problems, such as signal classification and more recently nonlinear and adaptive control problems.

Neural networks offer several advantages over conventional approaches, providing a general framework for modeling and control of nonlinear systems. Feedforward neural networks adopted in this investigation are the most popular architectures used in the neural control. They consist of several layers of processing units where neuron connections occur only between adjacent layers [5,6,7]. In this work, the proposed method consists in the decomposition of the considered nonlinear system into two subsystems. The second subsystem is a linear one, for which

the optimal gain is easy to compute. The first subsystem is a linear one, for which the optimal gain is obtained by solving the nonlinear Riccati equation. The second subsystem is also linear. However, its dynamical behavior depends on the state vector of the first one. In this context, a neural controller is identified for determining local optimal gains in terms of the first subsystem state vector.

This paper is divided into five sections. The second section presents the problem formulation. In section three, we present the decomposition of the system into subsystems. In section four, the proposed method based on neural networks is detailed. Finally, simulation results illustrate the proposed approach.

## 2- Problem formulation:

The fundamental motion of a rotary crane is the rotation and load hoisting [8,9]. For simplicity, we assume that the container load can be regarded as a material point and that frictional torques which may exist in torque-transfer mechanisms can be neglected.

The following notations, which are shown in figure 1, will be used:  $\theta_1$ , angle of rotation of the trolley drive motor;  $J_1$ , total moment of inertia of trolley drive motor, a brake, a drum and a set of reduction gears;  $b_1$ , equivalent radius of the drum of trolley drive motor which is reduced to the motor side;  $\theta_2$ , angle of rotation of the hoist motor;  $J_2$ , total moment of inertia of hoist motor, a brake, a drum and a set of reduction gears;  $b_2$ , equivalent radius of the drum of the hoist

motor which is reduced to the motor side;  $\phi$ , the load swing angle;  $m$ , total mass of the trolley and operator's cab;  $M$ , total mass of the contain load, the spreader, and the attached equipment;  $T_1$ , driving torque generated by the trolley drive motor;  $T_2$ , driving torque generated by the hoist motor; and  $g$ , the acceleration of gravity (figure 1). The process to be controlled can be described by a sixth-order nonlinear differential system, described by:

$$\begin{aligned}\frac{dx_1}{dt} &= x_2 \\ \frac{dx_2}{dt} &= u_1 + (-s_2 x_1 x_6^2 + g s_1 s_2 \cos(x_5) \sin^2(x_5)) / \alpha \\ \frac{dx_3}{dt} &= x_4 \\ \frac{dx_4}{dt} &= u_2 + (-s_1 (1 - s_2) \sin(x_5) x_1 x_6^2 \\ &\quad - s_1 \sin(x_5) \cos(x_5)) / \alpha \\ \frac{dx_5}{dt} &= x_6 \\ \frac{dx_6}{dt} &= -\frac{\cos(x_5) u_2}{x_1 + h} + [2x_2 x_6 (1 + s_1 - s_1 s_2 \sin^2(x_5)) \\ &\quad - s_1 \cos^2(x_5) + 1/2 s_1 x_1 x_6^2 \sin(2x_5) (s_1 - s_2) \\ &\quad + g \sin(x_5) (1 + s_1 - g s_2 \sin^2(x_5))] / \alpha (x_1 + h)\end{aligned}$$

where  $\alpha$  is a function of  $x_5$  :

$$\alpha = -1 - s_1 + s_2 s_1 \sin^2(x_5) + s_1 \cos^2(x_5)$$

Parameters  $s_1$  and  $s_2$  are given by  $s_1=66.66$  and  $s_2=76.66$ .

The state variables  $x_1$ ,  $x_2$ , and  $x_3$  are defined by:  $x_1=b_2\theta_2$ ,  $x_3=b_1\theta_1$ ,  $x_5=\phi$ .

The objective is to determine the optimal strategies of such system, in order to transfer the rotary crane from any initial state to the desired state described by:

$$X_d = [0 \ 0 \ 0 \ 0 \ 0 \ 0]^T. \quad (1)$$

We choose the following quadratic criterion defined as:

$$J = \frac{1}{2} \int_0^{\infty} (X^T Q X + U^T R U) dt$$

where  $X=[x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6]^T$ , the state vector,  $U=[u_1 \ u_2]^T$ , the control vector and  $R$  and  $Q$  are ponderation matrices:  $R=\text{diag}([20 \ 4])$ ,  $Q=\text{diag}([1 \ 10 \ 1 \ 40 \ 20000 \ 20000])$  [13].

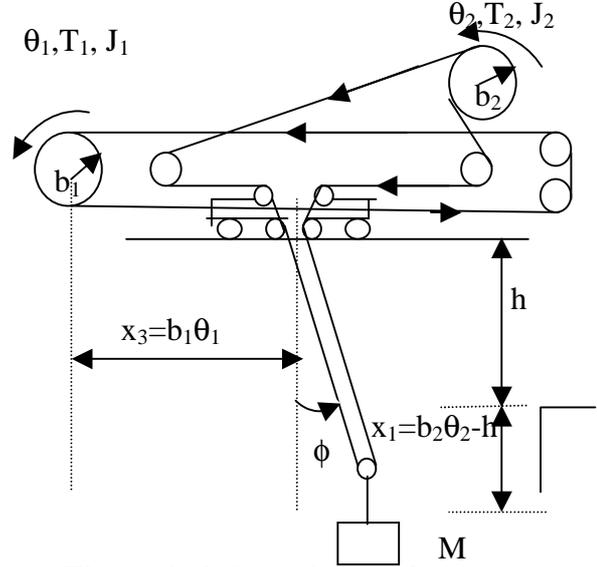


Figure 1: Schematization for the rotary

### 3-Decomposition of the system:

For simplicity, we assume that the load swing angle  $\phi$  is so small that all terms containing  $\phi^\alpha \phi^\beta$  ( $\alpha \geq 0$ ,  $\beta \geq 0$ ,  $\alpha + \beta \geq 2$ ) can be neglected and  $\cos \phi \approx 1$ ,  $\sin \phi \approx \phi$  hold. Consider the rotary crane described above (1). This system can be divided into two subsystems defined by equation (2) and (3) [13]. The subsystem 1 is defined by:

$$\frac{dx_1}{dt} = x_2 \quad (2)$$

$$\frac{dx_2}{dt} = u_1$$

The subsystem 2 is described by:

$$\frac{dx_3}{dt} = x_4$$

$$\frac{dx_4}{dt} = u_2 + s_1 x_5 \quad (3)$$

$$\frac{dx_5}{dt} = x_6$$

$$\frac{dx_6}{dt} = -\frac{1}{x_1 + h} [u_2 + s_2 x_5 + 2x_2 x_6]$$

It's clear that the subsystem 2, which depends on the state vector of subsystem 1, is linear with respect to its state vector and its control vector. However subsystem 1, which is also linear, is independent from subsystem 2.

#### 4-Optimal Control:

##### 4-1 Optimal control of subsystem 1:

Subsystem 1 being linear can be represented by the following condensed form of the state equation:

$$\frac{dX_1}{dt} = A_1 X_1 + B_1 u_1$$

where:

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{and} \quad X_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

The corresponding objective function to be optimized is described as follows:

$$J_1 = \frac{1}{2} \int_0^{\infty} (X_1^T Q_1 X_1 + U_1^T R_1 U_1) dt$$

where  $Q_1 = \text{diag}([1 \ 10])$  et  $R_1 = 20$ .

The optimal solution is then expressed as:

$$u_1 = -K_1 X_1$$

$$K_1 = R_1^{-1} B_1^T P_1,$$

where  $P_1$  is the solution of the following Riccati equation

$$P_1 A_1 + A_1^T P_1 - P_1 B_1 R_1^{-1} B_1^T P_1 + Q_1 = 0$$

The resolution of this equation gives:

$$P_1 = \begin{bmatrix} 4.3525 & 4.4721 \\ 4.4721 & 19.4650 \end{bmatrix}$$

Thus, the optimal gain is defined by:

$$K_1 = [0.2236 \quad 0.9732]$$

Finally, the optimal solution can be expressed as:

$$u_1 = -0.2236x_1 - 0.9732x_2$$

##### 4-2 Neural control of subsystem 2:

In this section, we propose to identify a neural network computing the optimal gain of the subsystem 2 in terms of the state of the subsystem 1.

The subsystem 2, whose state vector is  $[x_3 \ x_4 \ x_5 \ x_6]^T$ , is linear for given values of  $x_1$  and  $x_2$ . The main idea is to construct a database, for which we compute optimal gain for different values of  $x_1$  and  $x_2$ . Then, a neural network, which gives the optimal gain for each vector  $(x_1, x_2)$ , is identified. The subsystem 2 is represented by the following condensed form :

$$\frac{dX_2}{dt} = A_2 X_2 + B_2 u_2$$

where:

$$A_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & s_1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\alpha s_2 & -\alpha \beta \end{bmatrix},$$

$$B_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -\alpha \end{bmatrix}, \quad X_2 = \begin{bmatrix} x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}.$$

$\alpha$  and  $\beta$  are functions of  $x_1$  and  $x_2$ :

$$\alpha = \frac{1}{x_1 + h} \quad \text{and} \quad \beta = 2x_2$$

The corresponding objective function to be minimized is defined by:

$$J_2 = \frac{1}{2} \int_0^{\infty} (X_2^T Q_2 X_2 + U_2^T R_2 U_2) dt$$

where:  $Q_2 = \text{diag}(1 \ 40 \ 20000 \ 20000)$ ,  $R_2 = 4$ .

The optimal solution, which minimizes the quadratic criterion  $J_2$ , is a function of  $x_1$  and  $x_2$ .

So, the optimal gain depends on  $x_1$  and  $x_2$ . Then, one can easily write:

$$u_2 = -K_2 X_2$$

with:

$$K_2 = R_2^{-1} B_2^T P_2$$

where  $P_2$  is the solution of the following Riccati equation:

$$P_2 A_2 + A_2^T P_2 - P_2 B_2 R_2^{-1} B_2^T P_2 + Q_2 = 0$$

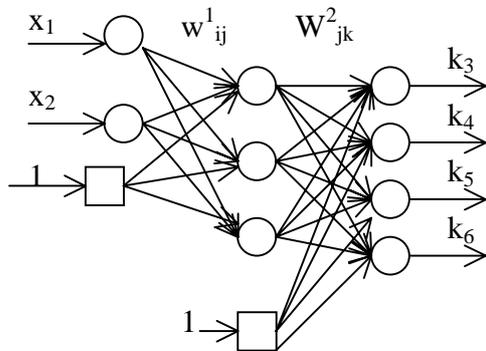
It is clear, that matrices  $K_2$  and  $P_2$  are functions of  $x_1$  and  $x_2$ .

The optimal solution can be expressed as:

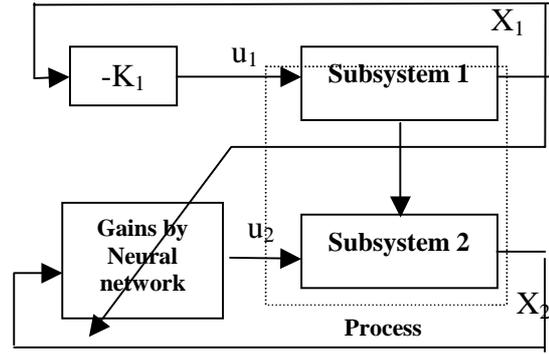
$$u_{21} = -(k_3 x_3 + k_4 x_4 + k_5 x_5 + k_6 x_6)$$

For 100 given values of  $x_1$  and 100 given value of  $x_2$ , we solve the Riccati equation. We obtain 10000 input-output pairs of training data. Then, a neural net has two inputs ( $x_1$ ,  $x_2$ ) and four outputs ( $k_3$ ,  $k_4$ ,  $k_5$ ,  $k_6$ ). The proposed net contains three layers: the input layer, the hidden layer and the output layer (see figure 2) [11, 12, 13].

The neural network contains three hidden neurons. We use the hyperbolic tangent function for the transfer function of neurons.



**Figure 2:** Adopted feedforward networks computing optimal gains of subsystem 2.



**Figure 3:** Structure of the neural controller

### 4-3 Simulation Results:

Results are presented in figure 4, it is clear that the state vector ( $x_1$  and  $x_2$ ) of subsystem 1 have rapid variation compared to the variation of the state vector of subsystem 2 ( $x_3$  to  $x_6$ ). This can be explained by the fact that the vertical movement of the load (subsystem 1) is practically independent from the swing (variables  $x_5$  and  $x_6$ ). However, the load horizontal movement (subsystem 2) is the principal factor generating the swing. Thus, the control vector of subsystem 2 is the principal factor responsible for the limitation of the hoisting. This is why subsystem 1 is more rapid than subsystem 2.

The obtained controller is applied to the full system (without approximation). Simulations results are presented in figure 5. Observing the state variations presented in figure 4 and figure 5, we notice that the system after and before approximation reached the desired terminal state after a certain transient. The criterion value given by the system after approximation is equal to 498.59. However, the criterion value given by the full system is equal to 498.37. This leads to an error of 0.044%. The obtained relative error on the state vector  $\frac{\|\Delta x\|}{\|x\|}$  is equal to 0.34%. Whereas,

the obtained relative error on the control  $\frac{\|\Delta u\|}{\|u\|}$  is equal to 0.042%. Analyzing these

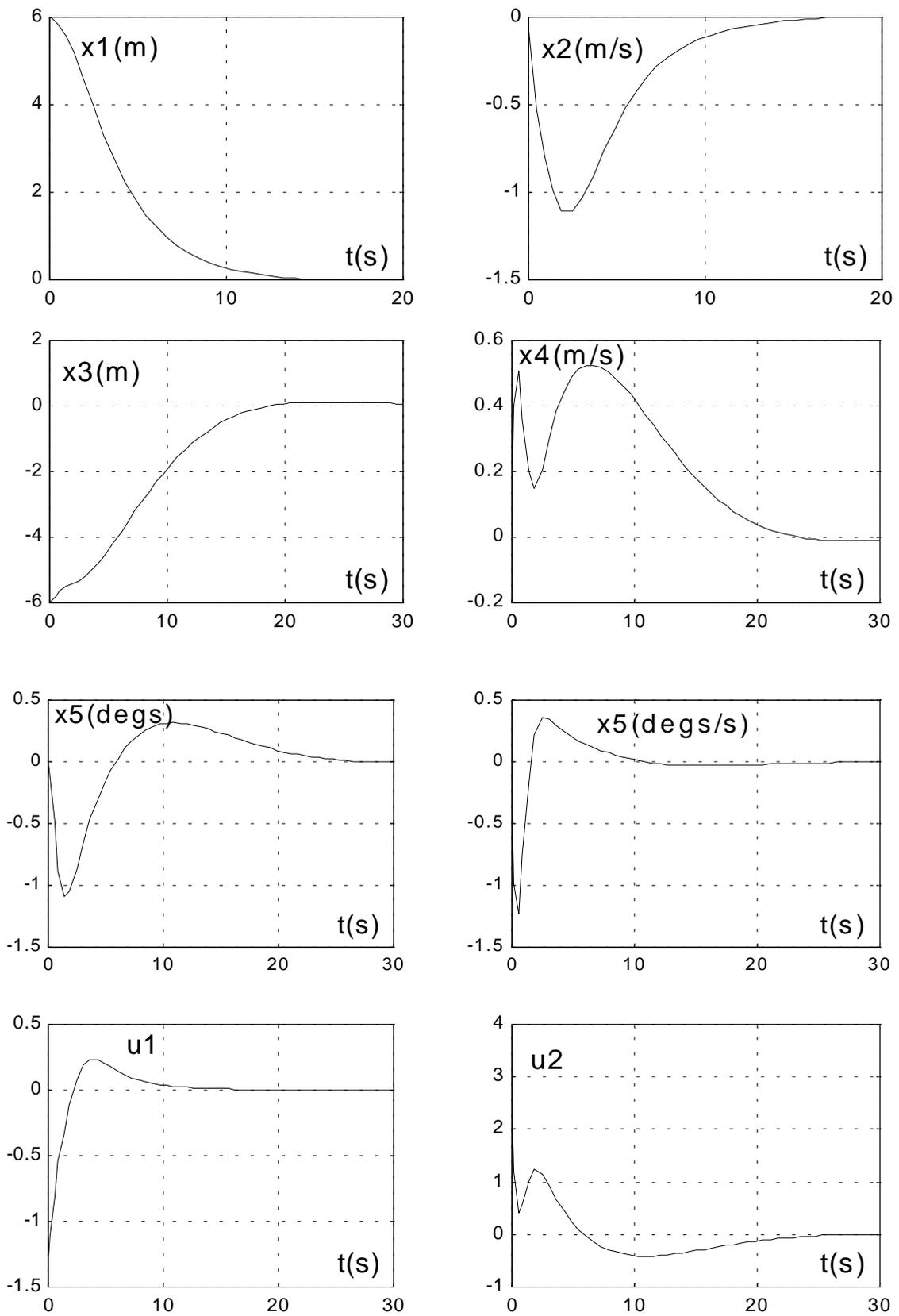
results and observing figure 4 and figure 5, it is evident that the approximations are very efficient.

## 5 Conclusion:

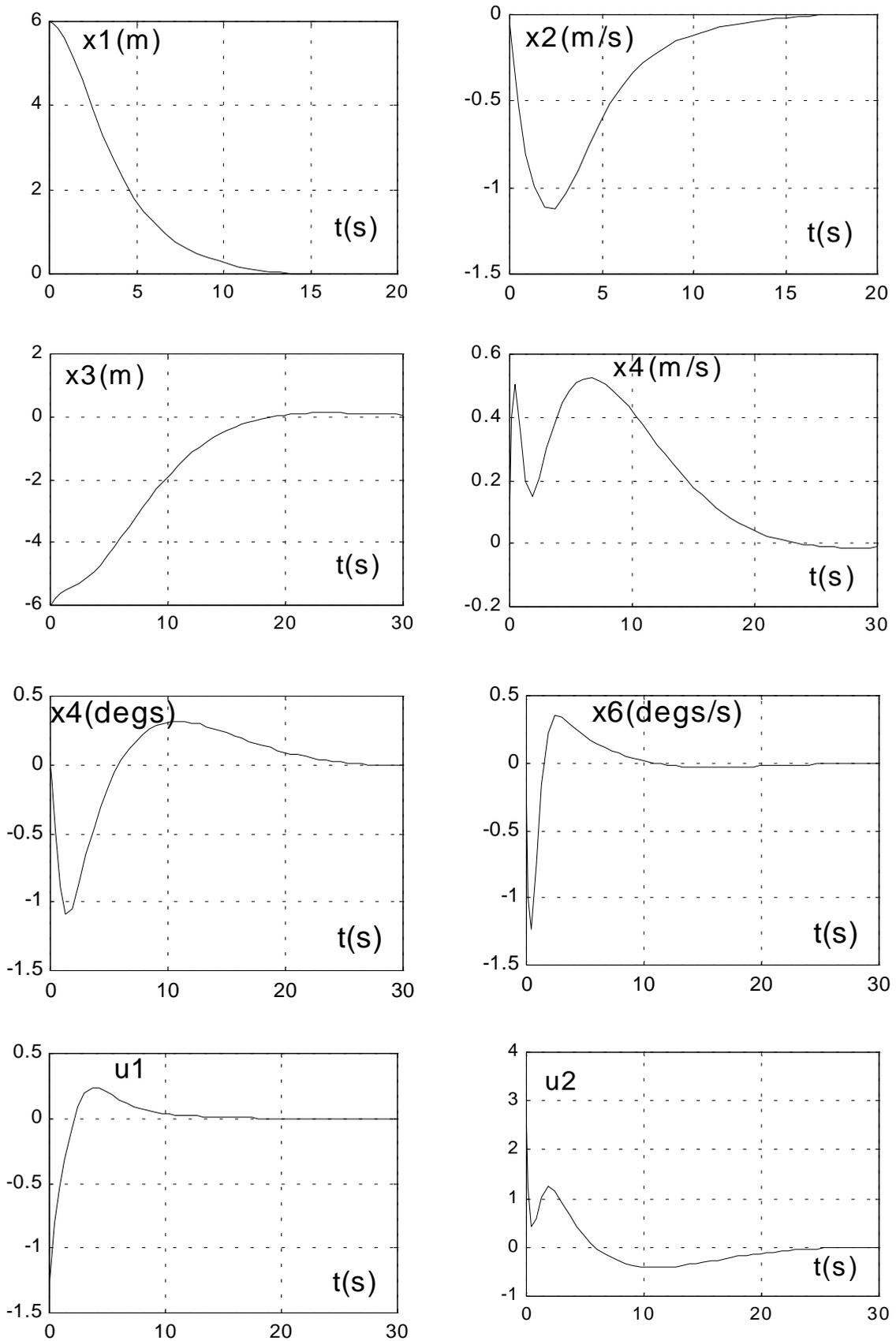
This paper concerns the determination of the optimal control strategies of a rotary crane. The system was decomposed into two linear subsystems. The dynamics of each subsystem depends on the state vector of the other one. A neural network controller is identified based on the computation of the gains of each subsystem in terms of the state vector of the other one. Simulation results show satisfactory performances. The optimal control is successfully applied to the full system (without approximation) yielding same results.

## References:

- [1]: A. Sideris, D. Psaltis and A.A. Yamamura: "A multi-layered neural network controller", IEEE Control System Magazine, April 1988, pp.17-21.
- [2]: J.A. Freeman, and D.M. Skapura: "Neural networks algorithms, applications, and Programming techniques", Addison Wesley, 1991.
- [3]: K.J. Hunt, D. Sbarbaro, R. Zbikowski, and P.J. Grawthrop: "Neural networks for control systems – A survey", Automatica, vol. 28, N°6, 1992.
- [4]: K.S. Narendra and K.Parthasarathy: "Identification and control of dynamical systems using neural networks", IEEE transactions on neural networks, vol. 1, N°1, March 1990.
- [5]: L. Jin, P.N. Nikiforuk and M.M. Gupta: "Adaptive tracking of SISO nonlinear systems using multi-layered neural networks", Proceedings of American Control Conference, Chicago, 1992, pp. 56-60.
- [6]: R.H. Nielsen: "Theory of the back-propagation neural network", Proceedings of the international joint conference on neural networks, 1989, vol.1, pp. 593-605.
- [7]: R.P. Lippmann: "An introduction to computing with neural nets", IEEE ASSP magazine, April 1987.
- [8]: Y. Sakawa and Y. Shindo: "Optimal control of container cranes". Automatica, vol. 18, pp. 257-266, 1981.
- [9]: Y. Sakawa, Y. Shindo and Y. Hashimoto: "Optimal control of rotary crane", Journal of optimization theory and application, vol. 35, No. 4, December 1981.
- [10] N. Derbel: "On the Optimal Control of Nonlinear Systems by Neural Networks", Conf. NNSC'2000, Zakopane, Poland, June 2000.
- [11] N. Derbel: "Optimal Control of Nonlinear Systems Using Neural Networks", Second Middle East Symp. on Simulation and modeling, MESM'2000, Amman, Jordan, Septmeber 2000.
- [12] N. Derbel: "On the Optimal Control Approaches of Nonlinear Systems", Int. Conf. on Artificial Computational on Decision, Control and Application, ACIDCA'2000, Monastir-Tunisia, Mars 2000
- [13] A. Ouezri, N. Derbel: "On the Improvement of the Dynamical Behavior of a Rotary Crane Optimal Control, Neuro-Fuzzy Approach". Int. Conf. Mechanical Engineering (ICAME'02), Hammamet, Tunisia, Mars 2002.



**Figure 4 :** Simulation results for the simplified system.



**Figure 5 :** Simulation results for the full system without approximations.



**SIMULATION  
IN  
BUSINESS  
AND  
MANAGEMENT**



# General Systems Theory and Knowledge Discovery in Databases: Two Faces of a same Coin?

Van Welden D. and Vansteenkiste G.  
Ghent University, Applied Mathematics, Biometrics and Process Control  
Coupure Links 653, B-9000 Gent  
Belgium

## ABSTRACT

General System Theory (GST) and Knowledge Discovery in Databases (KDD) have a lot in common. In this paper, this commonality is described from a high abstract level to a more concrete one.

The comparison starts by considering the philosophical aspects of GST and KDD. Next, their life cycle is compared and some emphasis shifts are explained.

## INTRODUCTION

In this paper a model is the key to the comparison of GST and the KDD. It is defined as any formalisation that aids in understanding a system under investigation. This definition is very abstract, which makes it applicable to both domains. Hence, a model in its most general form can be (the list is not exhaustive):

- a set of differential (or difference) equations, a block diagram, a Petri net, a Bond graph, any qualitative model (e.g., confluences), a time series model, ...
- a classification or regression tree, a dependency network, a hypothesis, a regression model (linear, non-linear, logistic, ...), a discriminant function, ...
- a set of rules (if-then rules, association rules, ...), a Markov chain, a look-up table or contingency table, a set of data (for lookup), a clustering model, a neural net, a genetic algorithm, ...

The first types of models are mainly known in the domain of GST, the second type in the domain of KDD, and the third types are commonly used in both domains. The above definition of a model is consistent with the definition of Minsky [1965] that states: “*An object 'A' is a model of an object 'B' for an observer, if the observer can use 'A' to answer questions that interest him about 'B'.*”

## GENERAL SYSTEM THEORY

Although it is well known that “the whole is more than the sum of its parts”, General System Theory has found its origin in the “organismic biology” or “system theory of the organism” of the biologist Bertalanffy (in the late twen-

ties), which was then termed ‘Allgemeine Systemtheorie’. In the late forties he stated that “there exist models, principles and laws that apply to generalized systems or their subclasses irrespective of their particular kind, the nature of the component elements, and the relations or “forces” between them. Of course, other people contributed as well such as Wiener (cybernetics), Ashby, Klir, etc.

The GST point of view is interdisciplinary and has brought a new perspective, a new way of doing science. It provides an organized body of knowledge for dealing with systems in general, in which there are basic concepts, a number of principles, some rigorous enough to be considered “laws,” and, a general framework for theory construction. In this framework, isomorphic schemes play an important role. Hence, GST touches virtually all traditional disciplines, from mathematics, technology and biology to philosophy and the social sciences and includes AI, neural networks, dynamical systems, chaos, and complex adaptive systems.

G Klir [1985] gives a concrete description of the key concept in GST : “*A general system is a standard and interpretation-free system chosen to represent a class of systems equivalent (isomorphic) with respect to some relational aspects that are pragmatically relevant.*” A taxonomy of general systems based on (abstract) relational aspects, which transcends the disciplines of sciences, allows a systematic approach in which the following constant traits of a directed system (with in- and outputs) can be considered:

- A set of external quantities and a given resolution level. In the case of a directed system the external quantities are classified as input and output quantities, and the resolution level is given for both.
- A given activity (experimentally determined or observationally given measurements), where the quantities are classified as input and output quantities for a directed system
- A given permanent behaviour, which is a time-invariant relation that holds between the principal quantities associated with the latest values of output quantities on the one hand, and the set of all the other principal quantities on the other.
- A given UC (Universal Coupling) structure (defined by subsystems, their behaviour, and their coupling). The couplings are directed from output

quantities of one element to input quantities of other elements, including the environment for directed systems.

- A given ST (State-Transition) structure (defined by a set of states and a set of transitions between them). A more rigorous formulation is given in chapter 3. In the case of directed systems, a single stimulus is associated with each transition.

Depending on the type of the system problem, each of the above traits can be used for a basic definition of a system. They also allow defining ‘analysis of systems’ and ‘synthesis of systems’. The former starts from a given (UC) structure to find the corresponding behaviour and/or the ST structure. The latter starts from a given behaviour or activity and looks for a compatible ST-structure and appropriate (UC) structure. As a particular behaviour can be realised by different structures consisting of the same type of elements, further requirements are needed to determine a structure (‘inductive bias’).

## KNOWLEDGE DISCOVERY IN DATABASES

More and more data is digitised and stored in databases thanks to the automation in generating and collecting of data (examples are tax returns, health care transactions, consumer behaviour, industrial transactions, etc.) and the maturity of database technology. The resulting data explosion cannot be processed by humans alone anymore despite the fact that quite some informative patterns may be hidden in the data. Consequently, tools and techniques have been designed that assist humans intelligently and automatically in analysing this abundant amount of raw data to obtain useful information via knowledge discovery in databases (KDD). Summarised, KDD is a data exploration methodology that is defined to be the non-trivial extraction of implicit, previously unknown, potentially useful, ‘relatively simple’, and not predefined information from large databases.

KDD is employed in finance (fraud detection, stock market prediction, credit assessment), CRM (Customer Relation Management) where one tries to see how to improve the targeting efficiency, or cross or up-selling of products, quality control, medicine (effect of drugs, diagnosing, hospital cost analysis), astronomy (cataloguing), molecular biology (finding patterns in molecular structures), text mining, web mining, etc.

*Economical criteria* are ‘les raisons d’être’ of KDD in business enterprises. Ultimately, it should give a good return on investment. Customer retention, fraud detection, risk analysis, management applications and market analysis applications are the driving forces behind KDD. *Practical criteria* express that it must have a potential for significant impact of an application. It should be advisable because no standard methods exist to solve the problem. In addition, there must be domain expertise available. Juridical criteria, such as privacy and tractability issues, play along. *Technical criteria* imply that there should be suffi-

cient data available because many data mining algorithms require large amounts of data.

KDD starts with the *goal definition*, which must be formalised and made executable so that it can be related to relevant data, which are hopefully present in the database. The *data pre-processing* step prepares and reshapes the data for subsequent processing. It involves data and attribute focusing, data cleaning, data projection, and data augmentation. The *data mining* step induces the model. It consists of a model specification, model fitting, model evaluation, and model refinement. *Consolidation of the newly found knowledge* and output generation consists of interpreting and documenting the found patterns, and if more model types were used, comparing them. Any conflicts (if any) that may arise with previous knowledge in the knowledge base must be resolved. A priori knowledge can be used in any step.

These steps can be illustrated with what an archaeologist does, see [Brachman 1996]. When an archaeologist wants to dig up interesting artefacts (goal definition), he looks at the (data) landscape to decide where to dig. This is based in part on what he sees and in part on his experience and background knowledge (focusing). Once at the site, he brushes away the dust (data cleaning and data projection), pieces fragments together that seem to fit (data mining), and decides what to do next in order to confirm an evolving hypothesis about the creator and the meaning of the artefacts (consolidate new found knowledge). Finally, the artefact is presented in a museum (output generation).

Remark the possible impact of different degrees of a priori knowledge in the process. More a priori knowledge helps in all steps to achieve a better (more valid) model (is the artefact a vase, an eating bowl, a religious symbol, etc.?). Remark that the model specification can be purely data driven too (just try to puzzle pieces together). Thus, data analysis and model development is highly intertwined.

Popular methods used in data mining are [Siftware 2002]:

- Bayes rule, Discriminant analysis, Nearest neighbour method, Density estimates, Tree classifiers: CHAID, CART
- neural networks (connectionist approach)
- Concept learning and rules, Tree classifiers: ID3, C4.5, C5.0, Case based reasoning, Genetic algorithms

## PHILOSOPHICAL ASPECTS OF GST AND KDD

### *Induction-deduction and applicability*

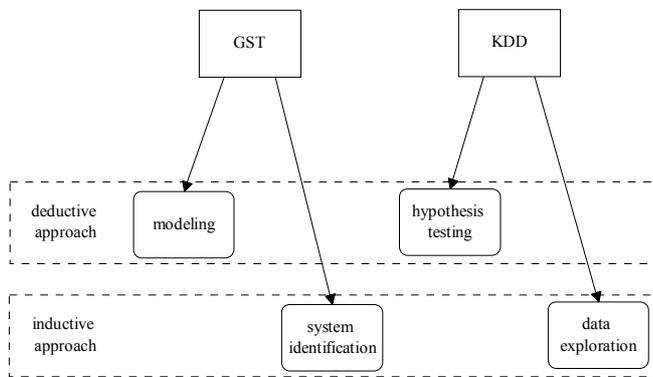
Philosophical issues concern the fact that system observations may be obtained either actively or passively. In KDD, the latter situation is the rule (although ML (Machine Learning) considers both cases), while in GST both settings deserve equal attention. Another, even more im-

portant issue, is the different emphasis and experience with regard to the two main approaches to model construction. These approaches are depicted in Table 1.

	knowledge based approach	data based approach
synonyms	modelling	system identification
	top-down modelling	bottom-up approach
reasoning	deduction	induction
does what?	encodes the (inner) structure of the system	encodes the behaviour of the system (via experimental data)
problem type	Analysis	synthesis

**Table 1 : The two main approaches to model construction**

A model constructed purely deductively (top-down) can generally be considered as a unique solution of modelling. Hence, the top-down approach can and should be used if there is enough a priori knowledge and theory to characterise completely the mathematical equations. In that context, the concept of model structure (not to be confused with system structure) becomes important.

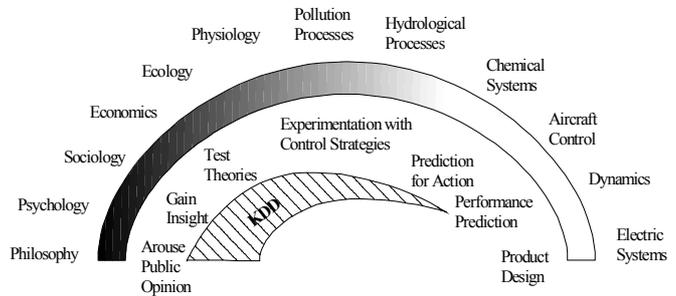


**Figure 1 : Deductive and inductive approach in GST and KDD**

The bottom-up approach tries to infer the structural information from experimental data and to come up with a usable model under a given experimental frame. This approach may generate an infinite number of models satisfying the observed input-output relationships. So, there is no straightforward procedure for determining the structure of a model. A set of guiding principles and quantitative procedures for inferring structure parts from data sets is needed (inductive bias, Michie [1994]). More specifically, it is desirable to have additional assumptions or constraints to help selecting an ‘optimal’ model.

Figure 1 shows the corresponding terminology used in GST en KDD.

The main emphasis difference is found in Figure 2 shows the mapping of KDD onto the GST spectrum. The latter stems from the paper of Karplus [1976].



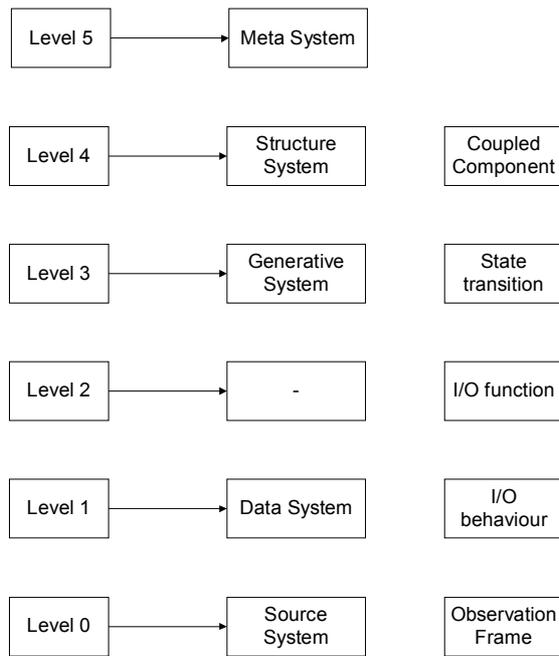
**Figure 2 : Spectrum of modelling**

It shows that KDD can be used for light gray systems (even white systems for fault diagnosis), but its major field of application is towards black box systems.

With regard to the black box aspect in GST, it is interesting to look at the definition of a system given by Klir [1985] earlier in this paper. It allows for both the deductive or inductive approach. From this definition, it can be seen that a class of systems can be uniquely defined by given activity (series of I/O records). This corresponds with a system identification approach in GST (and with the starting point of data exploration in KDD). The first three traits are commonly used in KDD, while the latter two, UC and ST structure (Universal Coupling and state transition structures), are more devoted to GST.

Another way to see the emphasis difference is by considering how knowledge can be organized. Knowledge about systems can be represented in associated hierarchical epistemological levels, see Figure 3. The levels are indicated at the left, the levels according to Klir [1985] in the middle and the levels according to Zeigler [1976] at the right. Source and data systems are predominantly of an empirical nature, while higher epistemological levels are predominantly of a more theoretical nature. The hierarchy of levels can be used for an inductive or a deductive approach.

An inductive approach relates to climbing up the hierarchy. In contrast, a deductive approach signifies descending the hierarchy. As each higher level contains some more knowledge than a lower level, one easily sees that inductive reasoning tries to find more information from facts, while deductive reasoning generates no new knowledge. Going down the hierarchy is relatively simple (deductive), while going up is difficult (inductive).



**Figure 3 : Simplified overview of epistemological level**

At level 0 (source system or observation frame), one defines what could be the constituents of relevant data relative to the object of investigation, the purpose of the investigation and the constraints posed on the system, i.e. time-base, inputs, outputs. At level 1 (data system or I/O behaviour), one does the actual data gathering, related to the source system previously chosen. The data is stored in an activity matrix. Zeigler considers an extra level (level 2), which stores the I/O function. At the next level (level 3), one can start with data processing, where the aim is to discover some support-invariant relationships among the data (here states come into play). Level 4 contains structural (components en coupling) and level 5 meta-knowledge (time-varying structure systems).

GST considers all epistemological levels, while KDD mainly situates itself on the lower 3 levels.

Additionally, a horizontal dimension indicates which types of system problems are defined (methodological problems or types). Klir considers the methodological distinctions as a secondary classification of system problems.

### Validity

The validity of a model plays a crucial role in all data experiments undertaken. Because each model is an approximation of reality, it implies the existence of a model error. The validity of a model is about how well a model represents the original system it stands for or when referring to the model error, validation deals with the evaluation of this error rather than with its existence. It is the process of determining whether or not the model is an acceptable description of reality (model-reality relation). The

validation process itself is valid if the conditions governing the model and the physical system are the same.

Validity can be measured by the extent of agreement between the original system and the model. This is already formalised in Coombs et al. [1954] where the authors describe two different routes: one is by experiment (validation) and the other by logical argument (deductive modelling). Both routes try to arrive at the same conclusions about the real world by different means. The notion of validity is extended by Zeigler [1976], who distinguishes different degrees of validity:

- A model is *replicatively valid* if it matches data already acquired from the original system.
- A model is *predictively valid* if it can match the data of the original system before these data are acquired from the original system. Predictive validity is stronger than replicative validity.
- A model is *structurally valid* if it is not only predictively valid, but also reflects the ways in which the original system operates to produce its behaviour.

Replicatively and predictively validity is important in both GST and KDD, while structurally validity is more an issue in GST.

### LIFE CYCLE OF GST AND KDD

A merged life cycle of GST and KDD is represented in Figure 4. In this life cycle, three major conceptual blocks are to be distinguished:

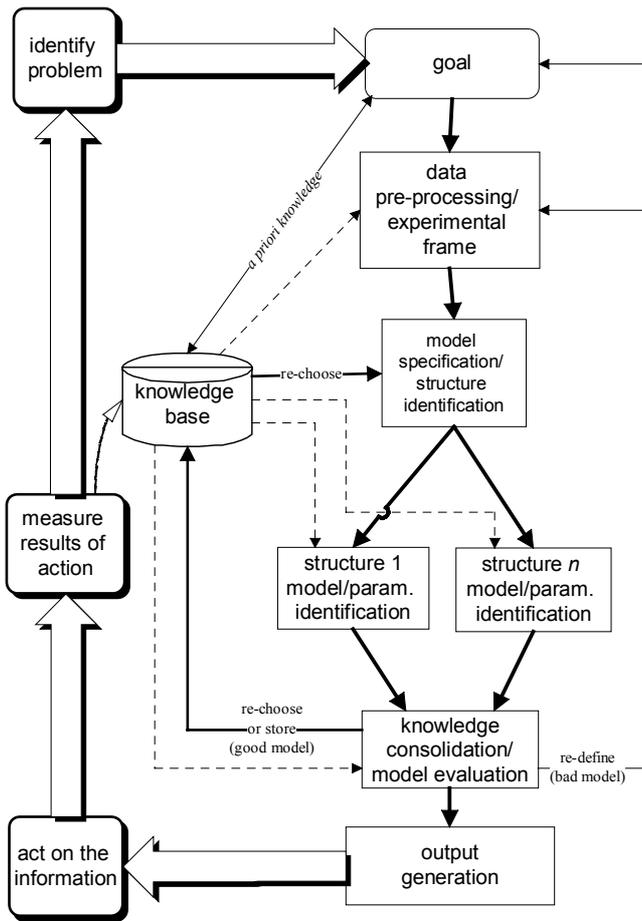
- A knowledge or model base
- The bussiness part
- The modelling construction part

Figure 4 also gives an integrated and broad picture showing the three main societies involved in the virtuous cycle, i.e., the management society, the database society, and the ML-statistic societies (ML and statistics are grouped here).

#### The Knowledge Base

There exist scheme's for modeling and simulation that take into account the re-use of models [Zeigler 2000]. Similar, there is in KDD a virtuous cycle of data mining' that shows its complete enterprise approach to KDD, [Berry 1997]. Figure 4 combines these two in an unifying attempt.

Models of standard components should be saved in libraries (model bases) so reuse is promoted. A model can have any number of components defining the interface and the behaviour. The description of the actual behaviour of a model will be called the models' realization.



**Figure 4 : General modelling and KDD**

It is essential that the model interfaces are standardized in some way by what is called a terminal (or port) and a model may have any number of them. A terminal can be simple or it can be structured, i.e., having a set or an array of subterminals. Terminals can be related to each other by connections. For a connection to be valid the terminals must have a similar structure.

Why do we need a knowledge/model base?

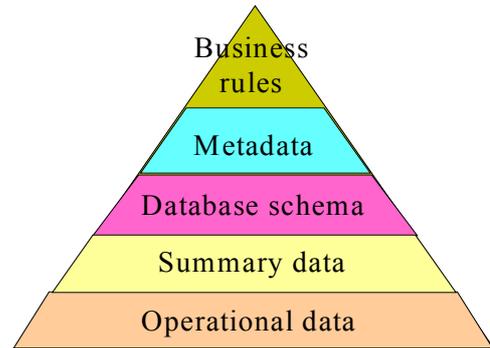
To support model reuse, the models have to be expressed in a language which admits structural decomposition with well defined interfaces between submodels.

To understand a large model is very difficult unless it can be divided into smaller parts that can be analyzed and comprehended separately. Models can be parameterized in order to make them more flexible and adaptable for different purposes.

Every model described by the user is actually a model type rather than a particular instance. Also terminals and realizations should be represented as types rather than instances. When a model is to be simulated, every submodel has to be instantiated.

Model inheritance is a concept entirely different from submodel decomposition. In fact there are orthogonal concepts for describing models in a structured way. A model defined as a subclass will inherit all properties from the super class while it is not appropriate to say that a submodel will inherit properties from its super model.

The concept of meta-data exists in both GST and KDD, but to our knowledge meta-data has more semantic richness in KDD. A first aspect of this richness can be illustrated by looking at the data types in a data warehouse, which are shown in Figure 5.



**Figure 5 : The use of meta data in data warehousing (from [Han 1999])**

Operational data is the data itself. In data warehousing, this also means where it comes from, when it was stored, etc. Summary data gives summaries of the data, so it is a kind of meta-data already. The database schema gives the physical layout of the data. The meta-data level itself is the logical model. The meta-data repository also encompasses business terms and definitions, ownership of data, and charging policies. For example, operational meta-data concerns:

- data lineage: history of migrated data and sequence of transformations applied
- currency of data: active, archived, purged
- monitoring information: warehouse usage statistics, error reports, audit trails

A second aspect of the richness of the term meta-data in KDD is found in the data mining approach. Here, meta-data says something about the measurement scale; whether variables are derived from others (data projection); if there are bounds on data; if there are structural zero's; whether there are distributional assumptions made (parametric versus non-parametric), etc. Meta-data is thus much related to the use of a-priori knowledge in data-pre-processing (structural zero's, derived variables), and in model specification and selection (which data mining method to use). Remark that the distinction between data and meta-data may be blurred.

### *The Business Part*

The ‘management’ part of Figure 4 (identify problem, measure results of action and act on the information) is very general. Thus, it is valid for any modelling attempt, be it via GST or via KDD. It is a process that applies knowledge that is gained from increased understanding of customers, markets, products, and competitors to internal (continuously evolving) processes. It is not the aim of this paper to elaborate on this part.

### *The Modelling Construction Part*

Of course, these concepts will have an impact on the life cycle of modeling. One distinguishes

- the goal formulation in both GST and KDD,
- the data pre-processing versus experimental frame definition,
- the model specification versus structure identification (use of shallow versus deep models)
- data mining or parameter estimation
- knowledge consolidation and model evaluation

### *Goal Formulation*

All problem statements start with the identification of the problem. They try to set a goal in order to solve the problem at hand. To be able to meet a goal, it has to be formalised so that a rigorous approach to problem solving can be obtained. For that purpose, a focus on part of reality has to occur, i.e., one has to define a system. Defining a system is studied in the general systems theory by Karplus [1976], by Klir [1969], and by Zeigler [1976] (list is not exhaustive). It involves defining the boundaries of the system, the interaction of the system with the environment, and last, but not least, defining the relevant variables (space-time resolution). These are concepts that are described in GST, and which are as well applicable to KDD because general system theory is applicable for any model (thus also to data mining).

With regard to goal formulation, more similarities are present: e.g., gaining insight (in KDD) corresponds with understanding (in GST). Both domains stress the principle that a model should not be more complicated than absolutely necessary (Occam’s razor). The three distinguished broad levels of intervention in GST, i.e., management, control and design, are equally applicable to the domain of KDD. Obviously, data mining can be used for gaining insight (model purpose), and the pattern found in KDD provides an opportunity to intervene at some level (as in modelling). However, GST puts some more emphasis on a correct classification and prediction. A speedy or less costly classification and prediction is less the issue in GST (except perhaps in the domain of control theory), but of more importance in KDD. Hence, the ‘Quick decision’ problem type, where accuracy may be less important than comprehensibility, is less stressed in classical modelling. A similar argument applies for the cost of a model.

### *Data Pre-Processing Versus Experimental Frame Definition*

Defining a system has as much to do with the definition of the object system as with an experimental frame, see Figure 3. An experimental frame isolates specific input/output behaviour both of the real system and its model, [Elzas 1984]. In KDD, this corresponds with the data pre-processing step. Attribute focusing is in fact nothing more than defining what are supposed to be relevant variables for the system under investigation. This is consistent with the notion of an experimental frame and with what Klir calls an observation channel. It also fits in the viewpoint, taken in systematic modelling, which states that a model is conceived as a collection of variables and relations among them, [Ören 1984]. The collection of relevant variables is attribute focusing, while identifying the relations is part of the data-mining step. The KDD approach to attribute focusing is related to the philosophical viewpoint of Klir. This can be induced from what Elzas [1984] wrote: “*Imagine that we are able to prove that the model is not sufficiently detailed in its system description for some goal that was set beforehand. Confronted with this situation Klir will remark that the reason for this deficiency is a lack of knowledge*”. Data focusing or sampling data is less used in systems theory, because the nature of data is often different than for KDD. In KDD, one usually starts with a collection of static independent records. Taking a relevant subset poses no major problems. However, in general systems theory, taking a subset of time-depending data is usually not done: data records do depend on each other. This does not mean that GST has no way of dealing with an abundance of data records. A kind of data reduction can be done in GST via the determination of the Nyquist frequency and via filtering techniques. Still, it illustrates the static(KDD) and dynamic(GST) modelling aspects of the respective domains. On the contrary, data projection can be applied for both domains. It can be meaningful to create a new variable, which summarises the behaviour of a set of other variables, and to use that new variable in the modelling phase that follows later (data projection also fits in the definition of an experimental frame).

### *Model Specification Versus Structure Identification*

The data-mining step in KDD is much related to modelling in system theory. Model specification can be compared with model structure identification. It involves deciding what type of model to use (in system theory: Bond graphs, Petri nets, block diagrams, ... in data mining: classification trees, hierarchical clustering, linear regression, neural networks, etc.). The model specification step is straightforward applicable to both domains and can be considered at different abstraction levels. On a very abstract level this involves choosing if one wants to use Bond graphs, Petri nets, block diagram, rules, neural nets, etc. (for systems theory) or trees, clustering, rules, neural nets, regression models, etc. (for KDD). On a more concrete level, one has to further specify the model. Examples are: linear, logistic, or non-linear regression (KDD), linear or

non-linear models (GST), state-space or transfer functions in block diagrams (GST), kind of neural net (both domains), kind of tree (decision, regression, ...) (KDD), order of a differential equation (GST), type of clustering (KDD), etc. In both domains, the goal has a large impact on the used model types: when comprehensibility is more important certain representations may be more preferred than others. For example in GST, a block diagram may be more comprehensible than a Bond graph (it is also relative to the field of expertise). A tree structure is more comprehensible than a neural network (KDD). Rules are more comprehensible than some other model types (both domains), and neural networks are usually the least comprehensible (both domains). The issue about comprehensibility has a lot to do with the greyness of the model: black box models are always less comprehensible than white box models.

The amalgamation of terminology from GST and KDD may shed new light on the terms 'model specification', 'model structure', and 'model complexity'. Model specification has a lot to do with the goal setting, while the model structure is more determined by the experimental frame. Model complexity is related to validation and parameter estimation.

Table 2 shows the terms in relation to each other and it gives an idea of the degree of abstraction that comes into

Model specification	Rules	Differential equations	Tree classifiers	Neural networks	Time series
Model structure	Conjunctive, disjunctive	Linear, non-linear, time-invariant, ...	Classification, regression, survival ...	Kohonen, back-propagation, ...	AR, ARMAX, MA, ...
Model complexity	Rule size	Order	Number of nodes	Number of neurons	Order

**Table 2 : Model specification, structure and complexity**

play.

Another aspect is that the KDD society is used to handle all kind of variables (nominal, ordinal, continuous, etc). The GST society also knows about this taxonomy of variables, but they do not have such a systematic approach to constructing appropriate models for them. In the KDD society (especially statistics), one uses different models for nominal variables, ordinal variables and continuous variables (e.g., linear regression versus logistic regression, chi-squared based models on contingency tables versus other). Furthermore, the size of the data set is used as a guideline too. For small data sets, exact methods are used, while for larger data sets, asymptotic methods can be used.

#### Data Mining Or Parameter Estimation

Model fitting involves parameter estimation (identification) in modelling. It is also called model calibration, [Elzas 1984]. This reduces to estimating parameters or coeffi-

cients in differential equations (GST), parameters in state-space models (GST), regression models (KDD), etc.

Model validation consists of comparing the behavioural data of the system under investigation and the calibrated (fitted) model. Usually, a train-and test method is used. A rule of thumb is that 2/3 is used for fitting the model (training) and 1/3 for validation (testing). Cross-validation is commonly used in KDD, while it is not so popular in GST. Even more specifically, bootstrapping is known too in KDD, but almost unknown in GST. Replicatively validity is known in GST: it consists of fitting the model on the training set. In KDD, this is better known as internal estimates (resubstitution estimates, [Breiman 1984]). Predictively validity is done on a test set; it gives true estimates. What Elzas calls 'realism', i.e., looking at structural isomorphisms at different degrees of lumping of sub-systems (and corresponding sub-models), is not considered in KDD. There is usually only one global level, and the validation takes place for the total (indivisible) system.

#### Knowledge Consolidation And Model Evaluation

KDD uses an interesting function for *evaluating* a model. The evaluation can be based on more than just accuracy performance. For example, KDD can take into account what is economically interesting (cost of model) or it

can take the faster model with regard to prediction. This may prove an important point for the modelling society when they want to evaluate their models in an economic context (cost of modelling), or when speed is of the utmost importance. KDD provides a more general framework for dealing with these situations.

Model refinement is equally applied in GST and KDD; when a model does not validate well, another model (structure) is chosen. When this fails too, one can go one step further back and redefine the experimental frame or even adjust the goal. The refinement of an existing model is a major issue in modelling. The systems theory as developed by Zeigler (use of a SES (System Entity Structure) and model base) has as purpose to construct and refine models, [Zeigler 1976], [Van Welden and Vansteenkiste 1992].

From Figure 4, it can be seen that many models may be used in parallel. The models can be of a different specification (e.g., neural nets and genetic algorithms), and they can

be situated on different epistemological levels (e.g., rules versus decomposed models). In the latter case, one speaks of shallow versus deep models (in GST).

Models are not only evaluated with regard to an interesting function, or validated with regard to a goal setting, but in KDD, model *specifications/paradigms* are also compared with each other. This belongs to the knowledge consolidation step. Here, model evaluation is on another epistemological level than in the data-mining step. Models are not only compared on a test set to validate the parameter estimation, but they are compared on yet another (independent) test (or evaluation) set to evaluate the chosen model specification. The GST community has not done so much work in this aspect as the KDD community, which focused and implemented this right away (see SAS-enterprise miner, [SAS 2002]).

In both GST and KDD, the consolidation with regard to storing the found knowledge is present. In GST, the newly found model is stored in a model base (called modelling in the large, see also [Van Welden et al. 1991]), while in KDD this is not stated so explicitly. Output generation has a lot to do with appropriate graphical representations of the data. In that aspect GST can learn from KDD, because the latter is well acquainted with all kinds of graphs (e.g., multidimensional graphs such as Trellis graphs). Multi-dimensional visualisation techniques are considered very important in KDD (e.g., in statistics, see [Friendly 1991]).

Finally, both GST and KDD acknowledge the necessity of feedback from steps that appear later in the cycle to steps that appear earlier. This is indicated in Figure 4. Therefore, the steps in the life cycle are more intertwined than one should expect at first sight.

#### FURTHER EMPHASIS SHIFTS BETWEEN GST AND KDD

GST focuses more on accuracy of a model, sometimes speed and less the cost of a model. It does not lay emphasis on data reduction as KDD does. Modelling of dynamical systems is the primary focus of research. Decomposition of models is common place.

KDD puts more emphasis on static systems (temporal databases can be used and sequential patterns detected, but this is not the same as modelling dynamical systems). KDD focuses more on model comparison in the large (via a third test set), on data warehousing and relies on a richer semantic meta-data structure. In KDD, one has more experience with very large databases and with high dimensionality problems. Data reduction is commonly used. The interestingness function is more general than the evaluation functions used in GST.

#### CONCLUSION

This paper pleads for a tighter integration of general systems theory and knowledge discovery in databases. GST and KDD put a different emphasis on modelling and their field of application is somewhat different, but the complementary aspect of both domains makes it fruitful to have a good cross-fertilization. Hence, there is room for elaboration from both sides.

#### REFERENCES

- Berry M. J. A., Linoff G. [1997], *Data Mining Techniques For Marketing, Sales and Customer Support*, Wiley Computer Publishing, 1997.
- Brachman R. J., Anand T. [1996], "*The Process of Knowledge Discovery in Databases*", *Advances in Knowledge Discovery and Data Mining*, ed. Fayyad et al, AAAI Press/MIT Press, Cambridge, England, p. 37-58a, 1996.
- Breiman L., Friedman J. H., Olshen R. A., Stone C. J. [1984], *Classification and Regression Trees*. Chapman & Hall, 1984.
- Coombs C.H., Raiffa H., Thrall R.M. [1954], "*Some Views on Mathematical Models and Measurement Theory*", *Decision Processes*, John Wiley, New York, 1954.
- Elzas M.S. [1984], "*System Paradigms as Reality Mappings*", *Simulation and Model-Based Methodologies: An Integrative View*, ed. Ören T.I. et al. NATO ASI Series F: Computer and System Sciences, vol. 10, Springer Verlag, p. 41- 68, 1984.
- Friendly M. [1991], *The SAS System for Statistical Graphics*. Cary, NC: SAS Institute Inc, 1991.
- Han J. [1999], Personal communication, (from a lecture for the IBM chair in Antwerp 1999).
- Karplus W.J. [1976], "*The Spectrum of Mathematical Modelling and Systems Simulation*", *Simulation of Systems*, ed. Dekker L., North-Holland Publishing Company, p. 5 – 13, 1976.
- Klir G.J. [1969], *An Approach to General System Theory*. Van Nostrand Reinhold, 1969.
- Klir G.J. [1985], *Architecture of Systems Problem Solving*. Plenum Press, 1985.
- Minsky M.L. [1965], "*Matter, Mind, and Models*", *Proceedings of the IFIP Congress*, 1, Spartan Books, p. 45-49, 1965
- Michie D., Spiegelhalter D.J., Taylor C.C. [1994], *Machine Learning, Neural and Statistical Classification*. Ellis Horwood, 1994.
- Ören T.I. [1984], "*Model-Based Activities: A Paradigm Shift*", *Simulation and Model-Based Methodologies: An Integrative View*, ed. Ören T.I. et al., NATO ASI Series F: Computer and System Sciences, vol. 10, Springer Verlag, p. 3-40, 1984.
- SAS [2002], see [www.sas.com](http://www.sas.com)
- Sftware [2002], See [www.kdnuggets.com/tools.html](http://www.kdnuggets.com/tools.html)
- Van Welden D., Verweij D., Vansteenkiste G.C. [1991], "*A Proposal for Incorporating Heuristic Knowledge in a Multifaceted System*", *Proc. of EUROCAST 91*, Krems (Wachau), Austria, April 15-19, p. 295-306, 1991.
- Van Welden D., Vansteenkiste G.C. [1992], "*A Mixed Deductive-Inductive Approach to Model Recognition*", *Proceedings of the 1992 European Simulation Multiconference*, York, UK, June 1-3, p. 112-118, 1992.
- Zeigler B.P. [1976], *Theory of Modelling and Simulation*. John Wiley & Sons, 1976.
- Zeigler B.P. [1984], "*System Theoretic Foundations for Modelling and Simulation*", *Simulation and Model-Based Methodologies: An Integrative View*, ed. Ören et al., NATO ASI SeriesF: Computer and System Sciences, vol. 10, Springer Verlag, p. 91-118, 1984.

# A SIMULATION BASED COSTING IN SERVICE INDUSTRY

**Behrouz Zarei,**

*Graduate School of Management and Economics,  
Sharif University of Technology, Tehran, Iran*

*Tel: +98 21 6022755*

*Fax: +98 21 6022759*

Email: [zareib@sharif.edu](mailto:zareib@sharif.edu)

## **Abstract**

Traditional costing methods for services are not accurate enough, therefore the paper proposes activity based costing as the framework and simulation modelling as a paradigm for costing purposes. In this regard, cost of a process can be simulated prior to real world experimentation, by specifying attributes of the arrival jobs and their requirements, and keep tracking of them as they are passing through the process. As a case study, the costing problems of the Iranian Information and Documentation Centre (IRANDOC) are discussed. For process representation, system analysis tools were used. In order to bring cost elements into the model a three-phase method including the cost identification, specification and the allocation was applied. ServiceModel<sup>TM</sup> simulated this process and some aspects of financial analysis including the cost of information item, departmental cost to process a specific amount of information, cost evaluation of training programs, cost of updating and releasing new databases are presented.

## **KEYWORDS**

Discrete event simulation, business processes, activity based costing, database management system

## **Introduction**

Information plays a crucial role in the world of science, technology and business. Information and documentation centres have an important role in the efficient organisation and dissemination of this information, which normally requires many employees and other physical resources. However, current costing methods do not describe them accurately. Because of the inherent *variability* of the required time, the manpower needed and other requirements for processing their entities, along with highly *integrated* data generated by the traditional accounting system.

Traditional cost accounting methods have been criticised by Johnson and Kaplan [1] as too late, too integrated and too distorted to support proper decision making. In deducing the pitfalls of cost

accounting techniques, various new management accounting techniques has been emerged in the mid 80's onward. Some of the many efforts in this field includes balanced scorecard [2], throughput accounting [3], target costing and value chain analysis [4]. In the same direction, Cooper and Kaplan [5] developed Activity Based Costing (ABC) to overcome the pitfalls of the traditional cost accounting. However, they have not addressed the effective implementation issues of this method in a specific environment [6].

Troxel and Weber [7] argued that, without an appropriate computer based model, managing many activities with variation in the cost model elements is extremely time consuming and expensive. Therefore, Spedding and Sun [8] and Takakuwa [9] developed simulation models to support implementation of the ABC in a semi-automated print circuit and in a flexible manufacturing system, respectively. However, the manufacturing environment is more structured than a service industry; making it easier to apply ABC. The paper extended the application of simulation in activity based costing in the service industry, and presents a case study of a database management system to clarify the issue. The ultimate objective of this approach is to produce better understanding of the cost elements and demonstrate a more realistic cost analysis. Briefly in our method, when an entity (*e.g.* a document) leaves each stage of the simulated processes, its relevant cost is calculated. Subsequently the cost of every process is obtained and accumulated as the total cost of processing of the entities. This paper emphasises on the framework, design and implementation issues of the methodology.

## **Cost items**

Traditional accounting systems classify the costs into three groups: direct costs, direct overhead, and indirect overhead. ABC is a procedure that often makes it possible to estimate product costs more accurately than traditional cost systems[10]. In service industry this separation is

more complex than a manufacturing environment. In order to understand the process cost, Enning and Bakker [11] proposed a three-step method including identification of components, specification of the actual costs, and allocation of the costs. These steps in the context of the IRANDOC processes are

1) *Cost identification, specification and allocation*

Before any attempt can be made to analyse the costs involved in PED, the cost components have to be identified. These are as follows:

- ◆ *Computer costs* including hardware and software.
- ◆ *Personnel*, which include operators, technical support, training and secretaries. The costs are wages, social security, pension etc.
- ◆ *Real estate costs* that include maintenance, management of real estate, security, cleaning etc. These costs are often passed on, based on the number of meters in use by certain departments.
- ◆ *Equipment*, which includes particular instruments such as photography machines in the audio-visual department, or printers in the computing department, which used in different stage of the processes.
- ◆ *Overheads* include general management, administration and personnel, the cost of telephone and mail etc. This type of cost is passed on as general overheads per employee to specific departments.

Although this list seems to be comprehensive for our purposes, it can be further aggregated or specified for others. For accurate analysis, specification of the costs is important. To determine how the cost elements are allocated, an Ishikawa diagram is used, to identify the possible causes (here, cost elements) of an effect (the cost of an information record).

In our methodology, the simulation model is used as a basis for calculating the cost as the information records move through the system. The simulation model is elaborated in [12]. In this sense, the costs cascade down to the final department. In this approach, variable overhead costs are traced to individual product which has the advantage of associating many of the costs that are defined in traditional costing systems as fixed overheads. This helps to clarify the relationship between the causes of cost and individual products. Understanding this relationship allows management to differentiate between value added and non-value added activities.

2) *Service time variations*

Beside the complexity of the problem, variation in inputs and service time was another reason to approach the system by simulation. There are at least three sources of variance in various stages of the physical model: inappropriate standards, breakdowns and random variations. Simulation seems to be a proper way of tackling the stochastic aspects of the process where statistical distributions are used to characterise the variation of the model elements.

**Simulation model**

ServiceModel™ [13] was used to develop a working simulation of the IRANDOC processes based on process diagrams and physical models. In the model, cost items (such as equipment and personnel) are defined as variables. Time and equipment requirements in different stages of the process were obtained by conventional work-study and are introduced as the model parameters. The hourly payments for equipment and manpower are defined as model constants and are obtained from current regulations, updated annually by the government. Other cost elements, explained earlier, are also included in the model. The screen layout of the simulation model can be seen in Fig 1. Different departments involved in the process are shown as different rooms, for simplicity, although in the real system, they include several offices.



**Figure 1:** Layout of the simulation model

Whilst the entities move through different departments, their related costs can be obtained. The cost contribution in each stage can be calculated as:

$$C=TR+M \quad (1)$$

Where  $T$  is the time that the resource is captured,  $R$  is the cost rate at which the resource are being used and  $M$  is the cost of the equipment or material used for the process. When the cost contribution of a product is calculated, the cost of all the activities need to be added together. Therefore, the cost of making  $i$ th product can be given as

$$C_i = \sum_{j=1}^n C_{ij} \quad (2)$$

Where  $C_{ij}$  is the cost requires to perform the  $j$ th operation obtained from Eq. (1) and  $n$  is the total number of operations needed to produce the  $i$ th database. Since the model contains a number of random variables, it was run 37 times to reach an acceptable level of confidence for the output results. Because, we were interested in the *steady-state* behaviour of the model, the data collection was postponed until year 4 of each run of the simulation.

### Simulation outputs

As it is stressed earlier, the main aspect of this methodology is to get a more realistic cost values, which in the current costing system used to be more integrated, late and intuitively. However, the simulation enables us to estimate the required cost for different departments, databases, periods or a combination of them (e.g. the cost of a particular database in summer at a specific department) successfully. Simulation supports of the ABC enable us to derive interesting reports that assist managers to understand the cost aspects of the processes. Some of which are as below.

#### 1) Information item cost.

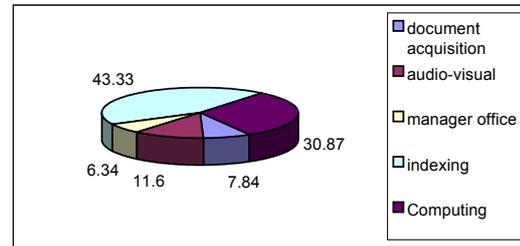
As we have already mentioned, the model can trace a particular entity and take the costs into account when it is passing the processes, which finally provides the total cost for processing a single information item. For instance, the cost of processing a PhD dissertation is broken into the cost elements using the simulation in Table 1.

Resources	Time (min:sec)	%Cost
Employees	150:07	64.54
Computer	47:30	19.46
Manager	10:00	8.18
Machine 1	3:30	2.39
Machine 2	3:10	1.93
Machine 3	2:15	1.54
Copier	1:12	1.31
Printer	3:10	0.65

**Table 1.** Cost elements for an information record

#### 2) Departmental costing.

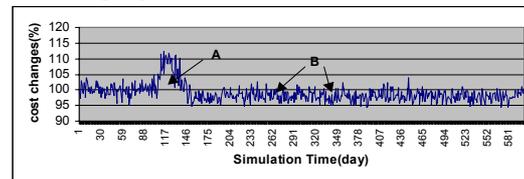
From the performance point of view it is important to determine the amount of money spent in each department. Simulation can help to estimate the required cost in different departments. For instance, the following figure depicts the simulated departmental cost for processing a governmental report.



**Figure 2.** Departmental cost for an information record

#### 3) Training programs evaluation.

Training programs are costly in the short-term but they allow the personnel to do their jobs better, faster and more reliably, which will ultimately decrease the costs. The simulation enables the organisation to determine the effects of new training programs on the cost of the information records. For example, an IT training program in the computing department is evaluated via the simulation. The effects of this program on the cost of an information entry into databases over the time can be seen in the following figure.



**Figure 3.** Effect of a training program on the cost of an information record

The program is supposed to start at day 100 of the simulation and as it can be seen it quickly leads to an above average increase in the cost of an information record. Thereafter, gradually the performance of the employees increases, as it was expected, so that the time and cost require to perform the activities decreases, until they reach a new steady state which is less than the same value prior to the training program. This training program can be evaluated by a cost-benefit analysis through comparing regions A (area between the cost curve and old average from the start point of the training program to the point that the cost is less than the old average) and B (area between old average and cost curve from the point that the cost is less than the old average to the point that the program is effective). Although such a result was expected, but correct estimation of the new steady-state value was of particular importance. This value makes it possible to evaluate the training program which was not clear before the simulation. Also the

model helped managers to recognise the effects of the program on the cost of those departments as they were affected by the rate of output in computing department.

#### 4) Database updating costs.

From the predicted rate of arrival and types of entities for the next period(s) of operation, simulation is able to estimate the cost of updating the databases. For instance, in recent years the universities, as a source of new entities for current databases, have dramatically expanded their activities. This has caused significant changes in the entities of the process, which in some cases leads to major shifts, such as expanding the capacity of some of the information processing institutes. This ultimately affects the cost of information in different levels. The simulation model generates various cost reports for possible scenarios of entities. This supports managers with a more realistic view of what may they face in the future.

#### 5) New database appraisal.

A typical contract in the centre is to produce databases for other organisations. In which case, the organisations usually send their proposal, and the cost of establishing the database is decided in the board of directors. But the accurate cost for this service is often unclear. By considering the attributes of the prospectus items, required processes and related costs, the simulation model is able to estimate the relevant cost more accurately. The simulation model is tested for a new chemistry database and it supports the executives to make an accurate decision about the cost of the proposal.

### Conclusions

It is argued that a simulation model can be used in order to establish activity based costing of these processes. As a case study, the information and documentation centres with huge number of entities complex processes including many stochastic elements for establishing and updating the databases were studied. In our case the model is used to respond to some cost-oriented questions. It is argued that a simulation model could consider the effects of stochastic elements and handle the complexity of the system which, ultimately may leads to a better projection of the problem. ServiceModel<sup>TM</sup> simulated this process and the results enabled us to outline some aspects of financial analysis to fix the traditional costing problems. Some instances of the simulation results are discussed in the paper. Our experience has already shown the importance of simulation, to some extent, in fixing the above problems. This model may be used by similar organisations for analysis of activity based costs.

---

### References:

- [1] T. Johnson and Kaplan R. 1987 "*Relevance Lost*", Harvard Business School Press, Boston, MA.
- [2] R. Kaplan and Norton D. (), "The balanced scorecard - measuring that drive performance", *Harvard Business Review*, **70**(1), pp. 71-79.
- [3] Waldron D. and Galloway D. 1994 "Throughput accounting 1-3, *Management Accounting*", Vol. 66, No. 10-11 and **67**(1).
- [4] Cooper R. and Slagmulder R. 1997 "*Target Costing and Value Engineering*", Productivity Press.
- [5] Cooper R. and Kaplan R. 1988 "Measure costs right: Make the right decisions", *Harvard Business Review*, pp. 96-103.
- [6] Zuk J., et al. 1990 "Effective cost modelling on the factory floor: Taking simulation to the bottom line", *Proceedings of 1990 Winter Simulation Congress*, pp. 590-594.
- [7] Troxel R. and Weber M. 1990 "The Evolution of Activity-Based Costing", *Journal of Cost Management* **4**(1), pp. 14-22.
- [8] Speding T. and Sun G. 1999 "Application of discrete event simulation to the activity based costing of manufacturing systems", *Production economics*, **58**, pp. 289-301.
- [9] Takakuwa S. 1997 "The use of simulation in activity-based costing for flexible manufacturing Systems", *Proceedings of 1997 winter simulation conference*, pp. 793-800.
- [10] Cooper R. 1996 "Activity-based costing and the lean enterprise". *Journal of Cost Management*, **9**, pp. 6-14.
- [11] Enning J. and Bakker A. 1995 "Analysis of Costs of Information Systems", *Studies in Health Technology and Informatics*, pp. 87-97.
- [12] Zarei, B. 2001 "Simulation for Business Process Re-engineering: A Case Study of A Database Management System", *Journal of Operational Research Society*, **52**(12).
- [13] -, 1995 "Service Model User's Guide", PROMODEL Corporation.

### AUTHOR BIOGRAPHY

Dr. Behrouz Zarei is graduated from Management School of Lancaster University (England) in 2001. He is assistant professor of management science in the Graduate School of Management and Economics in Sharif University of Technology. He is working on business process reengineering, management control systems and electronic government implementation.

# MODELLING OF THE INFLUENTIAL FACTORS OF FOREIGN DIRECT INVESTMENT IN THE UNITED ARAB EMIRATES' MANUFACTURING ENVIRONMENT

Mohamed Aljunaibi  
Val Vitanov  
School of Industrial and Manufacturing Science  
Cranfield University  
Cranfield, Bedfordshire MK43 0AL  
England  
Tel +44 (0) 1234 750111 ext 2384  
Fax +44 (0) 1234 752159  
E-mail: mohdaljunaibi@hotmail.com  
v.vitanov@cranfield.ac.uk

## KEYWORDS

Fuzzy logic, fuzzy modelling, foreign direct investment.

## ABSTRACT

This paper aims to develop a fuzzy logic based methodology for modelling foreign direct investment (FDI) in the manufacturing multinational companies (MNC's) operating in the United Arab Emirates (UAE). The intention is to investigate the country's general attractiveness to foreign investors through the identification of the influential factors for foreign investors and the assessment of the investment environment in the UAE. Fuzzy inference models have been used to describe the relationships between different influential factors and selected policies. Because of the nature of the expert information that is currently available, fuzzy sets and approximate reasoning have been used to obtain a global 'good enough' solution. Fuzzy rules have been developed on the basis of the expert knowledge of FDI obtained through extensive literature sources and interviews with experts from the field. The fuzzy model is designed to be consistent with evaluation process conducted by foreign investors. It integrates factors that influence foreign investors decision in order to obtain an estimate that will reflect the general attractiveness of the UAE's investment climate. It could be used further to select optimal policies for further improvement.

## INTRODUCTION

In the globalised world economy of the twenty-first century, the world market for foreign investments has become more competitive. FDI occurs when a foreign investor or entity acquires an asset in a certain country with the intent to manage that asset. The management dimension is what distinguishes FDI from other types of investment.

FDI can be an effective contributor not only to economic growth, but it is also important to management skills, transfers of technology and a higher standard of living.

Therefore, developing countries have made considerable efforts over the past decade to improve their investment climate by offering a wide range of investment incentives (government promotional policies). This paper attempts to use a fuzzy logic based model to investigate the relation between foreign manufacturing corporations and the host country.

## United Arab Emirates

The UAE is a country with one of the world's highest per capita incomes due to substantial revenues arising from crude oil production. Oil revenues have enabled the UAE to transform its economy from a simple localised economy into a developing one with great potential to diversify and grow. At present, the main problem faced by the U.A.E arises from its dependency on oil as a dominant source of national income because it is an exhaustible natural resource. This is a strong incentive for the UAE to find new options for income generation. In the above context, the UAE government is steering the economy towards expanding the manufacturing sector and attracting foreign investments in order to strengthen and balance its economy. To achieve that, the UAE has to be more consistent with modern trends and practices in the world economy. FDI appears to be an appropriate vehicle towards achieving this goal.

## FUZZY INFERENCE SYSTEMS

Fuzzy set theory was introduced for the first time by [Zadeh, 73], as an appropriate tool to deal with linguistic variables containing uncertain or vague information. The fuzzy set theory uses an expert opinion of the membership functions to express the actual value of any variable with values between [0, 1]. [Mendel, 95] explains the concept of a fuzzy logic system as a nonlinear mapping of inputs into crisp outputs.

Fuzzy logic appears to be a scientific tool that can deal with the dynamics of the complex relation between foreign

investors and host countries without a detailed mathematical description. Instead, expert knowledge is used in the form of membership functions and fuzzy rules to develop a fuzzy inference system (FIS). In this process, real values are transformed into linguistic values by an operation called fuzzification which transforms the crisp values into fuzzy sets, and then fuzzy reasoning is applied in the form of IF-THEN rules. A final crisp value is obtained by defuzzification which is the operation that transforms back fuzzy sets into crisp outputs.

When developing investment models, the information that is available from the experts is expressed in linguistic variables. A linguistic variable can be defined as a variable whose values are not numbers, but words or sentences in a natural or artificial language. The concept of a linguistic variable appears as a useful means for providing approximate characterization of phenomena that are too complex or ill-defined to be described in conventional quantitative terms [Zadeh, 75]. By using fuzzy logic, values of a linguistic variable can be quantified and extended to mathematical operations. This transformation from crisp to fuzzy allows to manipulate the information available in a simple and efficient way instead of using sophisticated mathematical techniques.

For the purpose of this investigation, variables were divided into fuzzy sets. For every variable, the adequate number and shape of the fuzzy sets had to be chosen on the basis of the expert knowledge available to the authors. Although any number and any shape of fuzzy sets are possible, an attempt was made to keep the model as simple as possible. In order to define the fuzzy sets for each fuzzy variable, membership functions for the linguistic values are used. For example, for the fuzzy variable "services and infrastructure", the variable set  $A = \{VP, P, F, G, VG\}$  denoting 'very poor', 'poor', 'fair', 'good', and 'very good', have been selected as an appropriate representation. The membership functions of the linguistic values are shown in figure 1.

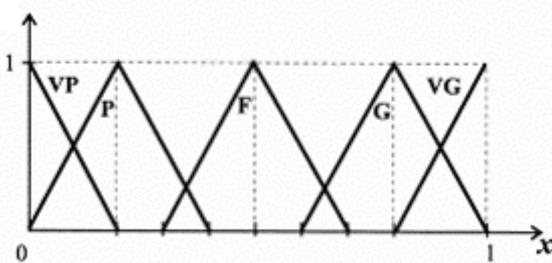


Figure 1: Membership functions for linguistic values

Similar membership functions have been previously suggested in research articles pertaining to application of fuzzy sets. The triangular membership functions have been chosen for application considering their intuitive representation that facilitates the knowledge acquisition and ease in computation. According to [Pedrycz, 94] triangular membership functions can approximate most of the practical situations.

A crisp numerical output of the fuzzy model can be achieved through defuzzification which forms the final operation that generates a single crisp value. There are several defuzzification methods [Driankov, 96], however, for the purpose of this research, the centre-of-gravity method which is compatible with our Mamdani inference scheme is chosen.

Fuzzy logic has been used before in several economic applications. For example [Basu, 84] introduced fuzzy revealed preference theory. Fuzzy logic has also been used to investigate the relation between interest rates and investments in Sweden by [Lindstrom, 97] and more recently [Facchinetti, 01] developed a fuzzy expert system for valuing strategic investments in Italy. One important advantage of the fuzzy approach is that it uses linguistic variables to perform computing with words. In this way, some important factors that otherwise cannot be quantified such as motives and opinions can be included into the model. Therefore, fuzzy logic performs successfully in modelling human knowledge and expertise [Zimmermann, 91].

### Foreign Direct Investment Modelling Methodology

As a first step, an extensive review of the main influential factors of FDI in UAE was carried out. In addition, several contacts were established with some experts in the Ministry of Finance and Industry (MOFI), the Ministry of Planning (MOP), Abu-Dhabi Chamber of Commerce and Industry (ADCCI), the Emirates Centre for Strategic and Scientific Research (ECSSR) and the Gulf Organisation for Industrial Consultants (GOIC). Then, a number of experts and foreign investors in UAE were interviewed in order to extract the main factors that were considered fundamental in the foreign companies' decision to invest. Data analysis and human expertise were both utilised in this research.

To effectively identify the influencing factors or variables, a number of possible influential parameters have been selected by taking a heuristic method based on expertise and common sense knowledge. Six principal factors have been selected in our model, namely, motives, determinants, incentives, restrictions, services/infrastructure and the availability and quality of resources. Note that our selection may be restricted by the availability of the data source, we shall design the model flexibly so that more factors can be adapted if necessary and possible. The model aims at assessing the FDI decisions in UAE in terms of six principal key factors. The final output is a crisp value between 0 and 100 that gives an indication of the foreign investor's propensity to invest.

The factors used in this model depend on the properties of the foreign company and the host country. FDI motives according to [Eiteman, stonehill and Moffett, 95] consist of a wide and complicated set of strategic, behavioural and economic considerations. Mathematical modelling of this problem is difficult, since the interactions between the foreign companies and the environment are uncertain, and

complex. Hence, an approach based on expert knowledge was applied to develop a rule-based fuzzy model. The model was developed in such a way that it can be expanded later. Six main factors have been identified and considered as principal drivers when making an investment decision. The factors included in the model are:

- Level of company's motives such as internationalisation factors and marketing motives.
- Level of host country determinants such as market size, growth potential and political stability.
- Investment incentives which are the host government's policies that motivate foreign investors to favour their countries.
- Investment restrictions which are the requirements or mandates that host government enforce on foreign investing companies.
- Services and infrastructure of the host country.
- Investment resources such as labour and raw materials.

These principal factors were chosen as a main set of factors considering that each foreign company has its own set of reasons which may or may not include the full set of factors.

The principal drivers were connected adopting a fuzzy modular approach as shown in figure 2.

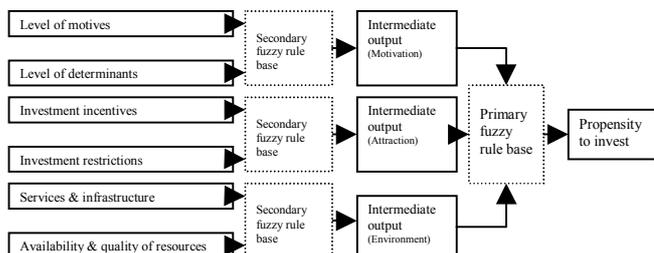


Figure 2: Basic configuration of the proposed model

The intermediate outputs are determined from subsets of the principal drivers. The value of the intermediate outputs is determined by a rule block which represents the core of the fuzzy model. The model is expected to come up with a value that reflects a measure of the propensity to invest as a function of six principal factors.

## DEPENDENCY ANALYSIS, RESULTS AND DISCUSSION

This step of the methodology includes validation and verification of the model. Fuzzy logic calculations have been done with the aid of MATLAB's fuzzy logic toolbox [The Mathworks, 95]. The intention was to demonstrate the advantages of using fuzzy interface systems when modelling investment decisions under the circumstances of uncertainty created by the limited amount of data and differences in experts' opinions.

As a first step, the overall correctness of the fuzzy rules with respect to the background expert knowledge was investigated. This stage identifies possible errors in the rule base and examines the reasoning paths of the inference mechanism. Several fuzzy operators and different shapes of the membership functions have been investigated. In the second step, a dependency analysis was performed by considering different situations that might be encountered in real life and finding out the model's output. This has been done by changing the values of the input factors and observing their effects on the model's output.

Figure 3 shows an example of the fuzzy reasoning of our model between intermediate output 2 and intermediate output 3 as well as the final output.

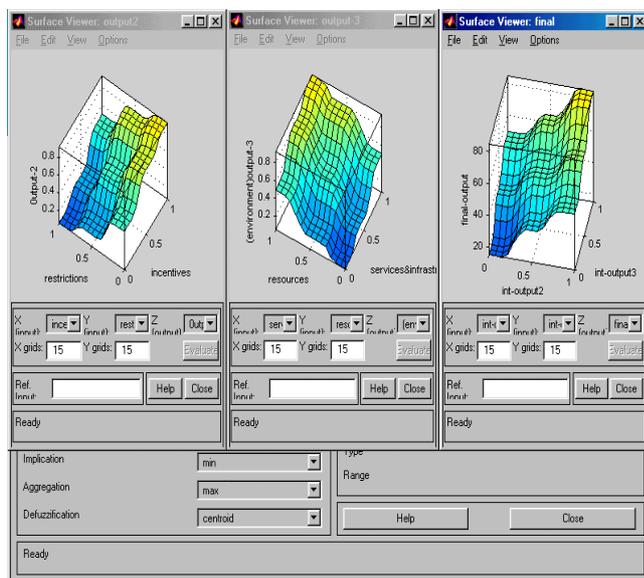


Figure 3: Example of the fuzzy reasoning results

The analysis has been conducted by varying one or more of the inputs and observing the effect on the intermediate outputs and the final output. Only two of the simulations have been presented and discussed in the present paper. At first, the input variable (Motives) has been varied from a low input value of 0.1 (case 1 in table 1) to a high input value of 0.9 (case 5 in table 1) while keeping the other input variables fixed as shown in table 1. The value 0.5 represents a medium value while the input values 0.3 and 0.7 represent a medium-low and medium-high values respectively. It can be noticed that the output value increased from 37 to 67.2 which means that as the value of the variable (Motives) increases, the value of the output increases too. This result gives us a partial confirmation of the model meaning that the output (Propensity to invest) is increasing with respect to the input variable (Motives). In this way, it becomes possible to understand how the model works by observing the result of our simulation and the variations in the intermediate outputs and the final output.

Table 1: Simulation where the input (motives) is varied and others fixed

Input	1	2	3	4	5
<b>Motives</b>	0.1	0.3	0.5	0.7	0.9
<b>Determinants</b>	0.5	0.5	0.5	0.5	0.5
<b>Incentives</b>	0.4	0.4	0.4	0.4	0.4
<b>Restrictions</b>	0.7	0.7	0.7	0.7	0.7
<b>Services&amp;infr.</b>	0.9	0.9	0.9	0.9	0.9
<b>Resources</b>	0.3	0.3	0.3	0.3	0.3
<b>Int. output 1</b>	0.2	0.37	0.5	0.63	0.8
<b>Int. output 2</b>	0.36	0.36	0.36	0.36	0.36
<b>Int. output 3</b>	0.67	0.67	0.67	0.67	0.67
<b>Output Propensity to invest</b>	37	47.7	51.1	54.3	67.2

In the second simulation, the output has been observed as a function of the input variable (Restrictions) with the other input variables fixed. As in the previous simulation, the variable (Restrictions) value will be varied from a low input value of 0.1 (case 1 in table 2) to a high input value of 0.9 (case 5 in table 2). As it can be observed from table 2, the output (Propensity to invest) decreased from 85 to 48.1 indicating that the output is decreasing with respect to the input variable (Restrictions).

Table 2: Simulation where the input (Restrictions) is varied and others fixed

Input	1	2	3	4	5
<b>Motives</b>	0.3	0.3	0.3	0.3	0.3
<b>Determinants</b>	0.3	0.3	0.3	0.3	0.3
<b>Incentives</b>	0.5	0.5	0.5	0.5	0.5
<b>Restrictions</b>	0.1	0.3	0.5	0.7	0.9
<b>Services&amp;infr.</b>	0.8	0.8	0.8	0.8	0.8
<b>Resources</b>	0.9	0.9	0.9	0.9	0.9
<b>Int. output 1</b>	0.35	0.35	0.35	0.35	0.35
<b>Int. output 2</b>	0.8	0.64	0.5	0.35	0.2
<b>Int. output 3</b>	0.92	0.92	0.92	0.92	0.92
<b>Output Propensity to invest</b>	85	64.3	60.7	57.7	48.1

It can be noticed that the intermediate output 2 has been affected by the variation in (Restrictions) and in turn the intermediate output2 affected the value of the final output.

## CONCLUSION

- Fuzzy logic can capture the expert knowledge for the rather complicated economic environment.
- A fuzzy modelling methodology to provide a good measure of the propensity to invest in a certain environment has been developed.
- The proposed model is capable of dealing with qualitative variables such as (level of motives) and (investment incentives).

- The propensity to invest can be determined by using the main influential factors which are motives, determinants, incentives, restrictions, services/infrastructure and the availability and quality of resources.
- The expert knowledge of FDI can be represented using four hierarchically organized fuzzy rule bases.
- Fuzzy logic appears to be an appropriate tool for FDI problem because a complete knowledge of the complex relationship between foreign multinational companies and their host countries is not available
- The developed model appears to be useful for both host governments and foreign investors.
- The fuzzy model can be modified easily following changes in the investment environment. Modification of the model by the introduction of new variables is an area for future research.

## REFERENCES

- Basu, K. 1984. Fuzzy Revealed Preference Theory. *Journal of Economic Theory* 32, no. 2: 212-27.
- Bolloju, N. 1996. Formulation of qualitative models using fuzzy logic. *Decision Support Systems* 17, no. 4: 275-98.
- Driankov, D., H. Hellendoorn, and M. Reinfark. 1996. *An introduction to fuzzy control*. New York: Springer.
- Dubois, D., and H. Prade. 1986. Fuzzy sets and statistical data. *European Journal of Operational Research* 25, no. 3: 345-56.
- Eiteman, D. K., A. I. Stonehill and M. H. Moffett (1995), *Multinational Business Finance*, Seventh Edition, Addison-Wesley.
- Facchinetti, G., C.A. Magni, G. A. Mastroleo, and M. Vignola, 2001. Valuing strategic investments with a fuzzy expert system: an Italian case. *IFSA World Congress and 20th NAFIPS International Conference* volume 2: 812-817.
- Lindstrom, T. 1998. A fuzzy design of the willingness to invest in Sweden. *Journal of Economic Behaviour & Organization* 36, no. 1: 1-17.
- Mamdani, E. H. 1974. Application of fuzzy algorithms for control of simple dynamic plant. *Proceedings of the Institution of Electrical Engineers* 121, no. 12: 1585-88.
- Mendel, J. M. 1995. Fuzzy logic systems for engineering: a tutorial. *Proceedings of the IEEE* 83, no. 3: 345-77.
- Pedrycz, W. 1994. Why triangular membership functions? *Fuzzy Sets and Systems* 64, no. 1: 21-30.
- The MathWorks Inc. 1995. *Fuzzy logic toolbox*. USA.
- Zadeh, L. A. 1975. The concept of a linguistic variable and its application to approximate reasoning. I. *Information Sciences* 8, no. 3: 199-249.
- Zha, X. F. 2000. An object-oriented knowledge based Petri net approach to intelligent integration of design and assembly planning. *Artificial Intelligence in Engineering* 14, no. 1: 83-112.
- Zimmermann, H.J., 1991. *Fuzzy Set Theory and Its Applications*. Kluwer, Boston.

# SIMULATING ELECTRICAL UTILITY CUSTOMER CENTER USING GIS

Nejad Mehic  
Information Research Center  
University of Bahrain  
PO Box 32038, Bahrain  
E-mail: [nedzad@batelco.com.bh](mailto:nedzad@batelco.com.bh)

Abdulla Saad Al-Huwaihi  
Ministry of Electricity  
Electricity Distribution Directorate  
PO Box 2, Bahrain  
E-mail: [alwasat\\_bh@hotmail.com](mailto:alwasat_bh@hotmail.com)

## KEYWORDS

GIS, geodata, distribution network, landbase map

## ABSTRACT

The paper describes a method simulating an electrical utility customer center operations in a typical town in Bahrain. As modern technologies arise, power utilities are seeking new methods to improve system reliability, power quality, productivity of day-to-day operations, and customer satisfaction. In the era of restructuring and modernization of electric utilities, the application of Geographical Information Systems-GIS technology in the power industry is growing and covering several technical and management activities. The integration of GIS with existing power analysis tools is tremendously improving planning and operation of the system. The simulation software will demonstrate abilities of GIS technology in planning electricity-generating facilities management of complains and related problems. This technology has many advantages over conventional method in recording, monitoring and managing information about customer's electrical installations, low voltage distribution line and transformers, high voltage transmission lines, and power generation stations.

## INTRODUCTION

Power management is an important part of any country's infrastructure. Huge amount of data is required to maintain the transmission lines through the terrain (Helmer 2001). Powerline characteristics are collection of data elements that describes the physical characteristics of the line, its current condition, and that which is being transported through it. Distribution systems constitute the link between electricity utilities and consumers. The physical characteristics of the lines are determined during construction represent dynamic data elements required for engineering planning and other analysis (Powershop 2002). The data elements which with both static and dynamic components represent a difficult challenge for proper maintenance and operations in a utility sector.

For efficient and reliable operation of a distribution system, a reliable communication network is required to facilitate project coordination of the maintenance and fault activities of the distribution system. Outages can be isolated faster than even before and maintenance crews

dispatched with critical information including location of the fault.

Major feature of efficiency is automation of industrial processes applied to data gathering, integration and processing, production of complete topographic products and their customized presentation, analysis and interpretation. This often includes modern data gathering remote sensing devices (Vann 2000). Utilities need detailed information about the location and condition of their transmission lines and ways to efficiently maintain and service them. To accomplish this, utilities need to regularly inspect and collect accurate spatial data of their facilities. Also the data must be incorporated in an enterprise system, which would support planners and managers in all the phases of utilities management.

Efficient functioning of this segment of the utility is essential to sustain the growth of power sector and the economy. However, in some systems this is characterized by unacceptably high losses (both technical and commercial), inadequate quality and reliability of supply, billing and revenue collection, frequent interruptions in supply and resultant consumer dissatisfaction etc. In these cases segment of power sector needs immediate attention and action to achieve a turn around and self-sustenance of power sector (Kumar and Chandra, 2002). Many recent studies conducted in the utilities have indicated that the data documentation in most of the utilities must be improved. The data of distribution systems is maintained through hand-drawn maps with facilities data printed in text form on them and available with the linesman in charge of the feeder. These maps are rarely updated.

## ROLE OF GIS IN MANAGING UTILITIES

Geographic Information Systems-GIS is an important tool managing in this area. GIS can be used in distribution systems management for handling customer inquiries, fault management, routine maintenance, network extensions and optimization, analysis and network reconfiguration, and improved revenue management (Harder 1999). GIS is a system of mapping of complete electrical network including low voltage system and customer supply points with latitude and longitudes overlaid on satellite imaging and/or survey of India maps. Layers of information are contained in these map representations where the first layer corresponds to the distribution network coverage. The second layer could correspond to the land background

containing roads, landmarks, buildings, rivers, railway crossings etc. The next layer could contain information on the equipment such as poles, conductors transformers etc. Most of the electrical network/equipment have a geographical location and the full benefit of any network improvement can be had only if the work is carried out in the geographical context. Business processes such as network planning, repair operations and maintenance connection and reconnection has also to be based around the network model. Even while doing something as relatively simple as adding a new service connection, it is very important to know that users of the system are not affected by this addition. GIS in conjunction with system analysis tools helps to do just this. (Kumar and Chandra, 2002).

The power and flexibility of GIS come from two characteristics: it is a computer-based system that operates with geographically referenced data (geodata) and it allows a user to attach information to each object in a table and refer to this table whenever required. The introduction of a GIS allows showing where lines have been put up, which lines connect which customers as well as background information like plots, houses etc. GIS also allows the user to perform complex analyses, making it a powerful tool for data query and management. Moreover, the interaction between the geographic and circuit information gives access to more usable information. GIS is an important tool where the value of visual feedback is used to supplement the detailed result tables coming from circuit analysis tools. Geographic based circuit maps allow user to improve circuit performance through more accurate placement of capacitors and this helps reduce power losses and improve operating voltages, Figure 1.

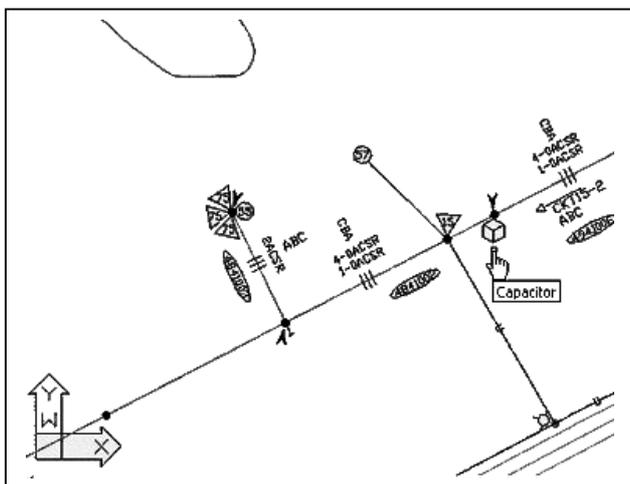


Figure 1 A section of a utility map

Moreover, GIS and circuit analysis tools such permits analyst to test various circuit configurations and view the results on a color-coded graphical display, and other graphical representations such as graphs. Having a visual display of circuit conditions make the analysis process quicker, easier, and more accurate. Furthermore, relational database access and object-oriented programming allow applications to be highly integrated and quickly developed.

## CUSTOMER CENTER

GIS technology is being used to support the planning and sustainable development of electricity-generating facilities. In developed and developing countries, distribution and access routes is critical to ensure effective infrastructure management and development (Townes 1994). The supply network can be seen as a network of wires, transformers and other components that spread out from the generating station and connect to households and factories across the town via substations to regulate and alter voltage as required. The electricity provider has to ensure that the entire network is operational at all times, regulate connections, and monitor their consumption for operations as well as billing purposes. There are engineers in the operations unit who monitor the current network load and demand as well as current status of generation units (Marmar, 1999). The electricity provider has to ensure that the entire network is operational at all times.

In the fault and customer service center, on daily basis, service technicians, engineers, and telephone operators together with a dispatcher record and process calls for help from across the city where each call is located on the map . For every complain call received, a customer complaint form is filled by the telephone operator on duty with the details given by the caller (e.g. no supply, defect, third party damage, fire, arson etc). After that, the field crew is contacted to investigate the complaint and any related problem. The crew will then try to find the source of the problem starting from a single unit (e.g. building) and than analysis a common component for all units (e.g. a substation) in the higher hierarchy. The identification of the problem automatically leads to selection of the crew, if available that is able to solve the problem. If the crew is not available, the caller has to wait until the crew is available. Upon investigation of the complaint on site, the field crew will inform the telephone operator about the nature of the defect/fault and the works to be done and the complaint form is completed. On the other hand, the planning section is responsible to design the plan for a new layout in a town and has to decide whether the current main network can support the extension with or without modification. The huge task of effective infrastructure management and development is much easier in every component of the network (every line, pole, meter) is available and trackable in a digital database.

The customer fault complaint center must be also able to handle supply outages (forced-emergency and planned) complaints and related problems. Forced outages are due to sudden failures of some components of the network or due to the third party damages which cause power interruptions to concerned customers. Planned outages are due to preplanned works to be done on the network of new connections or alternations of expansions of the network and the customer must be informed well in advance of the shutdown date. The fault report center must deal with the outages and customer calls on a 24 hour basis with the help of field crew and telephone operators who work on shift system.

## SIMULATION SOFTWARE

In order to prove capabilities of GIS technology for electrical utilities applications, a simulation software package is planned to model the electrical utility operations. The task at this stage is to identify the exact requirements of software and hardware combination which will satisfy the systems needs and to suggest a way of converting available landbase maps into digital format. It is important to note that at the present, the digital coverage of the utility distribution network is a part of the total existing utilities network all over the country, so it is essential to capture and digitize the additional network connectivity data. Using existing digitized maps, the electrical network must be modeled with precise positing of substations, feeder, transformers, switches and so on with individual attributes on each of these entities (Sharma and Khare 1999). The next step should be to integrate all software and associate data for producing consistent reporting.

Utility companies must consider different ways to define technical specifications for GIS data and related implementation schedule (Balakrishnan 2002). This includes considering impact of technology on the cost of data acquisition (especially important with large scale maps), impact of maturity of GIS data vending industry (it is expected that this industry will drive down the price of GIS data and increase the quality of data quality), impact of better and cheaper GIS software, and cost of upgrading the landbase at the future date. In general, the GIS data is made of two components, geometry features like street boundary, building features etc (geometric data) and information associated with the geometrical data like the locality name, street name etc.(attribute data).

A common practice (Piplapure 2002) to create models to describe real life problems. Models are generally created as the scaled down versions of the existing systems. The developed model allows visualization of the real-life situations, possibility to study the system and to analyze the effects of real-life changes that may take place. Dynamic models are generally used to simulate, study, and analyze engineering environments which normally undergo frequent changes. One example of situation where dynamic modeling is required is with performance studies of electricity distribution networks related to customer complaints. These networks have to operate under constantly varying conditions. To create dynamic models, the simulation software must have built-in graphical capabilities to describe the existing conditions and to quickly present the changes in these conditions. Also, the software must have extensive libraries with a set of symbols and intelligence to attach them on each of network elements. The power network consists of elements such as transformers, switches, capacitors and they are connected together

The simulation software package is design to maintain geographical information and records on network connection and maintain a customer/service unit records

containing customer's name, address, ID number, electricity meter, utility service size, installed load, maximum demand for a part of a typical town. An electronic map of a part of selected town was provided with spatial features stored in a coordinate system, which references a particular place on the earth. Descriptive attributes in tabular form are associated with spatial features. Spatial data and associated attributes in the same coordinate system are layered together for mapping and analysis. GIS and circuit analysis tool permits analyst to test various circuit configurations and view the results on a color-coded graphical display, and other graphical representations such as graphs. Having a visual display of circuit conditions make the analysis process quicker, easier, and more accurate. Furthermore, relational database access and object-oriented programming allows applications to be highly integrated and quickly developed.

The simulation software should have to following capabilities:

- a) Allow users to use equipment symbols to model the network
- b) Allow users to create their own symbols and or exit the existing
- c) Provide users with pull down menus for equipment types selection and position
- d) Allow users to annotate the equipment attributes using pull down menus which contains ratings of various types of standard equipment with default values
- e) Allow user to assign unique numbers (Ids) to the equipment
- f) Provide built-in facility to check correctness and reliability of the model, checks electrical connectivity of the equipment and check proper placement of nodes and equipment

The simulation software is able to accomplish the following tasks on a given area:

1. Trace and locate customers and utility network linkage
2. Maintain and identify the records of different types of important customers such as hospitals, airports, water/sewage pumping stations, government buildings and establishments, VIP premises, traffic light, factories and companies
3. Identify and highlight the network connectivity of important customers at any point along the network linkage
4. Interact with customers for all sorts of outages complains
5. Generate reference for each complain order sheet
6. Access the customer's service unit history database file in order to check for any previous outage and reasons behind these outages
7. Record the time of outage or incident as reported by the customer, the time of the emergency field crew was informed about it and the time of commencement of the repair-work

8. Retrieve and list all the distribution points from each customer to the low voltage transmission lines distribution
9. Locate all electrical equipment such as transformers, capacitors and other accessories on the low voltage distribution lines
10. Locate all distribution points from the high voltage transmission lines to the transformer Identify all distribution points from the substation to the transformers
11. Trace hierarchically linkage for any individual service unit
12. Maintain the records of equipment and their types located at the substations
13. Retrieve and list all distribution points, which are supplied from the power station to the substations

The fully developed system should be able to maintain and retrieve customer service unit database file including customer name, customer location and address, utility service size, installed load etc. Also the system will be able to maintain and retrieve distribution information such as pole number, wallbox number meter number, distribution point location etc. Additional information includes geographical information on network connectivity such as substation number and name, substation location etc. The system should maintain located at each substation.

The fully developed and retrieve records on equipment and their types system based on developed simulation software will have the following prime benefits:

1. It will replace a paper mapping system, which not only required manual updating and renewal, but also requires needless duplication, and extensive physical storage space. The maps are also extremely vulnerable to damage.
2. It will be highly flexible tool for managing assets, providing detailed information about attributes, geographic positions, spatial relationships and maintenance histories.
3. The system will provide a digital model of supply network that will help users analyze operational systems, such as street light networks, power networks, and signal and communications networks.

The system could also include methods to calculate voltages and losses in the network and cost benefit ratio incorporating future development/improvement schemes, capacity utilization and estimation, load analysis and the electrical distribution network, calculation of losses in the network under loading conditions, calculation of voltages at various points on the network, and optimal placement of capacitors in the network (Sharma and Khare 1999).

## CONCLUSION

The paper describes a simulation of a typical electrical utility customer center using GIS technology. The fully

developed system based on developed simulation software will allow migration to an open GIS environment that will provide a better level of service to customers, more accountable management of assets, better management of distribution network, faster management process and will result in overall benefit to customers. GIS ability to better manage the infrastructure, provide better and more timely maintenance and to provide the link between service requirements and customer needs. Use of GIS in utilities will grow over the next decade with the need to integrate a number of previously different systems.

## REFERENCES

- Balakrishnan, S. 2002. "Data Specification for Utility GIS and Corresponding Cost Benefits in the Year 2002", *GISDevelopment*.
- Harder, C. 1999. "Enterprise GIS for Energy Companies", ESRI Press, California.
- Helmer, T. 2001. "ArcGIS 8 Enabling Delivery Resources Planning", ESRI International User Conference.
- Kumar, V. and A. Chandra. 2002. "Role of Geographical Information Systems in Distribution Management", *GISDevelopment*.
- Marmar, S. 1999. "Implementing a GIS in an Electrical Utility: The Growth Pains", *GISDevelopment*.
- Piplapure, A. 2002. "Modelling the Electricity Distribution Networks, GIS Development Magazine", April.
- Powershop: Map India 2001. "Role of GIS as a Decision Support System in Power Transmission".
- Sharma, M. and M. Khare. 1999. "Enhancing the Analysis and/or Design of Electrical Distribution Network by Virtue of its Integration with GIS", *GISDevelopment*.
- Towns, D., 1994. "Customer Service Enhancement in an Electricity Utility through GIS", *AM/FM International*, 831-837.
- Vann, R., 2000. "Populating a GIS of Utility Corridor Assets Using an Integrated Airborne Data Acquisition System", ESRI International Conference.

## AUTHORS BIOGRAPHIES

**NEDZAD MEHIC** was born in Sarajevo, Bosnia and Herzegovina. Presently, Dr. Mehic is Associate Professor at the Computer Science Department at the IT College of the University of Bahrain and the Director of the Information Research Center, Deanship of Scientific Research, of the University of Bahrain, Kingdom of Bahrain. His current research interests include Internet technologies, Java, e-commerce, and GIS.

**ABDULLA SAAD AL-HUWAIHI** is the manager of the Operations and Maintenance Department at the Electricity Distribution Directorate of the Ministry of Electricity and Water, Kingdom of Bahrain. Mr. Huwaihi holds B. Sc. degree in Electrical Engineering from the Alexandria University, Egypt (1980). His current duties include responsibility for all operations and maintenance of distribution networks in the Kingdom of Bahrain.

# COMBAT SIMULATION TERMINOLOGY SHIFT TO PROCESS CRISIS MANAGEMENT

Bedřich Rýznar and \*Jiří F. Urbánek

Ground Forces University of the Czech Army in Vyškov and \*Brno University of Technology

V.Nejdlého, 682 01 Vyškov; \*Technická 2, 616 69 Brno, Czech Republic,

[ryznar@vvs-pv.cz](mailto:ryznar@vvs-pv.cz)

## KEY WORDS

Integration, Simulation Technology, Crisis Process Management

## ABSTRACT

Based on the tasks of modern armies, the introduction of simulation technologies is the most significant trend in the development of military instruction and training techniques for the preparation of commanders, staffs and troops. The experiences from the using of Ground Forces Training Centre of the Czech Army in Vyškov are presented here. This training and combat simulation centre is equipped with complex and efficient modelling and simulating (M&S) system. It combines virtual/real environment of combat situations, military means and crisis management systems. It makes the training, battle procedure, operations research, technological development etc. near to real conditions. The Ground Forces University of the Czech Army in Vyškov develops study programme "Process Management" already more than six years. The harmony and debugging of word asked process approach (here represented by Theory of Processes - Urbánek 2002) to army terminology is imperative task here, not only at experimental or scientific level. This approach is applied at army environments. It is task for real military education and culture. The conception - "crisis process" is introduced in several simulation environments, function modalities and procedures. The combat simulation methods are tested and evaluated at their dependences, relations and integration to the crisis processes and their environments.

## INTRODUCTION

University established system's approach is necessary to upgrade by means of process approach. The reason is for it a training of future crisis situation. Process approach contrasts to system approach by great emphasis on the integration to real-time, to real-space, to real-agents and to information-environment. The process environment is key entity, which include asked properties and behaviour for crisis situation solution. Its solution asks above all the manager qualification, decisiveness, flexibility, creativity, learning ability and operative ability. The Management can be defined as a "anthropogenic control" here. A Manager qualification is more difficult then a specialization in crisis situation solution. But the both are coessential. The qualification can be obtained by means a study and the specialization during the training. The environment (ENV) is necessary to comprehend as the most universal entity, as relation field - a Blazon (Urbánek 2002). Every one process

is an operation at defined ENV. A programme controls the process. Natural law programs the natural processes. However, contemporary always more active anthroposophy intervention to the man near process system results that the processes at anthroposophy ENVs proceed according to man-made programme. Interdisciplinary complication of decision-making process management requires heavy expenses quantity of the information. The ability of fast, pragmatic, efficient, actual and effective information processing, oriented to the control decision-making, results in an influence of process programme. Control decision carries out the process-management. Decision-making of process management (and-crisis-management especially) proceeds at operational levels almost pure. Decision-making qualification can be innate, as a result of a life in pertinent ENV. However, the environment of uncontrolled calamities, wars and catastrophes of all sorts is not native to any animal race. Management qualification is then necessary to obtain by the education and training in simulation ENVs. Discriminative level of information reception and informative details, necessary for a solution of future crisis-situation, isn't possible to predict in advance. Principal source of future right decision-making is the education and training in simulation environments. Simulation ENV is possible to identify with virtual (electronic) ENV at up to date practical implementation. Then, we can talk about computerised-aided decision-making. But anthrop entities are not possible to eliminate, because it acts by man-made processes influence. Biggest difficulty of management decision-making is a reception of excessive details of simulation ENVs, which can grow up to information redundancy. But automatic filtration "unsubstantial details" is a trip to the infernos. The causality of future crisis situation is guided by stochastic as far as by chaotic principles (Mandelbrot 1982). Even unsubstantial detail can shows as a cause of fatal incidences at operation level. That is why the quality of management decision-making asks a talent also. The obligation of education system is to awake this talent and then lifelong amplify its on principles of the self-education, self-organization, self-teaching and autogenously training. Thereby, it proceeds the cycle closing, in which must be educated and trained good (not only crisis) management. Simulation-ENV, in which this cycle acts, it can be discriminated according the branches (the industry, army, etc.). But it must have common important thing – the process approach. It significant contributes to the improvement of army forces preparedness to fulfil tasks not only in traditional war operations, but also in solving various crisis situations. The education of army managers needs the implementation of process approach principles to the military practice (Rýznar 2001) (Vrab 1998).

## COMBAT SIMULATION CENTRE REALIZATION

Based on the experience of modern armies, the introduction of simulation technologies is the most significant trend in the development of military instruction and training techniques for the preparation of commanders, staffs and troops. With reference to the adopted “NATO Modelling & Simulation Master Plan“ for a particular field, modelling and simulating (M&S) provide immediately available, flexible and cost-effective ways of simulating the real environment and in this way substantially extend activities in the field of training, operations research, technological development etc. Moreover, they can contribute to the improvement of NATO forces’ preparedness to fulfil tasks not only in traditional war operations, but also in solving various crisis situations.

Therefore, based on the design documentation developed and approved on December 15, 1999, the construction of the Ground Forces Training Centre of the Czech Army (GFTC) in Vyškov was begun in accordance with “The Concept of Introducing Simulation Equipment into the Czech Army“, signed by the Minister of Defence in July 1998. The Military Academy in Vyškov was charged by the Czech Army General Staff with management of the construction. (Rýznar 2001).

The Ground Forces Training Centre of the Czech Army is, after the National Centre of Simulation Technologies (NCST) at the Military Academy in Brno, the second Czech Army training centre to be equipped with more complex and expensive simulation equipment. The GFTC complements the National Centre of Simulation Technologies in Brno in its functions and tasks.

In the concept applied, the GFTC is divided into three training areas (see the following scheme at the Figure 1.):

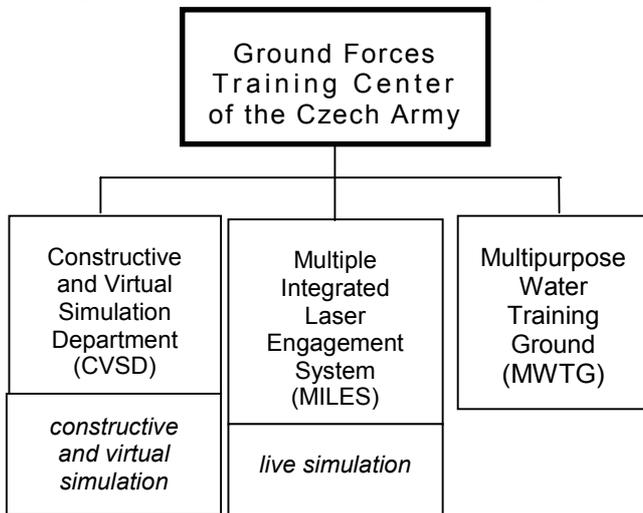


Figure 1: Organizational structure of the GFTC.

**Constructive and Virtual Simulation Department (CVSD)** as the most important part of the GFTC is intended for the training of commanders (at mechanized infantry squad or tank crew, platoon and company level) and the exercise of crews participating in the battle in the virtual environment. It can be connected to the battalion tactical simulator, which is a part of the NCST at the Military Academy in Brno, designed for the exercise of commanders and staffs.

The CVSD will enable:

- The training of squad, tank, platoon and company leaders/commanders to command and lead units in battle;
- The exercise of tank and infantry combat vehicle crews in battle situations;
- The training of units to fulfil tasks of combat fire and gunnery training;
- The exercise of drivers and the coordination of their operation within the framework of the crew.

All of this can be implemented at various command levels depending on the training cycle (squad or crew, platoon, company) using various training methods (exercises, joint and command exercises, combat shooting, tactical exercises with battle firing etc.) in unopposed as well as opposed forces exercises, under various terrain, weather and other conditions.

The elements of constructive and virtual simulation will be used to carry out the training in the CVSD. These types of simulation can be described briefly as follows:

➤ **Virtual simulation** is a type of simulation using virtual representations of real installations and/or objects to model the combat environment and combat operation. Based on the type of simulation required, particularly visual, but also others, the elements are exposed to the participants in the exercise. These elements are similar to those in a realistic operation in actual combat equipment (see Figure 2).



Figure 2: 3D method used to illustrate the battlefield during virtual simulation.

➤ **Constructive simulation** is able to supplement virtual simulation in an appropriate way. It is a large-scale computer-based simulation of combat operations intended for the training of commanders and staffs at battalion and higher levels, in general. It deals with the illustration of the combat environment, which is common to the previously mentioned levels. Commanders and staffs participating in the exercise solve tactical situations (fulfil combat tasks) in organic command centres or in mock-ups in training centres. The activities of subordinate units are simulated using a model containing programmed algorithms.

ModSAF simulation system is a representative of constructive simulation, which will be used in the CVSD. It is an open system of constructive simulation program modules arranged in a hierarchic way (called libraries in

ModSAF). Most of its libraries of computer-generated forces serve to generate simulated units and to generate their operation semi-automatically. In other words, when a soldier, vehicle, artillery battery or unit move, observe, fire, communicate etc. in the simulated battlefield, based on tactical principles adopted, the synthetic dynamic battlefield is shared by all the simulators participating in the joint exercise. Another method of simulating units requires input from the operator to expose simulated forces and equipment. It is usually employed in operations, which call for relatively higher levels of tactical decision-making as a rule, carried out by platoon leaders or higher commanders. For instance, in the battalion exercise company commanders are operators. They put data, e.g. company tasks, their “decisions” etc., into the simulator. In this way the operation of units, which are employed in the exercises for commanders at various levels, is simulated. The subordinated, supporting, and co-operative units and equipment are simulated to the commanders in the exercise.

➤ The Multiple Integrated Laser Engagement System (MILES) is based on **live and realistic simulation**. Standard field training is carried out with this type of simulation. It is usually in opposed forces training – in the training field or in the military training area. The participants in the field training use combined-arms armament and equipment fitted with devices, which use material to simulate firing, other combat activities against the enemy and the effects of these activities. For instance, in tactical training, laser generators in the invisible spectrum of frequencies are mounted on weapons (Figure 3) and the results of fire are recorded by special-purpose detecting elements (detectors) on the combat equipment which soldiers of both sides in the range of mechanized (tank) companies are equipped with. These detectors record the beams and monitor the activities of soldiers.



Figure 3: Laser generator mounted on rifle and detectors.

**The connection of constructive and virtual simulation** should be commonplace during training in the CVSD. Important actions taking place in virtual simulation must be accepted and displayed in constructive simulations and vice versa for this requirement. Advanced simulation technologies enable the sharing of a common synthetic dynamic battlefield not only for models and simulators, but also to integrate into it a field exercise. Therefore, it is possible to expect connection with live simulation technologies in the future.

The connection of individual simulation models (communication between individual simulators) should be enabled through use of a shared database and proven and internationally standardized technologies of distributed interactive simulation (DIS protocol) or, in the future, High Level Architecture (HLA). These must use all elements of both types of simulation used. Accomplishing the mutual interactivity of connected simulators should create the required conditions for realism of the dynamic battlefield in relation to the real environment where the actions are in progress concurrently. These actions are sometimes more and sometimes less interrelated and certainly influences one another. Another problem concerning the connection is the necessity of cross-correlation (interdependence) of both terrain databases used. This issue must be solved conceptually by laying down mandatory standards and regulations for the preparation of these databases.

Organizationally, the CVSD will be divided into:

➤ **The workplace of the training director** and his personnel – which provides the organizational and technical support of the training including the analysis of results in a given exercise. After Action Review (AAR), uses the recording device (logger) which records complex conditions of individual material and units and information transmitted in the simulated hierarchy of command and control. After the exercise is finished, the director can replay the exercise and analyze its most important parts.

➤ **The workplace of training personnel** (Figures 4, 5, 6) – which consists of halls and classrooms with various types of simulators where virtual simulation technologies (simulators for the training of T-72 crews, BMP-1 and BMP-2 crews) as well as constructive simulation technologies (ModSAF Simulation System) are applied.



Figure 4: Simulator for the training of T-72 crew.



Figures 5: Post of tank commander in simulator and image of his view



Figure 6: Instructor's Desk for the T-72 Simulator

## CONCLUSION

The construction of the Ground Forces Training Center is planned in the stages. The completion is planned for 2005. Testing of the CVSD is running from the end of 2001. Inauguration of the MILES simulator and Multipurpose Water Training Ground is planned for the second half of 2002 (Rýznar 2001).

## REFERENCES

- Mandelbrot, B.B. 1982. „The Fractal Geometry of Nature”. W.H. Freeman, San Francisco.
- Rýznar, B. 2001. “Ground Forces Training Center of the Czech Army in Vyškov”, The International Defence Training Journal MS&T, Issue 3, ISSN 1471-1052
- Urbánek, J.F. and Skála, Z. 2001. “Theory of Processes<sup>®</sup> - new Language of Human Sustainable Development”, *International Containment & Remediation Technology Conference and Exhibition, USA, Orlando.*
- Urbánek, J.F. and Rýznar, B. 2002. “Environmental Theory of Processes”, MESM'2002 International Conference, Sharjah, U.A.E., in print.
- Urbánek, J.F. 2002. „Theory of Processes – a Management of Environments“, Czech language book, edition of the CERM Brno, ISBN 80-7204-232-7.
- Vrab, V. 1998. “The Concept of Staff Officers' and Commanders' Training Using Computer Simulation and Simulators”, Inaugural dissertation, Military Academy, Brno:

## AUTHOR BIOGRAPHY

BEDŘICH RÝZNAR (Lieutenant Colonel of Czech Army) was born 1958 in Zlín (Czech Republic) and went to the Military University of Army Ground Forces Vyškov, Faculty of Military Control Systems. Here he studied the Tactics and obtained his degrees in 1981 and PhD in 1993. He graduates of Military Academy in Brno 1989 and Economics University in Praha 1997 beside.

He teaches as the Associate Professor at Military University of Army Ground Forces Vyškov, Faculty of Military Control Systems, Department of Military Tactics in the branches Military Tactics and Operation Science. His specialised fields of scientific qualification are process simulation. He is University Vice-dean, Member of University Scientific Board and Chairman of the Committee for State Examinations. He lectures of minister's study subjects: Operation and Tactics.

Publication activities: 25 specialised papers at International Conferences and in the journals. 2 books.

**SIMULATION  
IN  
MANUFACTURING**



## Methods and Tools for the future Factory and Production Systems Planning

*Dr.-Ing. Dipl.-Wirtsch.-Ing. W.Sihn,*

*Dr.-Ing. Dipl.-Kfm. Ralf von Briel,*

**Fraunhofer Institute for Manufacturing Engineering and Automation**

*Nobelstr. 12, D-70569 Stuttgart*

*Tel.: 0711/970-1964, Fax.: 0711/970 1002*

*E-mail: [whs@ipa.fhg.de](mailto:whs@ipa.fhg.de), [rvb@ipa.fhg.de](mailto:rvb@ipa.fhg.de)*

### **Abstract**

The problem of significant increase of market turbulence has been discussed by many authors who however don't even try to differentiate this general concept. In this paper, the causes for turbulences will be presented based on an enterprise survey. This survey implies that the strategies for turbulence control are not sustainable strategies, explicitly at the level of control and single order. Therefore, the structures and process flows in the enterprise must be continuously adapted in order to forecast the market turbulence and thus to increase the transformability of the enterprise. This, however, can only be possible if processes of production system planning undergo a fundamental paradigm shift and organizational structure and deployed technologies are improved. Thereby, the once project-oriented character of Factory Planning must give way to its continuous character. In this paper, new approaches to the Digital Factory, in particular the Planning Desk as tool for the concept planning and the Laser Scanner as a tool for the analysis of the actual state, will be integrated in the planning chain of the Digital Factory and the application fields. Potentials for further development will also be addressed.

### **Keywords**

Factory Planning, Continuous Structure Planning, Digital Factory, Factory Planning Desk and Laser Scanner

### **1 Turbulence – well known phenomenon or a new trend?**

#### **1.1 Turbulence: What does it actually mean?**

American researcher Henry Mintzberg shows by means of an analysis of various publications that the enterprises and their managers have already been complaining about the high market dynamics and great cost pressure and pressure

of competition for 30 years [MB94]. Thus, he remarks that market turbulence is apparently familiar to the enterprises and therefore respective methods and procedures should be available to cope with this problem.

This, however, does not reflect the reality. During the past years, many enterprises had to face chaotic situations [WI94, RE99], triggered by the turbulent sales and procurement markets. While Mintzberg claims that the market fluctuations and turbulences have always occurred, the authors [KU98, MI97] claim that the high rate of change of market turbulence acts as a trigger for the pressure of adaptation in enterprises. The general development trends like reduction of product life cycles and the increase of variant figures indicate the increase in turbulence. However, a detailed analysis of the causes for turbulence remains undone. Therefore, systematic planning mistakes and unadapted application of methods are often referred to as turbulences. In the situations where the actual figures never reflect the target figures for production, or the uncultivated product master data causes insufficient deadline adherence, the imperfect quality of results do not originate from the increase in turbulences but from lacking methods. Therefore, it is not possible to conclude the increase of turbulence from the reduction of planning time or from the increased number of variables. The results of the special research project 467 "Transformable Business Structures for Multi-Variant Serial Production" show rather that in the actual turbulences it appear if there is, in addition to the normal change, a "surprise effect" occurring and the appropriate outcome is no longer possible without great effort [WK00].

The causes of turbulence, or "turbulence sources", can be divided into internal and external sources depending on

---

where the change takes place. Furthermore, the turbulences can be classified according to their reaction time as short-term - on the level of single orders and resources, as medium-term - on the level of the production schedules and as long-term on the level of product programs and network structures. On the basis of this definition and systemization, fields of turbulence can be derived and differentiated for the enterprises and their structures.

The results of the survey of more than 200 enterprises (< 200 employees), carried out within the scope of the special research project 467, has identified 10 crucial turbulence sources in the short- and medium-term field.

## 1.2 Implication of this definition

Based on this definition, the turbulence sources cause an increase of turbulence only when the previous approaches do not provide a market-adapted reaction. The conventional methods, like increasing inventory stock, extending the delivery date or early make-to-stock production make it possible to counteract the turbulence sources. However all these methods neither induce a market-adapted response time nor set marketable and acceptable prices.

According to many authors [WI02, WR02, WE02] market-adapted response to the presented turbulence sources makes it necessary to adapt the production structures more frequently in the future. Only if it proves successful to increase the readiness for change in the structures of enterprises it will be possible to control the turbulences at a higher level.

The results of the study confirm this theory. It indicates that the reduction of order processing time holds the most significant potential for the future in a common enterprise. Hereby it also implies that the conventional methods like increasing inventory stock do not prove to yield sufficiently adapted response times.

The study also shows that the reduction of planning time for product development and production -start-up and the operative reduction of cycle time are equally important. This also indicates that reduction of planning time involves increasing the frequency of planning and adaptation of the structures.

Setting the structure of the production system as a variable enables turbulence control to a higher degree. Many enterprises claim that this increased flexibility, referred to as transformability of the enterprise, must be significantly enhanced in the future.

The survey also reveals that only 31% of the enterprises interviewed describe the actual state of their enterprises as “transformable” while 51% are aiming to become. Similarly, 4% of the enterprises claim to be “highly transformable“, whereas 39% strive to be.

On the basis of a detailed definition of turbulence, it has been stated that the turbulence could only be controlled if the structural variability or transformability of the enterprises would be significantly improved. In the following paragraphs new methods and tools of the Digital Factory will be described based on the concept of continuous structure adaptation.

## 2 Turbulence control through transformability

As structure adaptation is a compensation mechanism for the increased turbulence, enterprises must now determine the frequency and level of structure adaptation. Set-up on the level of single orders can be selected as an analogy according to frequency. Set-up, that makes the next order possible on the level of workstations, must take place more frequently, the shorter the processing time or smaller the processing amount of the order. A “continuous set-up” is available on the level of single order if it takes place after every produced part i.e. the batch quantity of order equals 1. If this analogy will be transferred to the next higher level of order profile, the production system must always be set up if the order profile relevant for this system would be changed. Analogically, continuous set-up or continuous adaptation of production system will only take place if the adaptation of production system results at the inspection point of the order profile i.e. in the monthly planning meeting or on similar occasion. Analogically to the set-up on the level of single orders, time for the set-up of production systems in the factory will be critical due to the frequency [BR01].

Thus, there is a correlation between turbulence control and transformability. In the case of increased turbulence and therefore more frequent adaptation of production structure, the time required for the “set-up of the factory” will be a critical success factor. Furthermore, the turbulence can increase so heavily that the continuous adaptation of the production structures will be necessary. [WIME02, WK00]. The continuous planning must also take place on the level of location if the turbulence further increases and the order profile changes continuously and unpredictably.

According to Eversheim and Warnecke [EWSC96, WA95] Factory Planning encompasses “planning and design of

---

manufacturing facilities together with monitoring the realization up to the production start-up” and embraces the entire structure of resources of the production from “rescheduling of machines and their supporting functions up to building a new plant”. Should the production systems be set up faster, the Factory Planning plays an essential role in increasing the transformability of factories and production systems. However on the contrary to the conventional set-up, the factory set-up must be a systematic task. Therefore improving technology as a system element is not sufficient for the acceleration of this operation. Furthermore, the adaptability of all elements of the entire production system i.e. technology, employees, organizational and facility structures, must be increased.

Despite this integral optimization task, the processes play an essential role, as they are measurable improvement steps in the enterprise. When viewing processes according to their improvement potential the Digital Factory represents central element for increasing the transformability.

Therefore, the Digital Factory is not an end in itself but must aim at increasing the transformability of production systems and factories whilst it enables reduction of the planning time required for the factory changeover for the similar or improved planning quality and cost level. All measures performed in the field of the Digital Factory of the enterprise must be evaluated according to this aim.

### **3 Digital Factory as a new method to increase transformability**

Despite the fact that there has been a great variety of definitions of Digital Factory [BI01, FG98, WR01, ME02], all of them indicate its three key components: Simulation, Data Integration and Visualization. The main task of simulation is to support technical, logistic and business planning processes by means of dynamic models. Data Integration should serve as a continuous solution between these two levels as well as between the planning phases. Visualization should provide access to all planning levels and planning data. The Digital Factory enables the following functions:

- All processes and data flows of the factory will be digitally modeled, tested and simulated before the actual operation.
- The models are geared together at different abstraction levels (technical simulation, logistic simulation, simulation of business processes).

The data of the entire factory planning will be stored in data bank systems that will be accessible by all planning participants worldwide.

- Using Virtual Reality, factories will be visualized in 3D models, in which a 3D Walk Through will be possible before building of the factory.
- New production processes will be developed either on the basis of existing configurable modules or through default or self-learned rules.

The aim of the Digital Factory is to combine the existing elements in an improved way or to a greater extent. It is then no longer essential to improve the single activities or production steps in the factory planning. However, if the development of new methods of the factory planning would be taken into consideration the Digital Factory alone would not be sufficient. The simulation of the material flow is nowadays an essential element of Factory Planning.

The development of the Digital Factory should also be open to new approaches. The necessary improvement of the planning processes and their efficiency can be achieved on condition that the new methods of factory planning would be consecutively associated with the approaches of the Digital Factory

Following, selected new methods of Factory Planning and their integration in the Digital Factory will be described. On the one hand, the Factory Planning Desk will be presented as a system for participative planning while on the other hand the need for further development in the field of actual state analysis will be explained by means of Laser Scanning.

#### **3.1 Planning Desk as a system for Participative Planning**

According to Breit [BR99], planning time can be significantly reduced through integration, parallelization, avoiding iterations and improved cooperation. These mechanisms will be supported in the participative planning. The main aim of participative planning is therefore to improve the integration in the concept planning due to know-how of the different planning partners. Furthermore, the information transfer between the interfaces will be improved, the parallelization of the planning processes enabled and the unnecessary duplication of effort, resulting from the repeated data input and transfer, will be significantly reduced. The functionalities, efficiency and the technical layout have been described in detail in various publications [BI99,WK00,SI00], therefore

---

only substantial further developments in the participative planning will be presented in the following paragraphs.

Besides these developments, the efficiency of the planning desk has been increased while taking advantage of noticeably improved computer hardware.

Planning Desk as tool for participative planning will first be integrated in the planning chain based on the basic principle of separation between visualization and data management. Within the framework of “Dynamic Factory Structures“ sub-projects, the interface between the Planning Desk as tool for participative planning and a data-driven simulation tool has been established, which enables a parallel simulative and subjective evaluation.

Thus, dynamic evaluation will already be possible in the concept phase. Besides the simulation of the material flow, the interface between Layout Planning and the associated technical control mechanisms of the material flow systems and manufacturing facilities will be established. Therefore the controlling objects (actuator, sensor, bus systems) must be combined with the layout elements. Thus, a further logical approach will be integrated and visualized by means of a corporate platform.

A further integration of the Planning Desk will be to connect it to worldwide available data and models of planning objects. Based on Internet technologies, a network will be established similar to the „Napster concept“, where each participant is able to place data models and then receives a corresponding number of models from the network. This enables faster distribution of models and at the same time reduces the effort of developing new models.

Based on this enhancement of the Planning Desk, a further reduction of the time required for establishing planning variants will be possible in the future. The required interface between visualization and evaluation will already be possible in the concept phase through the integration with simulation. The connection between Layout Planning and Control Planning allows early coordination between different planning tasks. The Napster principle also enables to support the spatially divided planning. Therewith the efficiency of the Planning Desk can be further increased and its integration into the Digital Factory as a concept tool can be achieved. If the 80:20 principle of the product development also applies for Factory Planning, significant potential in the Digital Factory can be developed through acceleration and improve-

ment of the participative planning implemented in the concept phase.

### **3.2 Integration of Laser scanning**

The Laserscan 3D measurement system is a starting point for the Digital Factory planning in existing facilities. The weak points, hampering the application of methods for Digital Factory planning, can therewith be reduced. The Laserscan forms a basis for the actual spatial layout by the help of which it is possible to derive the changes. However, according to Aggteleky [AG90] and Bischoff [BI01], the representation of the actual state is only the starting point for the planning of the adaptation measurements. The essential planning task is the determination of the target state of the adapted factory. Therefore, the scatter-plotted data must be transformed in order to not only make the distance measurements available in the actual state but moreover to enable the changes and adaptations. Therefore, two feasible approaches have been established, one of them already being fully automatable, the second still requiring a great deal of manual modeling effort

In the first approach, 3D CAD machines data will be recorded in the scatter plots. Thus, a model will be developed containing both the information of the scatter plots and the information of the new or adapted elements. This approach is particularly beneficial if the layout and collision in 3D space of the new objects will be tested at the actual state and the empty hall geometry has been laser-scanned. However, this approach does not produce a complete CAD model consisting of all information of the 3D scatter plots and the new objects. Therefore, this approach can only be applied if there is an actual 2D CAD drawing and the changed objects, already tested for their collision in 3D space, can be inserted in this drawing. Within the scope of the special research project 467 “Transformable Business Structures for Multi-Variant Serial Production“ the scatter plot representation will be integrated in the team-based Factory Planning on the basis of the Factory Planning Desk. Thereby it is possible to use the scatter plot either as a background planning wherein the new object can be inserted or to cut out the elements of the scatter plot and integrate them in the existing planning methods. Thus, new machinery can be placed in the existing factory building or existing machinery can be located in new buildings. The acceleration of Digital Factory Planning by means of the Planning Desk and reduction of modeling effort lead to considerable reduction of the total planning time. It enables a further step towards Digital Factory Planning.

In the second approach, that is currently being developed, CAD elements will first be generated from the scatter plots. This however still requires a great deal of manual modeling efforts. In order to fully automate the CAD model generation, it is possible to access the data of the generated CAD elements of the scanned products. [WK00]. On the contrary to the product-oriented CAD transfer, where the main task was to define the free form surfaces, the task here is to select and allocate the standard geometries in the elements of the scatter plot. The bar charts do not consist of a pre-defined number of polygons, but they will be generated in the CAD model as independent and scaleable objects with the dimensions of the bar chart of the scatter plot. This approach leads not only to a reduction of the polygons to be presented but moreover it allows establishment of structured and object-oriented CAD models, in which the scaleable general model of the Digital Factory is to be integrated. Various CAD Systems, like TriCAD, with a large scaleable library of CAD elements already make the integration possible. The second approach allows development of a complete CAD model which can serve either as an input for the team-based planning or is available for further planning systems. Thus, the further step towards acceleration of the planning processes and complete digitalization of a factory has been made. This again will noticeably reduce the planning time and effort. The third significant development for the future will be the question of whether the laser scan could be beneficial for the actual acquisition of the moveable models in the factory and if this information can be used for job control. The next aspect is therefore whether the elements from order and material tracking must be adapted in the future through the application of the new technology. This however still remains a great novelty both for the Laser Scanner as well as for the application.

#### 4 Conclusion

Despite many efforts, the problem of turbulence remains unsolved. According to the survey made, enterprises are forced on finding new mechanisms to control turbulences. A sustainable approach would be to frequently and rapidly adapt factory structures. The Digital Factory offers an efficient approach to reduce planning time in the Factory Planning stage. It will only reach its full potential in the long run if it focuses on this aim and remains open to new methods and developments.

#### 5 Literature

- [AG90] Fabrikplanung – Werkentwicklung und Betriebsrationalisierung. Bd. 2: Betriebsanalyse und Feasibility Studie: technisch-wirtschaftliche Optimierung von Anlagen und Bauten. 2. Aufl. München u.a.: Hanser, 1990
- [BI01] Bischoff, J.: Die Fabrik: Planung, Betrieb und Wandlungsfähigkeit. In: Das große Handbuch Produktion / Burkhardt, W. (Hrsg.). Landsberg/Lech: Moderne Industrie, 2001
- [BI99] Bischoff J. et al.: Team Desk: A Framework and Tool for a Continuous Factory Planning SPIE: Intelligent Systems in Design and Manufacturing II, Boston 1999
- [BR01] v. Briel, R: Methoden und Werkzeuge für zukünftige Fabrikplanung. In: Wirth, S. et al.: Vernetzt planen und produzieren VPP 2001: Tagungsband Chemnitz 20.-21. September 2001. S. 83- 87.
- [BR99] Breit, S.: Methodik zur umsetzungsorientierten Gestaltung von Umstrukturierungsprojekten in der Produktion. Aachen: Shaker, 1999. Aachen, RWTH, Diss., 1999
- [EH95] Ehrlenspiel, K.: Integrierte Produktentwicklung – Methoden für as, Produkterstellung und Konstruktion. Hanser Verlag, 1995
- [EVSC96] Eversheim, W; Schuh, G: Die Betriebshütte. Produktion und Management, T. 2. 7. Aufl. Berlin u.a.: Springer, 1996
- [FG98] Flaig, T.; Grefen, K: Virtual Environment for Integrative Factory and Logistics Planning with Virtual Reality. In: Changing the Ways we Work. Shaping the ICT-Solutions for the Next Century, Conference on Integration in Manufacturing, 6.-8. Oktober 1998 in Göteborg, Schweden / Martensson, N.; Mackay, R.; Björgvinsson, S. (Hrsg.).
- [MB94] Mintzberg, H: That's not „Turbulence, Chicken Little, It's Really Opportunity“. In: Planning Review, 11/12, 1994, S. 7-9.
- [ME02] Menges, R.: Wettbewerbsfähigkeit durch virtuelle Fabrikplanung. In: Westkämper, E.; Schraft R.-D. (01): Zukunftsfähige Fabrik – Vorsprung im Wandel. Fraunhofer Ipa Seminar F74 5. März 2002, Stuttgart. S-40-62.
- [RE99] Vom Wandel der Zeit: Wandel als Chance für unsere Unternehmen im globalen Wettbewerb. In: ZWF 1994 (1999) 1-2, S.14
- [SI00] Sihm, W. et al. Planning around the Desk: A New Approach towards a Continuous Factory Planning Process. 33<sup>rd</sup> CIRP International Seminar on Manufacturing, Stockholm, Schweden, 2000
- [WA95] Warnecke, H.-J.: Der Fertigungsbetrieb. Bd. 1-3. Berlin u.a.: Springer 1995
- WI01] Wiendahl, H.-P.: Die wandelbare Fabrik – Kundenorientierte Produktion In: BFT 3 (2001)
- WI94] Potentiale der Chaosforschung in den Produktionswissenschaften. Tagungsunterlage zum Work-

---

shop am 28. Februar und 1. März 1994 am Institut für Fabrikanlagen (IFA), Universität Hannover.

WIME02] Wiendahl, H.-P.; Hernandez, R.: Fabrikplanung im Blickpunkt- Herausforderung Wandlungsfähigkeit. Wt Werkstatttechnik online (2002) Jg. 92 H 4. S. 133-137.

WK00] Westkämper, E. et al. (00): Turbulenz in der PPS–eine Analogie. In: wt Produktion und Management 90 (2000) Nr. 5, S. 203-207

WK00] Westkämper; E.: Kontinuierliche und partizipative Fabrikplanung. Wt Werkstatttechnik 90 (2000), 3, S. 92-95

WK01] Westkämper, E: Fabrikdigitalisierung. In: wt Werkstatttechnik online 91 (2001) Nr. 6, S. 304-307

WO01] Wortmann; D.: Traum Fabrik – Von der Simulationstechnik zur Digitalen Fabrik, Maintal ZWF 6/2001 96. Jahrgang S 342 –344. Carl Hanser Verlag München.

WR00] Wirth, S.: Visionen zur Fabrik von morgen: In Logistik/Heute (2000) H. 10; S.57.

# Simulation Analysis of Production Planning in Manufacturing Industry

Razman Mat Tahar, Rahela Abd Rahim,  
Hawa Ibrahim, Noraziah Hj Man & Fatinah Zainon

School of Quantitative Sciences  
University Utara Malaysia, 06010  
Sintok, Kedah, MALAYSIA

## KEYWORDS:

Simulation, Manufacturing, Production,  
Performance, Efficiency

## ABSTRACT

Simulation is one of the most prevailing tools available for the design and operation of complex processes and systems. This paper demonstrate the use of simulation model in manufacturing of an electronic device called interrupter. The simulation model developed in this manufacturing environment is used as a planning tool to help the management in making decision. The model assist in making decisions on the production management of plant operations and it optimizes the utilization of resources and provide estimates for the plant throughput.

## 1. Introduction

Simulation has been successfully applied in many different areas such as manufacturing, healthcare, transportation, supply chain etc[1,2]. This study is specifically focus on the use of simulation modeling in manufacturing industry. The product manufactured is an electronic device called interrupter. The interrupter size is

about 1 cm and used in many electronic equipments.

## 2. Objectives

The objectives of the study is to develop a simulation model of a manufacturing plant. In order to achieve the stated aim, the following objectives were identified and pursued:

- To create a simulation models that are capable to mimic the plant operations
- To use animation that is capable to offer a clear description of the models
- To increase factory throughput
- To propose a new operating policy to decrease bottleneck and to increase the performance of the plant.

## 3. Production Process Flow Description

The process begin at Station 1 with the arrival of interrupter's housing for marking and ending at Station 8 where the interrupters are tested for their functionality. The process to produce an interrupter is described as in Figure 3.1 below.

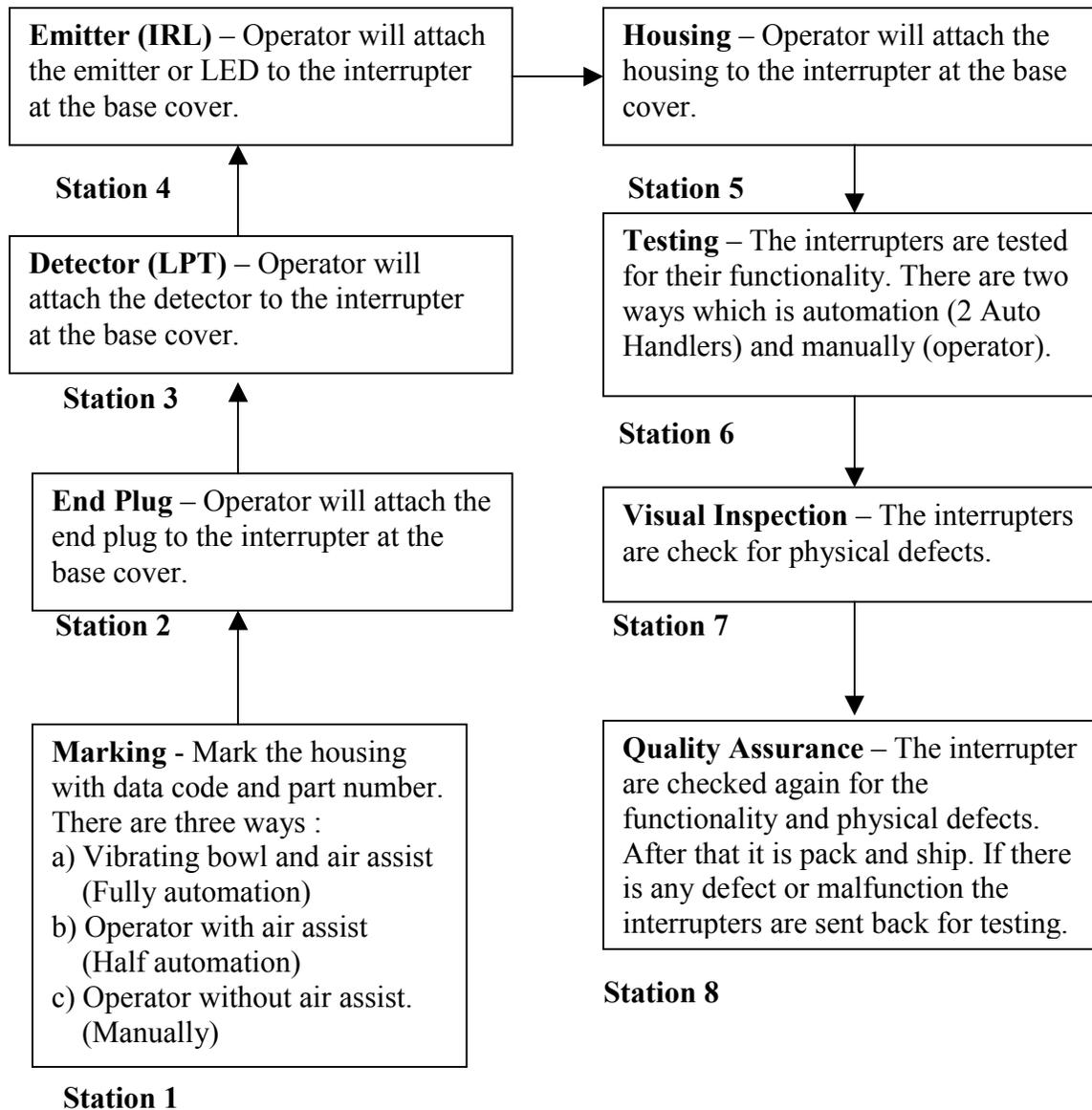


Figure 3.1 Process of Producing Interrupter

#### 4. Development of Simulation Model

Model is developed using ARENA simulation software. In the model, the interrupter represents the entity that moves throughout the system. In the ARENA software, the Input Analyzer is used to fit the appropriate distribution to

the empirical data collected. Most of the data were directly available, via operation or equipment specification documents. The simulation model presents an animated display of the production system. This animation can be run concurrently during simulation. The modeler can also include animation

in a real-time display of model statistics, such as dynamic plots, histograms, and time clocks, during the simulation in order to illustrate system performance. Statistics such as cycle time, resources utilization, queue time, and throughput can also be computed from the simulation output.

## 5. Summary of Simulation Results

The statistics collected from the simulation model include plant throughput, resources utilization and time spend in the queue.

### 5.1 Plant throughput

The output generated from the simulation output was 20,160 interrupters in one shift as compared with 19,289 from historical data. This shows a difference of 5%.

### 5.2 Resources Utilization

The resources utilization at plant is not equally distributed. The highest resources is at Station 1 and Station 2 with 85% each. The lowest is at Station 6 with 17.1%.

### 5.3 Queue Time

The highest queue time spend by the interrupter is both at Station 1 and Station 2 with an average of 54%.

## 6. Model Experimentation

A model experimentation is a test of different scenarios in which a meaningful changes are made to the input variables of a validated simulation model so that we may observe and identify the reasons for changes in the performance measures. Only two different scenarios as suggested by the management are listed here.

**Scenario 1 :** Reduce 75% the quantity of the box of interrupters to be filled. after marking process. As a result, the throughput increases by 10% and resources utilization are more equally distributed

**Scenario 2 :** Increase the time between arrival for both manual marking and half auto marking by 25%, and reduces the time between arrival of automation marking by 30%. This scenario is chosen in order to reduce the workload at manual marking and half auto marking and shift this workload to automation marking. It is verified that with this scenario, lesser interrupters have to queue to be process. However not much different with the total throughput.

## 7. Conclusion and Recommendation

In this study, a simulation model has been developed for a case study of manufacturing sector. The model can be used as a tool for making decisions of a production plant. Investigations on planning and changes can be tried on the model without disturbing the existing operations.

### Acknowledgements :

The authors would like to thank to the Universiti Utara Malaysia for funding this research.

### References:

1. Banks, J., J. S. Carson, and B. L. Nelson. 1996. *Discrete-event system simulation*. 2d ed. Upper Saddle River, New Jersey: Prentice-Hall.
2. Kelton, W. D., R. P. Sadowski, and D. A. Sadowski, 1998. *Simulation with Arena*. New York: McGraw-Hill.

# PETRI NET SIMULATION AND EFFICIENT EXPLOITATION OF FLEXIBLE MANUFACTURING SYSTEMS

Mihaela-Hanako Matcovschi  
Dep. of Automatic Control and Industrial Informatics  
Technical University "Gheorghe Asachi" of Iasi  
Blvd. Mangeron 53A, 6600 Iasi, Romania  
E-mail: mhanako@delta.ac.tuiasi.ro

Omran Hilal  
TOLEDO Electrical & Mechanical Works Est.  
PO Box 43542, Abu Dhabi, UAE  
E-mail: omranh@toledo.ae

## KEYWORDS

Event-oriented model, Performance analysis, Simulation interface.

## ABSTRACT

This paper emphasizes the role of simulation-based methods for the efficiency analysis in the exploitation of flexible manufacturing systems in connection with the deadlock avoidance problem. In order to incorporate both quantitative and qualitative information, relevant simulation tests must use timed models. The case study presented herein illustrates the utility of a Petri net simulator with appropriate facilities for the selection of the most adequate policy in controlling a production flow.

## 1. INTRODUCTION

*Flexible Manufacturing Systems* (FMSs) have become prominent in modern industry because the users can simply modify either physical or logical structure of the system. In FMSs resources are assigned to operations and released from operations. Failure to suitably assign shared resources can have serious effect on system performance, resulting, in extreme cases, in a sequence of  $n$  operations such that the  $i$ -th is waiting on the  $(i+1)$ -st and the  $n$ -th is waiting on the first. In this case the FMS enters *deadlock* or *circular wait state* and ceases to operate. Therefore, *deadlock free operation* is an important structural property of FMSs and represents a major objective of their control.

It is known that Petri nets (PNs) offer an effective framework for modeling FMSs where the deadlock avoidance problem can be rigorously approached. The purpose of our paper is to underline the role of simulation based on *P-timed PN models* in the efficiency analysis for the exploitation of the FMSs equipped with mechanisms dedicated to deadlock avoidance.

The paper is organized as follows: Section 2 comments, in Petri net terms, the deadlock avoidance strategies for FMSs within the context of time dependent performances. In Section 3, the characteristics a PN simulator must possess in order to allow a production efficiency study are pointed out. Section 4 contains a brief overview of the Petri Net Toolbox for MATLAB designed and implemented at the Department of Automatic Control and Industrial

Informatics of the Technical University "Gheorghe Asachi" of Iasi. Several examples of such studies carried out in this toolbox are presented in Section 5. Finally, some conclusions are delivered in Section 6.

## 2. DEADLOCK AVOIDANCE VERSUS TIME DEPENDENT PERFORMANCES

Since FMSs are driven by discrete events, they may be easily modeled and analyzed by means of *Petri nets*. Consequently, a number of techniques were proposed for dealing with FMS deadlock within the PN formalism, such as (Desrochers and Al-Jaar, 1993), (Ezpeleta et al., 1995), (Lewis et al., 1995), (Lewis et al., 1998), (Viswanadham et al., 1990), (Wysk et al., 1991), (Zhou and DiCesare, 1993). The basic idea is of *qualitative* nature and consists in using untimed PN models in order to ensure the robustness with regard to unexpected changes that might affect the durations of the operations performed by the resources in the FMSs. If an untimed PN is deadlock-free, then the FMS modeled by this PN will never be confronted with the deadlock problem, no matter how long each operation takes, provided that the sequencing of the operations along the entire production flow is not altered.

On the other hand, an extremely important criterion in exploiting an FMS consists in reducing the mean time spent by a part within the system (i.e. the mean time requested to obtain a final product from raw material). Therefore, such a performance analysis needs a timed PN model able to incorporate the *quantitative* information related to the processing time specific to each operation (Desrochers and Al-Jaar, 1993).

In many cases when deadlock appears in the untimed PN model, the addition of deterministic timing information results in a timed PN where deadlock does not occur because of a fortunate compensation of the time durations. These deterministic durations disable the firing sequences which lead to deadlock in the coverability tree of the untimed PN model. However, such a fortunate operation can be simply affected by perturbations of the time durations and, consequently, the occasional avoidance of deadlock might disappear. Hence, the correct exploitation of an FMS must combine the robustness in deadlock avoidance with the improvement of time dependent criteria.

### 3. ROLE OF SIMULATION IN ANALYZING PRODUCTION EFFICIENCY

In order to develop a relevant study of the quantitative behavior of an FMS, a stochastic timed PN model is requested so as to incorporate the significant details issued by the practical exploitation, when the durations of the operations often vary around a mean value. The analytical investigation based on such a model is far from a convenient approach and the usage of an appropriate simulation tool represents the only feasible solution. Furthermore, analytical tools can be helpful only for restricted classes of timed PNs, such as the max-plus algebra for event graphs (Bacelli et al., 1992), (Cohen et al., 1998), Markov chains for stochastic Petri nets (where all delays assigned to transitions are assumed to obey an exponential distribution), e.g. (Cassandras, 1993), (David and Alla, 1992).

A widespread technique for constructing the PN model of an FMS uses separate places for operation performance and resource availability, fact that requires the assignment of P-timing. Thus, the simulator must be able to operate with place-timed PN models so as one can allocate different probability distributions to each place in the net. Moreover, the simulator must provide statistical information to characterize the whole simulation period by means of standard performance indices such as mean values for arrival distance, throughput distance, queue length, etc.

It is worth mentioning that the capability of a simulator to operate only with stochastic transition-timed PNs represents a considerable impediment because the typical place-timed models cannot be directly used and their conversion into

T-timed PNs brings the major disadvantage of enlarging the topology.

### 4. USAGE OF PETRI NET TOOLBOX FOR MATLAB – BRIEF OVERVIEW

The *Petri Net Toolbox* for MATLAB (Mahulea et al., 2001), (Matcovschi et al., 2001) was designed and implemented at the Department of Automatic Control and Industrial Informatics of the Technical University “Gheorghe Asachi” of Iasi. Its integration with MATLAB (The MathWorks Inc., 2000a), (The MathWorks Inc., 2000b) offers multiple advantages with regard to the simulation aspects we are interested in, among which:

- (i) It is able to operate with nets having *infinite capacity* places;
- (ii) *Deterministic* or *stochastic* delays can be allocated to places or transitions in a net;
- (iii) Priorities and/or probabilities may be assigned to concurrently enabled transitions;
- (iv) The *coverability tree* in text or graphical mode may be easily constructed and allows studying the behavioral properties of the PN model (Murata, 1989);
- (v) An analyzer of time-dependent performances provides statistical information about transitions and places characterizing the whole simulation period.

### 5. A CASE STUDY

Simulation experiments addressed in the Petri Net Toolbox for MATLAB allow the user to choose the most efficient solution for deadlock avoidance based on the computation of the mean production cycle time, as shown below.

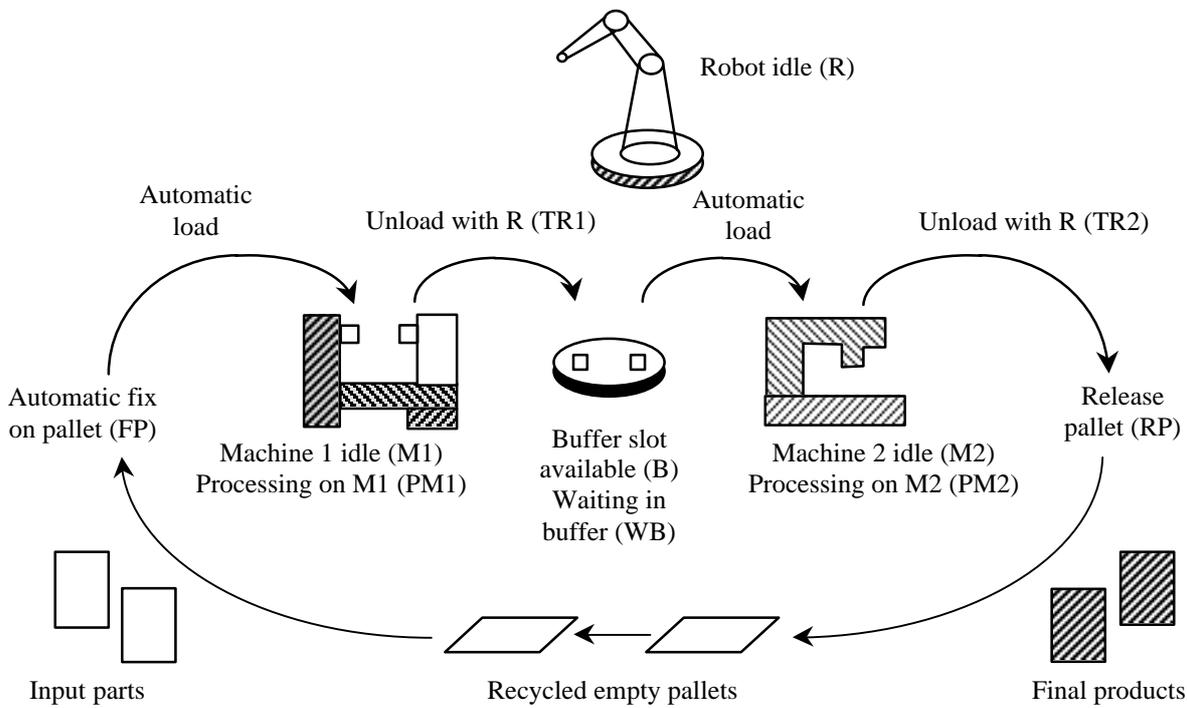


Figure 1. Schematic Representation of the FMS Used as Case-Study

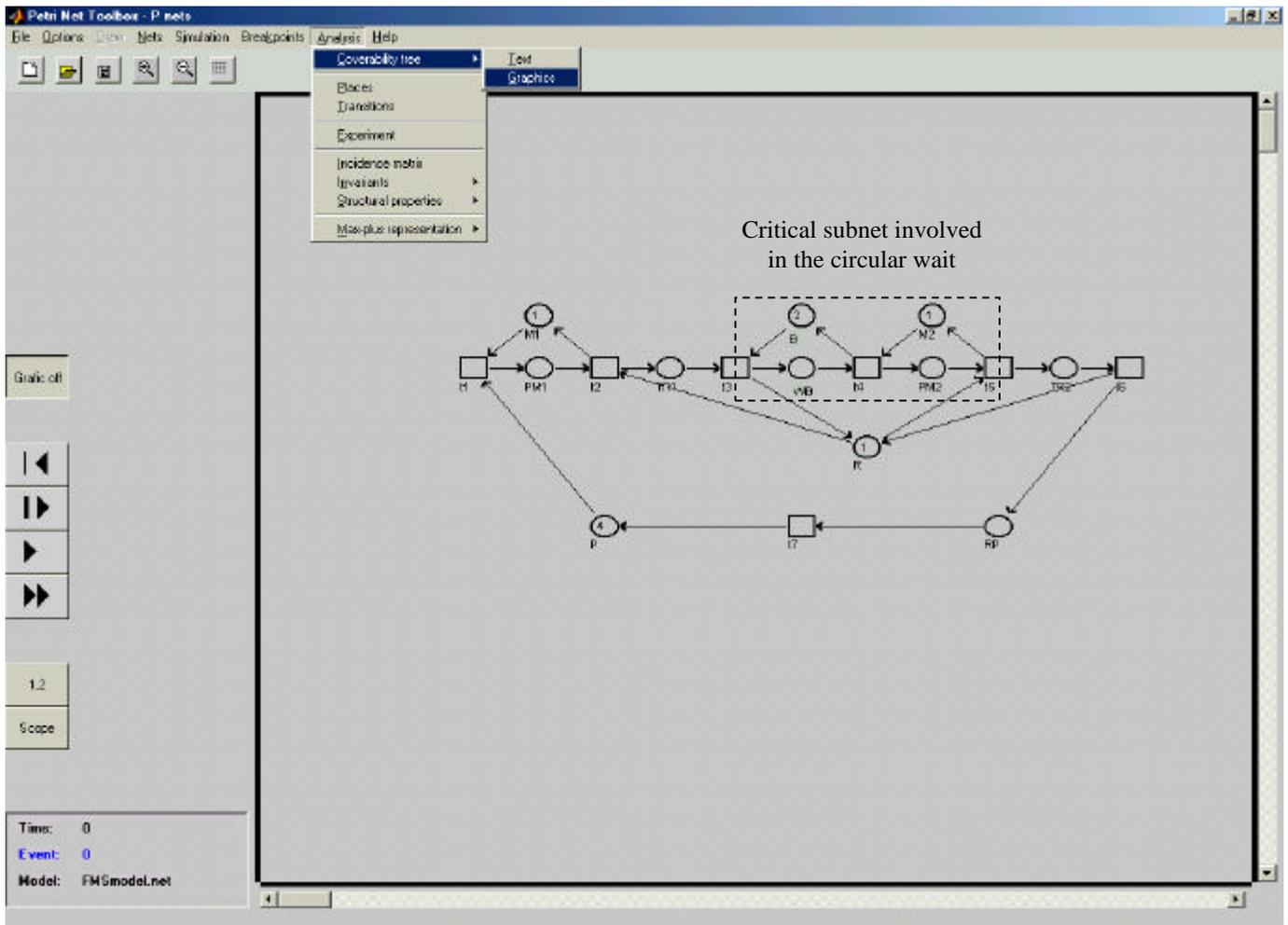


Figure 2. PN model of the FMS Used as Case-Study

The FMS presented in figure 1 was selected as an illustrative example and consists of two different machines (a lathe (M1) and a drilling machine (M2)), a robot (R) and a buffer (D) with two slots between the two machines (adaptation from (Zhou and DiCesare, 1993)). Every input part must be processed by M1 first and then by M2 in order to get the final product. Both machines are automatically loaded and are unloaded by the robot. A variable number of pallets can be used to fix on the processed parts

The PN model of this system drawn in the *Drawing Panel* of PN Toolbox is presented in figure 2. The places are labeled according to the abbreviations used in figure 1. The initial marking corresponds to the situation when all resources are idle.

Several scenarios for analyzing the behavior of this FMS have been considered. The following strategies for deadlock avoidance have been envisaged in our simulation study:

(1<sup>0</sup>) the limitation of the number of pallets in the system according to the token-capacity of the critical subnet in figure 2 (Zhou and DiCesare, 1993), (Lewis et al., 1998);

(2<sup>0</sup>) the usage of a *kanban* (Sugimori et al., 1977) or “*lookahead*” feedback (Lewis et al., 1995);  
 (3<sup>0</sup>) setting to TR2 a higher priority than to TR1 (Lewis et al., 1995).

Both deterministic and stochastic P-timed PN models have been analyzed in order to get a deeper insight into the time-dependent properties of the FMS. All simulation experiments described further on have run for 250,000 s. The performances of the system may be expressed in terms of the statistics offered by the PN Toolbox for the places and the transitions in the corresponding model.

We first deal with the *deterministic* P-timed model for which the delays corresponding to the operations have the following numerical values:

- automatic fix on pallet:  $d(FP) = 7.5$  s;
- unload from a machine with the robot:  $d(TR1) = d(TR2) = 22.5$  s;
- processing on M1:  $d(PM1) = 40$  s;
- processing on M2:  $d(PM2) = 80$  s;
- release pallet:  $d(RP) = 10$  s.

Place Name:	PM1	RP	TR1	FP	WB
Arrival Sum:	2749	2746	2748	2746	2748
Arrival / Time:	0.01	0.01	0.01	0.01	0.01
Arrival Dist.:	90.93	91.03	90.94	91.03	90.95
Throughput Sum:	2749	2746	2748	2749	2747
Throughput / Time:	0.01	0.01	0.01	0.01	0.01
Throughput Dist.:	90.95	91.03	90.95	90.93	90.99
Waiting Time:	51.48	10.00	22.50	7.54	67.97
Queue length:	0.57	0.11	0.25	0.08	0.75
Place Name:	PM2	TR2	M1	B	M2
Arrival Sum:	2747	2746	2748	2747	2746
Arrival / Time:	0.01	0.01	0.01	0.01	0.01
Arrival Dist.:	90.99	91.02	90.94	90.99	91.02
Throughput Sum:	2746	2746	2749	2748	2747
Throughput / Time:	0.01	0.01	0.01	0.01	0.01
Throughput Dist.:	91.02	91.03	90.93	90.95	90.99
Waiting Time:	91.00	22.50	39.47	113.96	0.02
Queue length:	1.00	0.25	0.43	1.25	0.00
Place Name:	R				
Arrival Sum:	5494				

Figure 3. Statistics Offered by the Analyzer in PN Toolbox Corresponding to the Simulation of the FMS model with 3 Pallets and Deterministic Delays

In applying strategy (1<sup>0</sup>), in order to ensure sequential mutual exclusion, a total of 3 pallets must be utilized. The overall temporal statistics provided by the analyzer in PN Toolbox for the places in the net are presented in figure 3. The mean time necessary for processing a part is 91.03 s. The mean time M1 and M2 are busy processing a part are 51.58 s and 91 s respectively, both of them longer than the effective operation times. The mean time a part has to wait in the buffer is 67.97 s, the mean number of parts waiting in the buffer being 0.75. The utilization of the robot is 0.49.

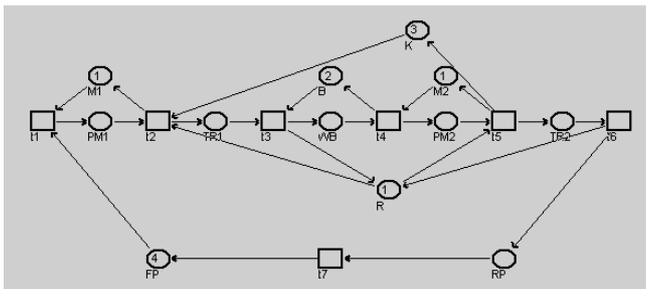


Figure 4. PN model with Kanban Control of the FMS with 4 Pallets

When using 4 pallets to fix on the parts, for the considered numerical values of the delays deadlock occurs. Figure 4 presents the PN model corresponding to the usage of a kanban (K) for limiting the number of tokens (parts) in the critical subnet involved in the circular wait, explicitly delineated in figure 2.

The results obtained through simulation when applying strategies (2<sup>0</sup>) and (3<sup>0</sup>) were similar. The mean time for

processing a part decreased to 80.03 s. The mean time M1 and M2 are busy processing a part are 62.47 s and 80 s respectively, the first one longer than the effective operation time on M1. The mean time a part has to wait in the buffer is 114.94 s, the mean number of parts waiting in the buffer being 1.44. The utilization of the robot is 0.56.

The cases of utilizing 5 pallets with strategies (2<sup>0</sup>) and (3<sup>0</sup>) were also analyzed. All the performances of this system were the same as when using 4 pallets, except for the mean time a part has to wait on M1 which increased up to 80 s, meaning that M1 was, in fact, continuously busy.

In order to test the sensitivity of the performances of this FMS to the variations in the delays assigned to the operations in the system, a *stochastic* P-timed model was next taken into consideration. Assuming that practical experience is able to provide a set of reasonable lower and upper bounds for all the delays the uniform distribution was chosen. The following numerical values were selected for the simulation experiments:

- automatic fix on pallet:  $d(FP) \in Unif(6, 9)$ ;
- unload from a machine with the robot:  $d(TR1), d(TR2) \in Unif(20, 25)$ ;
- processing on M1:  $d(PM1) \in Unif(35, 45)$ ;
- processing on M2:  $d(PM2) \in Unif(75, 85)$ ;
- release pallet:  $d(RP) \in Unif(8, 12)$ ;

where  $x \in Unif(\min, \max)$  denotes a random variable uniformly distributed in the interval  $[\min, \max]$ .

In all the studied cases, the performances of the FMS did not vary significantly from those obtained in the corresponding deterministic case. A similar study within

the stochastic context, but testing exponentially distributed delays with mean values equal to the deterministic delays used above, provided the same sort of results as in the case of uniform distributions.

## 5. CONCLUSIONS

The paper presents the role of simulation-based study of efficient exploitation of FMSs in connection with the deadlock avoidance problem. It emphasizes the need for employing timed models in order to incorporate both qualitative and quantitative information associated with FMSs and perform an effective analysis. The examples herein presented demonstrate that the selection of the most adequate policy in controlling the considered production flow cannot be achieved without relevant simulation tests.

Thus, the simulation-based analysis leads to the conclusion that the most efficient approach to exploit the considered FMS consists in employing 4 pallets to fix the processed parts on and ensuring deadlock avoidance through kanban control or priorities set between the operations that use the shared resource. A deeper insight into the efficiency of an FMS utilization might be reached by assigning to the places in the PN model that correspond to the operations cost functions for the utilization of the resources, constructing the P-timed version of Stochastic Reward Petri Nets (Muppala et al., 1994).

## 6. REFERENCES

- Bacelli, F., G. Cohen, G.J. Olsder and J.P. Quadrat. 1992. *Synchronization and Linearity, an Algebra for Discrete Event Systems*. Wiley, New York.
- Cassandras, C. G. 1993. *Discrete Event Systems: Modeling and Performance Analysis*. R.R. Donnelley & Sons, Boston.
- Cohen, G., S. Gaubert, and J.P. Quadrat. 1998. "Max-plus algebra and system theory: where we are and where to go now". *IFAC Conference on System Structure and Control*, Nantes, France, 8-10 July 1998.
- David, R. and H. Alla. 1992. *Du Grafcet aux réseaux de Petri (2e édition)*, Hermes, Paris.
- Desrochers, A.A. and R.Y. Al-Jaar. *Modeling and control of automated manufacturing systems*. IEEE Computer Society Press, Rensselaer, Troy, New-York, 1993.
- Ezpeleta, S.D., J.M. Colom, and J. Martinez. 1995. "A Petri Net Based Deadlock Prevention Policy for Flexible Manufacturing Systems". *IEEE Trans. on Robotics and Automation*, RA-11, 173-184.
- Lewis, F., A. Gurel, S. Bogdan, A. Doganalp, and O. Pastravanu. 1998. "Analysis of Deadlock and Circular Waits Using a Matrix Model for Flexible Manufacturing Systems". *Automatica*, vol. 34, no. 9, 1083-1100.
- Lewis, F., H.H. Huang, O. Pastravanu, and A. Gurel. 1995. "Control Systems Design for Flexible Manufacturing Systems". In: (Raouf, A. and M. Ben-Daya, Eds.) *Flexible Manufacturing Systems: Recent Developments*, Elsevier Science, 259-290.
- Mahulea, C., L. Bârsan and O. Pastravanu. 2001. "MATLAB Tools for Petri-net-based Approaches to Flexible Manufacturing Systems". In: (Filip, F.G., I. Dumitrache and S. Iliescu, Eds.) *Proceedings Volume from the 9th IFAC Symposium on Large Scale Systems LSS 2001*, Bucharest, Romania, 18-20 July 2001, 184-189.
- Matcovschi, M., C. Mahulea, and O. Pastravanu. 2001. "Exploring Structural Properties of Petri Nets in MATLAB". In: *Proceedings Volume from the 7th International Symposium on Automatic Control and Computer Science and Parallel Workshop on Control Theory SACCS 2001*, Iasi, Romania, CD-ROM.
- The MathWorks, Inc. 2000a. *Using MATLAB*. Natick, Massachusetts.
- The MathWorks, Inc. 2000b. *Building GUIs with MATLAB*. Natick, Massachusetts.
- Muppala, J.K., G. Ciardo, and K.S. Trivedi. 1994. "Stochastic Reward Nets for Reliability Prediction", *Communications in Reliability, Maintainability and Serviceability: An International Journal published by SAE International*, vol. 1, no. 2, 9-20.
- Murata, T. 1989. "Petri Nets: Properties, Analysis and Applications". In: *Proceedings of the IEEE*, no. 77, 541-580.
- Sugimori, Y., K. Kusunoki, F. Cho and S. Uchikawa. 1977. "Toyota Production System and Kanban System Materialization of Just-In-Time and Respect-For-Human System". *International Journal of Production Research*, vol. 15, no. 6, 553-564.
- Viswanadham, N., Y. Narahari and T.L. Johnson. 1990. "Deadlock Prevention and Deadlock Avoidance in Flexible Manufacturing Systems Using Petri Net Models". *IEEE Trans. on Robotics and Automation*, RA-6, 713-723.
- Wysk, R.A., N.S. Yang and S. Joshi. 1991. "Detection of Deadlocks in Flexible Manufacturing Cells". *IEEE Trans. on Robotics and Automation*, RA-7, 853-859.
- Zhou, M.C. and F. DiCesare. 1993. *Petri Net Synthesis for Discrete Event Control of Manufacturing Systems*. Kluwer Academic Publishers, Boston.

## ABBREVIATIONS

FMS – Flexible Manufacturing System  
PN – Petri Net

## AUTHOR BIOGRAPHY

**MIHAELA-HANAKO MATCOVSCHI** received the M.S. in Mathematics from the University "Alexandru Ioan Cuza" of Iasi, Romania, in 1986, the M.S. in Control and Computer Engineering and the Ph.D. in Industrial Engineering from the Technical University "Gheorghe Asachi" of Iasi in 1999 and 2001, respectively. She is currently a Lecturer at the Department of Automatic Control and Industrial Informatics of the Technical University "Gheorghe Asachi" of Iasi. Her current research interests include discrete event systems and applications of neural networks in Control Engineering.

**OMRAN HILAL** owns a M.S. degree in Electrical Engineering. As a head of department at the TOLEDO Electrical & Mechanical Works Est., Abu Dhabi, UAE, he is responsible for design estimation and project management.

# TPM<sup>2</sup>: AN INTEGRATED DECISION SUPPORT PARADIGM FOR PRODUCTIVITY IMPROVEMENT USING SIMULATION

Mousalam A. Razzak  
Intelligent Quality Systems Ltd.  
333 14<sup>th</sup> St. 4<sup>th</sup> Floor  
Toledo, OH, 43624 USA  
1-419-724-0505  
[mrzzak9@hotmail.com](mailto:mrzzak9@hotmail.com)

George C. Daley  
Effective Metrix, Inc.  
Toledo, OH, USA

John P. Dismukes and  
Lawrence K. Miller  
Manufacturing Value Chain  
Science Center  
The University of Toledo  
Toledo, OH, USA

## KEYWORDS

Total Productive Manufacturing Methodology (TPM<sup>2</sup>), Total Productive Maintenance (TPM), Lean Manufacturing (LM), Theory of Constraints (TOC), Discrete Event Simulation.

## ABSTRACT

Global competition is forcing manufacturers to continuously seek productivity improvement solutions. Conventional productivity metrics, such as throughput and utilization rate, only measure part of equipment performance and are not very informative when it comes to identifying areas of improvement. Although advanced techniques, such as Total Productive Maintenance (TPM), have delivered significant improvement gains, results have only been proven at the equipment level. These results are insufficient to manufacturers who not only need productive equipments, but more importantly, efficient and productive systems. In this paper, the newly developed Total Productive Manufacturing Methodology (TPM<sup>2</sup>) is introduced as a paradigm which integrates system level productivity metrics and discrete event simulation to implement a closed loop continuous improvement cycle. This approach provides a basis for the next generation of lean manufacturing practices. Also presented, is the proposed implementation of these techniques through commercial web-based applications.

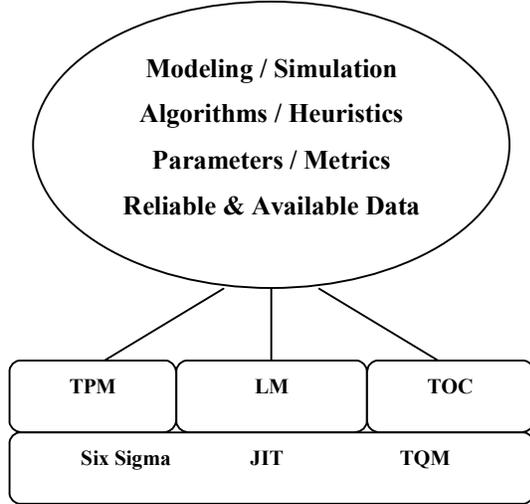
## I. INTRODUCTION

Keeping ahead of the game is tougher than ever in today's manufacturing environment. Competition is worldwide and markets are fast becoming price sensitive. These challenges are forcing companies to implement various productivity improvement efforts to meet the needs

of ever changing market demand. One of the major features of TPM (Ljungberg 1998), launched by Seiichi Nakajima (Nakajima 1988) in the 1980's, is maximizing asset utilization or overall equipment effectiveness (OEE) by gauging equipment productivity and focusing improvement tasks to eliminate equipment losses. A second theory that follows the same line of thought is the Theory of Constraints (TOC) (Dettner 1997), which proposes that physical constraints arise due to bottleneck equipments. TOC further proposes that eliminating losses at a bottleneck will significantly enhance productivity, and that non-bottleneck equipments do not impact productivity to the same extent as bottlenecks do.

Conclusions from both theories, TPM and TOC, ignore the interdependency among production steps. Although the idea of insignificant contribution by non-bottlenecks holds when compared versus gains from focusing on bottleneck equipments, the impact of production disturbances is becoming more felt as a result of the transition of manufacturers to continuous production. As illustrated in Figure 1, a third theory for effective discrete part manufacturing, with minimum build-up of work-in-process, is Lean Manufacturing (LM) (Womack 1996) which addresses productivity optimization in factories with continuous flow of sequential, discrete manufacturing operations. Other well known productivity practices in Figure 1 underpinning these three theories are Six Sigma, JIT, and TQM (Evans and Lindsay 1999). Waste-free, single part flow or continuous production, the ultimate goal of Lean Manufacturing, is increasingly gaining a positive reputation as being the solution to implementing Just-In-Time manufacturing in production systems (e.g. factories and supply chains).

## Integrated TPM<sup>2</sup> Decision Support Paradigm



**Figure 1: Integrated TPM<sup>2</sup> Decision Support Paradigm Linking Equipment and Factory Level Productivity To Facilitate System Optimization Via Simulation**

Recent publications by Scott (Scott 1999) recognize and analyze the need for a coherent systematic methodology for productivity measurement and analysis at the factory level. As Scott (Scott 1999) pointed out, manufacturing is a complex web of highly inter-dependent activities: interactions among machines, tools, materials, people, testers, processes, departments, and company. Overall factory efficiency is about combining activities, the relationships between different machines and processes, integrating information, decisions, and actions across many independent systems and sub-systems.

However, these inter-dependent activities cannot be isolated, as eloquently discussed in the tutorial first chapter in the manufacturing textbook, "Factory Physics: Foundation of Manufacturing Management" (Hopp and Spearman 2001). Too often they are viewed in isolation, and there is a lack of coordination in deploying available factory resource (people, information, materials, tools) to manage work efficiently. The gains made in OEE, while important and ongoing, are insufficient. It is necessary to focus one's attention beyond the performance of individual tools towards the performance of the whole factory. The ultimate objective is a highly efficient integrated system, not brilliant individual tools. A literature survey indicates that up to this time, there has been no single, well defined, proven methodology for analysis of overall factory performance. The TPM<sup>2</sup>

paradigm summarized in this paper fulfills this need, as described below.

More recently, researchers at the University of Toledo, in collaboration with manufacturers of Tier 1 and Tier 2 suppliers of the automotive industry in the US, have designed a new set of algorithms allowing manufacturers for the first time to analyze system level productivity (Razzak 2002) as a function of system complexity and productivity of individual units. The new approach serves as a tool that drives incremental improvements to the manufacturing system through a closed loop cycle that integrates simulation results to identify improvement opportunities.

## II. SYSTEM LEVEL PRODUCTIVITY

In order to take all equipments into consideration when analyzing a production system, it is required to use a system level metric (Razzak 2001; Su et al. 2002b) analogous to OEE, and designated as Overall Throughput Effectiveness (OTE). OEE is the ratio of actual good parts produced by a unit production process (UPP) to the theoretical number of parts the UPP is capable of producing during a certain time period,

$$OEE = \frac{\text{Number of Good Parts Produced by UPP in Total Time}}{\text{Theoretical Number of Parts Produced by UPP in Total Time}} = A \times P \times Q = \frac{P_g}{R_{th} T_T} \quad (1)$$

where, for a UPP, A is availability efficiency, P is performance efficiency, and Q is quality efficiency.  $P_g$  is throughput,  $R_{th}$  is the theoretical processing rate of the UPP, and  $T_T$  is the total time of observation.

Similarly, OTE is the ratio of good parts produced by a system to the number of parts the system theoretically could produce during a certain time period,

$$OTE = \frac{\text{Number of Good Parts Produced by System in Total Time}}{\text{Theoretical Number of Parts Produced by System in Total Time}} = A_F \times P_F \times Q_F = \frac{P_G}{R_{TH} T_T} \quad (2)$$

Parameter definitions at the system level are analogous to equipment level parameters, where subscript F implies factory or system level.

OTE and other system level metrics (Dismukes 2001, Su et al. 2002a) are determined as a function of material flow between UPPs. OTE of a system is calculated first by calculating OTE of sub-

systems formed by grouping UPPs according to material flow in predetermined group structures similar to those in Figure 2. Sub-system groups are combined further in the same fashion until the system is reduced to a single unit, which yields an overall value of OTE (Razzak 2001; Dismukes 2002b).

Calculating system productivity in terms of OTE provides robust metrics that are very useful especially when multiple part types are produced by the same system. A dynamic production system is one that is producing multiple part types each with a different recipe of production such as seen in the semiconductor industry. In these environments bottlenecks are continuously shifting from one equipment to another. Non-bottleneck equipments may not have a direct impact on throughput especially when space is allocated for building a limited level of inventory.

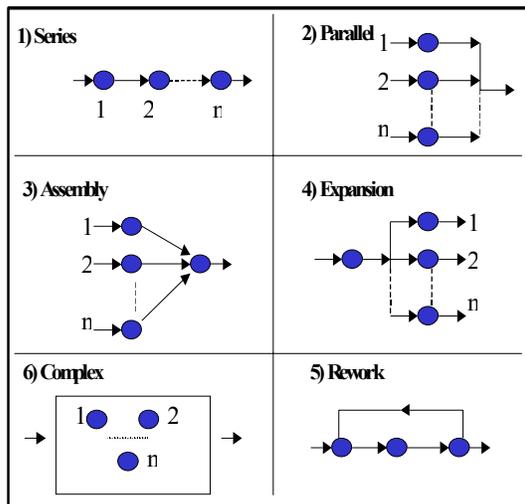


Figure 2: Generic Sub-System Combinations

### III. TPM<sup>2</sup> CLOSED LOOP IMPROVEMENT METHODOLOGY

The closed loop improvement methodology presented in this paper (Razzak 2001) utilizes some of the principles and practices exhibited in the teachings of TPM, TOC, and LM. However, it goes beyond these three practices in presenting a systematic process comprising four stages, each with multiple subtasks. As schematically illustrated in Figure 3, the TPM<sup>2</sup> closed loop methodology is a powerful tool for managing and optimizing productivity improvement activities.

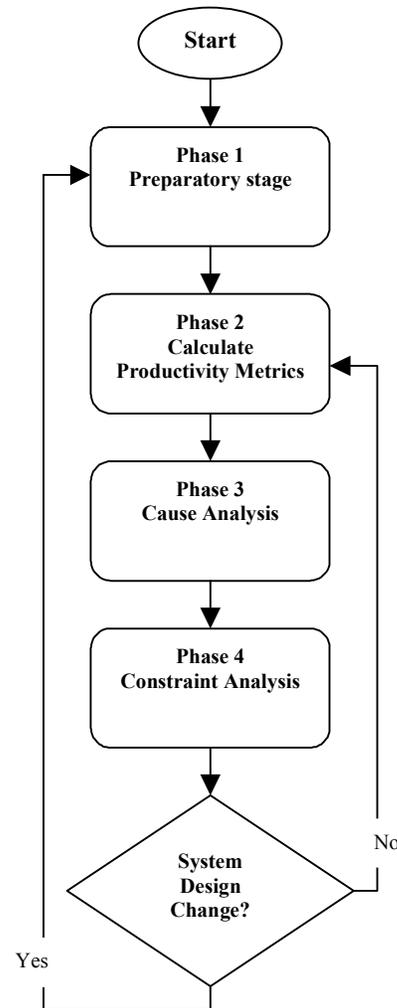


Figure 3: TPM<sup>2</sup> Closed Loop Methodology

The following is a summary of each step:

- Phase 1: Preparatory stage
  - Flowchart system architecture.
  - Define productivity parameters.
- Phase 2: Calculate Productivity Metrics
  - Collect production data.
  - Apply productivity metrics.
  - Calculate productivity at the equipment, sub-system, and system level
- Phase 3: Cause Analysis
  - Identify bottleneck and associated losses
  - Identify upstream losses
  - Identify downstream losses

- Phase 4: Constraint Analysis
  - Identify constraint.
  - Decide how to relieve constraint.
  - Manage improvement decision.
  - Eliminate constraints.
- System Design Changed?
  - Repeating the cycle takes one of two different paths depending on the improvement implemented. If no changes are required in the system design then the cycle may be repeated from phase 2. On the other hand, if improvement introduces a change to the system design, then the cycle needs to be repeated from phase 1.

#### IV. SIMULATION IN TPM<sup>2</sup>

During the last two decades, attempts have been made to model, analyze, and design different aspects of manufacturing systems. Modeling methods reported in the literature include:

- Graph with Results and Actions Interrelated (GRAI)
- Integrated computer-aided manufacturing Definition (IDEF0)
- Structured Analysis and Design Technique (SADT)
- Structured Systems Analysis and Design Method (SSADM)

These models share common elements, which are built on graphical tools and a structured approach of a hierarchy of diagrams, text, and glossary. In general, all these methods have advantages and disadvantages in modeling a manufacturing system. They can identify major elements of the system and model it in a hierarchical structure, by breaking down the whole system into low-level units of a factory. Flows of material, control and information can also be simulated. However, as mentioned by some researchers, these methods are “no more than static graphical representation”. Although this argument was not sufficiently justified, it can be said that, for modeling a real-world manufacturing system, a general structure is not enough. (Huang, et. al. 2002b)

In addition, these modeling techniques focus on the availability of the equipment, which is only one aspect of the system performance. More comprehensive integrative productivity metrics,

such as Overall Equipment Effectiveness (OEE) are needed to assess system productivity.

Analysis resulting from creating multiple what-if scenarios using discrete event simulation is considered the most reliable method to date in studying the dynamic performance of manufacturing systems closely reflecting real factory behavior (Huang 2002a). However, existing commercial simulation packages are still using traditional metrics such as utilization rates to measure performance of manufacturing facilities and lack the ability to readily provide information on OEE and OTE.

Ongoing research and development efforts at the University of Toledo and Effective Metrix Inc. envision the integration of the capabilities of discrete event simulation into a web-based commercial tool designed to provide real-time productivity improvement solutions.

Case studies from real-world manufacturing facilities provided promising results that demonstrate the usefulness of applying the methodology. One of the case studies carried out recently involved analyzing productivity of a glass manufacturing facility. After accumulating actual shop floor data, the productivity algorithms and metrics were used to determine the current status of the individual equipments and the overall system. The next step was the analysis of simulation results to identify potential areas of improvement. These opportunities were determined according to the different sets of loss eliminations introduced as variant parameters used by simulation.

To view a detailed description of the case study refer to “Manufacturing Productivity Improvement Using Effectiveness Metrics and Simulation Analysis” by Huang et al. 2002.

#### V. CONCLUSIONS

The TPM<sup>2</sup> paradigm integrates practices and principles featured in different advanced productivity improvement techniques (e.g. TPM, TOC, LM, Six Sigma, JIT, TQM) with system level productivity metrics. The resulting paradigm allows for a holistic analysis of the manufacturing system as determined by productivity of individual equipments and material flow.

Manual implementation of the closed loop methodology becomes more difficult as the system

becomes more complex. Also, the fast paced production environments require expediting the decision making process of identifying improvement tasks. Therefore, a computer-based application is recommended to facilitate this task, which would offer ongoing productivity assessment and improvement by utilizing simulation capabilities.

A web-based application utilizing the presented methodology resulting from research at The University of Toledo is currently planned for development and implementation by Effective Metrix, Inc. to monitor and analyze manufacturing productivity for several clients in the plastics and composites industries. The advantage of this deployment will be based on the convenience of accessing real-time production data and system level analysis from anywhere in the world through a browser interface, allowing for remote monitoring and management. Another advantage is the low upfront capital expenditure needed since the application is offered as a service according to the Application Service Provider (ASP) model.

## REFERENCES

- Dettner H. W. 1997. "Goldratt's Theory of Constraints", ASQC Quality Press, Milwaukee.
- Dismukes J. P.; Q. Su; S. H. Huang; M. A. Razzak; G. Wang; and R. Kothamasu, 2001. "Hierarchical Methodology For Productivity And Improvement Of Production Systems", *International Patent Convention Application*, Filed December 19, 2001.
- Dismukes J. P.; S. H. Huang; Q. Su, G. Wang; M. A. Razzak; and D. E. Robinson. 2002a "Measurement and Improvement of Factory Level Productivity", *NSF Design and Manufacturing Conference* (Jan).
- Dismukes J. P., Q. Su, S. H. Huang, M. A. Razzak, G. Wang, R. Kothamasu, 2002b. "Hierarchical Methodology for Productivity and Improvement of Production Systems", *International Patent Convention Application, PCT WO 02/50699 A1, June 27, 2002*.
- Evans J. R and W. M. Lindsay 1999. "The Management and Control of Quality". South-Western. pp. 42-43.
- Hopp W. J. and M. L. Spearman 2001. "Factory Physics: Foundation of Manufacturing Management", 2nd Edition, McGraw-Hill, Part I.
- Huang S. H.; J. P. Dismukes; Q. Su; M. A. Razzak; R. Bodhale; and D. E. Robinson 2002a. "Manufacturing Productivity Improvement Using Effectiveness Metrics and Simulation Analysis", June 2002, *International Journal of Production Research*, In Press.
- Huang S. H.; J. P. Dismukes; Q. Su; G. Wang; M. A. Razzak; and D. E. Robinson 2002b. "Manufacturing System Modeling For Productivity Improvement", June 2002, *Journal of Manufacturing Systems*, In Press.

- Ljungberg O. 1998. "Measurement of Overall Equipment Effectiveness as a Basis for TPM Activities", *International Journal of Operations and Production Management*, Vol. 18 No. 5, pp.495-507.
- Razzak M. A. 2001. "Metrics For Factory Level Productivity", *MS Thesis, University of Toledo* (May).
- Razzak M.A.; G. Daley; and J. P. Dismukes 2002. "Factory Level Metrics: Basis for Productivity Improvement", *Proceedings of the International Conference on Modeling and Analysis of Semiconductor Manufacturing 2002*, pp. 158-162, Tempe, Arizona, April 10-12.
- Scott, D. 1999, "Can CIM Improve Overall Factory Effectiveness?" *Pan Pacific Microelectronics Symposium*, Kauai, HI
- Su Q.; J. P. Dismukes; S. H. Huang; and M. A. Razzak 2002a. "Factory Level Metrics for Manufacturing Productivity," February 2002, *IEEE Transactions on Robotics and Automation*, In Review.
- Su Q.; J. P. Dismukes; M. A. Razzak; and V. A. Ames 2002b. "Metrics for Measuring Productivity of Flexible-Sequence Cluster Tools", *Proceedings of the International Conference on Modeling and Analysis of Semiconductor Manufacturing 2002*, pp. 3-13, Tempe, Arizona, April 10-12.
- Womack J. P. and D. T. Jones 1996. "Lean Thinking", Simon & Schuster.

## BIOGRAPHY

**Mousalam A. Razzak** is Program Manager at Intelligent Quality Systems Ltd., 333 N 14<sup>th</sup> Street, 4<sup>th</sup> floor, Toledo, Ohio, 43624. He holds dual M.Sc degrees in Industrial Engineering and in Chemical Engineering from The University of Toledo.  
(mrazzak9@hotmail.com)

**George C. Daley** is President and CEO of Effective Metrix, Inc., 333 N. 14<sup>th</sup> Street, 4<sup>th</sup> Floor, Toledo, Ohio. Dr. Daley holds a Ph.D. in Systems Engineering from The University of Toledo.  
(cgdaley@sprynet.com).

**John P. Dismukes** is Professor, Department of Chemical and Environmental Engineering, and Director, Manufacturing Value Chain Science Center (MVCS Center) at The University of Toledo. He holds a Ph.D. in Inorganic Chemistry from The University of Illinois, and has 30 years industrial R&D experience.  
(John.Dismukes@utoledo.edu).

**Lawrence K. Miller** is Assistant Professor, Department of Electrical Engineering and Computer Science at The University of Toledo. He holds a Ph.D. in Computer Science from The University of Houston, and conducts research in modeling and simulation of information and material flow.  
(lmiller@eecs.utoledo.edu).



**SIMULATION  
IN  
IMAGE  
PROCESSING**



# SPREAD SPECTRUM DIGITAL IMAGE WATERMARKING: SIMULATION AND PERFORMANCE EVALUATION

*Mohammed E. Al-Mualla and Sultan R. Sharar*

Etisalat College of Engineering,  
P.O.Box: 980, Sharjah, United Arab Emirates,  
Tel: +97165611333, Fax: +97165611789  
e-mail: [almualla@ece.ac.ae](mailto:almualla@ece.ac.ae)

## KEYWORDS

Spread spectrum, image watermarking, copyright protection

## ABSTRACT

The significant increase in the distribution of images in digital format, and the ease with which such images can be copied and modified has created a need for copyright protection methods. One possible solution is the use of digital image watermarking. This paper simulates and investigates the performance of three spread-spectrum digital image watermarking techniques: the Cox method, the Barni method, and a new block-based version of the Cox method. It is shown that this new version provides a good compromise between robustness and quality of the watermarked image, and at the same time it has good payload capacity.

## INTRODUCTION

In recent years there has been a significant increase in the distribution of images in digital format. The ease in which such images can be copied and modified has created a pressing need for the development of copyright protection methods. A promising technique in this direction is *digital image watermarking*.

In digital image watermarking, copyright protection information are embedded in the image in the form of a *watermark*. The image must not be visibly degraded by the presence of this watermark. The watermark must be resistant to unauthorised detection and decoding. In addition, the watermark must be tolerant to normal image processing techniques (e.g. compression), as well as to intentional attacks (attempts to remove the watermark).

Digital image watermarking techniques can be classified according to a number of different criteria. One such criterion is whether the original image is needed in the watermark extraction process. If the original unwatermarked image is needed to recover the watermark from the watermarked image, then the technique is called *complete*, otherwise it is *incomplete*.

There are many digital image watermarking techniques reported in the literature (Katezenbeisser and Petitcolas

2000). An important category of techniques are called *spread spectrum* techniques. In such techniques, the image is viewed as a communication channel, the watermark is viewed as a signal that is transmitted through it, and attacks and signal distortions of any type are viewed as noise. In an analogy with secure spread spectrum communication, the watermark is spread over the image (in the spatial-domain or in the frequency-domain) such that the energy in any one element (spatial pixel or frequency coefficient) is very small and undetectable. During the watermark extraction process, these many weak signals are concentrated into a single output with a high signal-to-noise ratio.

This paper simulates and evaluates the performance of a number of spread spectrum image watermarking techniques.

## SIMULATED ALGORITHMS

### The Cox Method

This is a complete frequency-domain technique reported in (Cox et al. 1997). In this method, the original image  $D$  is transformed into the frequency-domain using the discrete-cosine-transform (DCT). The largest  $N$  frequency coefficients (excluding the DC coefficient) are taken to form a vector  $V=[v_1, v_2, \dots, v_N]$  into which the watermark is inserted. The watermark  $X = [x_1, x_2, \dots, x_N]$  is generated randomly from a Gaussian distribution with mean of 0 and variance of 1. The watermark is then inserted into the coefficients vector as follows:

$$v'_i = v_i(1 + \alpha x_i) \quad (1)$$

where  $\alpha$  is a scaling parameter that determines the amount by which  $X$  alters  $V$ . The altered coefficients are then inserted back into the DCT of the image. Inverse DCT is then applied to obtain the watermarked image  $D'$ . One or more attacks may then alter  $D'$  producing a new image  $D^*$ . In the extraction process, the DCT is applied to the original image  $D$  and the locations of the largest  $N$  coefficients are identified. Coefficients in these locations are then extracted from the DCT of the original image  $D$  and the DCT of the attacked image  $D^*$  to give vectors  $V$  and  $V^*$  respectively. The watermark is then extracted as follows:

$$x_i^* = (v_i^* / v_i - 1) / \alpha \quad (2)$$

This extracted watermark is then correlated with the original watermark.

### The Barni Method

This is an incomplete frequency-domain technique reported in (Barni et al. 1998). This method is very similar to the Cox method. The major difference here is that the watermark is spread over a predetermined set of coefficients. Thus, the DCT coefficients of the original image are first re-ordered using a zig-zag scan. The coefficients vector  $V$  is then formed by taking the  $(L+1)^{th}$  to the  $(L+N)^{th}$  coefficients from this re-ordered set. The embedding equation in this case is:

$$v'_i = v_i + \alpha |v_i| x_i \quad (3)$$

Since the original image is not available during the extraction process, it is impossible to get an estimate of the watermark. Instead, the correlation between the extracted watermarked, and possibly corrupted, coefficients  $V^*$  and the original watermark  $X$  itself is taken as a measure of the watermark presence.

### The Block-Based Cox Method

In addition to the above two techniques, we introduce a new technique called the block-based Cox (BCOX) method. As the name implies, this is a block-based version of the Cox method. The original image in this case is divided into blocks. The  $N$ -length watermark is broken down into smaller watermarks of length approximately equal to  $M=N/B$ , where  $B$  is the number of blocks in the image. The embedding process treats each block individually. Thus, each block is independently DCT transformed and its watermark is inserted. In the extraction process, a watermark is extracted from each block and all the watermarks are combined to form an  $N$ -length extracted watermark.

## RESULTS AND DISCUSSIONS

The three algorithms Cox, Barni, and BCox were tested using a number of standard test images. The performance was evaluated using different parameters and under different types of attacks. Due to space limitations, only a small subset of results is shown here. Thus, unless otherwise stated, the results in this paper refer to the  $256 \times 256$  Cameraman image with parameters  $N=1000$ ,  $L=1000$ ,  $\alpha=0.1$ , and  $8 \times 8$  blocks. For more extensive results, the reader is referred to (Sharar 2002).

Figures 1 and 2 show the effect of changing the strength  $\alpha$  of the embedded watermark on the normalized correlation measure and the peak-signal-to-noise-ratio (PSNR in dB) measure, respectively. A higher normalized correlation measure indicates a higher confidence that the watermark is present in the watermarked image. It also indicates a higher robustness against attacks. A higher PSNR indicates that the quality of the watermarked image is closer to that of the original image.

As expected, increasing the watermark strength improves the chances of detecting the watermark and also increases its robustness against attacks (see Figure 1), this is however at the expense of more degradation in the quality of the watermarked image (see Figure 2). In general, the most robust technique is the Cox method, followed by the BCox, and then the Barni method, which trades robustness for incompleteness. On the other hand, the Barni method has the least effect on the quality of the watermarked image due to the use of the offset parameter  $L$  which preserves some important low frequency coefficients.

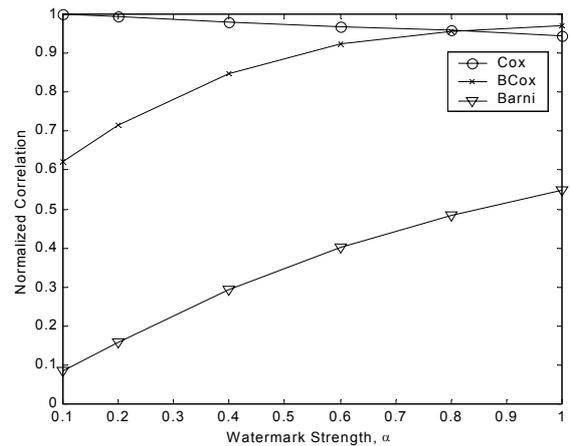


Figure 1: Effect of  $\alpha$  on Correlation

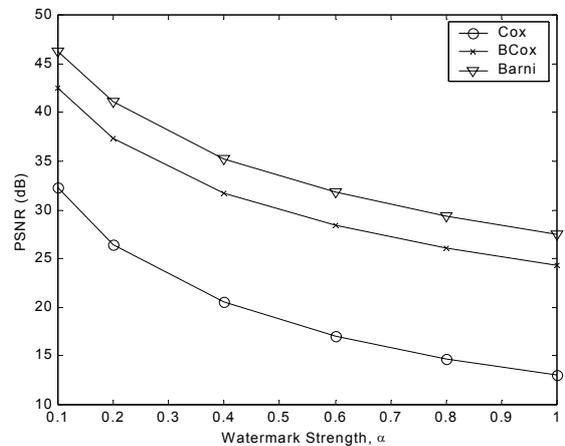


Figure 2: Effect of  $\alpha$  on PSNR

Figure 3 shows the effect of changing the watermark length  $N$  on the normalized correlation measure. As can be seen, the Cox and BCox methods are not sensitive to changing  $N$ . With the Barni method, however, there is a significant drop in correlation as  $N$  increases. This indicates that the Barni method has less capacity. That is, it can only support shorter watermarks.

Figures 4 shows the effect of changing the watermark length  $N$  on the PSNR measure. As expected, increasing the watermark length degrades the quality of the watermarked

image since now more coefficients are modified. The BCox method seems to be particularly sensitive to this effect although the quality is still better than the COX method even with a watermark length of  $N=10,000$ .

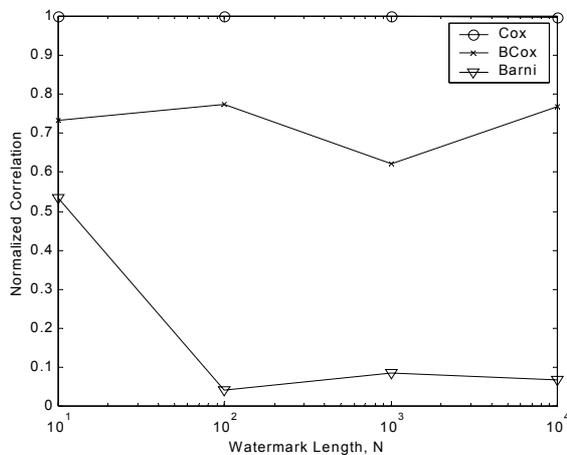


Figure 3: Effect of  $N$  on Correlation

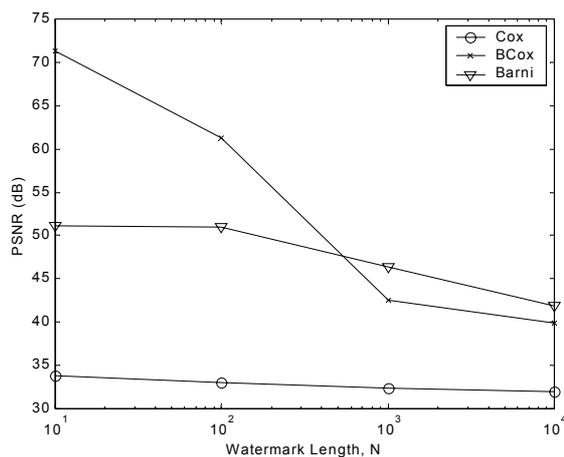


Figure 4: Effect of  $N$  on PSNR

Figure 5 shows the performance of the three algorithms under the JPEG attack. In this case the watermark is inserted in the image, the image is then subjected to JPEG compression (a JPEG compression quality of 1 corresponds to the highest compression and consequently to the worst visual quality), and then the watermark is extracted. As indicated by the very low correlation measure, Figure 5 confirms our previous conclusion that the Barni method is the least robust technique against attacks.

Figure 6 shows the effect of the JPEG attack on the quality of the watermarked image. As expected, higher compression (indicated by a lower quality number) leads to more degradation in the quality of the watermarked image.

## CONCLUSIONS

This paper investigated the performance of three spread-spectrum digital image watermarking techniques. It was

found that the COX method is the most robust technique. This is, however, at the expense of a lower-quality watermarked image. It was also found that the Barni method has the least capacity as it can only support short watermarks. In general, the BCox method provides the best compromise between robustness and quality of the watermarked image, and at the same time it has good payload capacity.

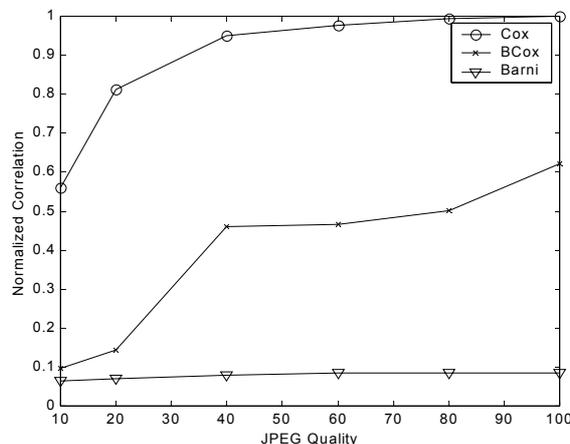


Figure 5: Effect of JPEG Attack on Correlation

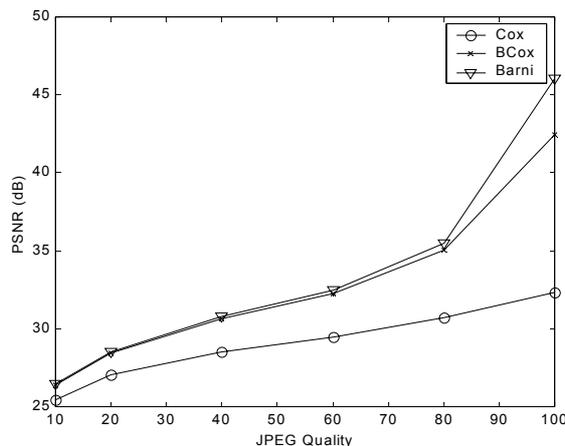


Figure 6: Effect of JPEG Attack on PSNR

## REFERENCES

- Barni, M, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain System for Robust Image Watermarking," *Signal Processing*, Vol. 66, 1998, pp. 357-372.
- Cox, I. J., J. Kiliant, T. Leighton, and T. Shmoon, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Transactions on Image Processing*, Vol. 6, No. 12, 1997, pp. 1673-1687.
- Katezenbeisser, S and F. Petitcolas. *Information Hiding: Techniques for Steganography and Digital Watermarking*. Artech House, 2000.
- Sharar, S. *Digital Watermarking of Images and Videos*. Final Year Project Report. Etisalat College of Engineering, May 2002.

# DATA EMBEDDING IN IMAGES USING CONVOLUTIONAL CODING

Nidhal Abdulaziz,  
K. Khee Pang  
*Dept of Elect & Comp Systems Engineering,*  
Monash University, Abu-Dhabi, UAE  
Clayton, VIC 3168, Australia  
E-mail: [nidhal.abdulaziz@eng.monash.edu.au](mailto:nidhal.abdulaziz@eng.monash.edu.au)  
E-mail: [khee.pang@eng.monash.edu.au](mailto:khee.pang@eng.monash.edu.au)

Abdullatif Glass  
Department of Planning and Training  
Technical Studies Institute,  
E-mail: [aglass@ieee.org](mailto:aglass@ieee.org)

## KEYWORDS

Data hiding, image coding, source and channel coding, wavelets.

## ABSTRACT

This paper presents an approach for image data embedding based on discrete wavelet transform and convolutional coding. We address the problem of embedding images of  $1/4^{\text{th}}$  the host image size into a host image. The algorithm starts with transforming each image into the discrete wavelet domain where their coefficients are grouped into vectors equivalent to the codeword length of the signature data. Maximum likelihood decoding using Viterbi algorithm is employed in extracting the hidden information. The results of embedding a coded and uncoded signature are compared and an average of 5dB SNR is needed for the uncoded case to obtain the same bit error rate. Comparison between our simulation results and theoretical ones are established for the curves of bit error rate versus signal to noise ratio. The method implemented provides robustness against widely varying signal distortions including compression and noise addition.

## INTRODUCTION

Data hiding can be thought of as a special communication problem, where the host image constitutes the channel for transmission of the watermark data and is subject to various types of attacks, and the signature data are the information to be sent from sender to receiver. It has been pointed out in the literature that similarities between the data hiding problem and digital communication can be utilized to improve the performance of the system (Tewfik 2000). Note that attacks such as lossy compression, enhancements, or transformations can be treated as noise addition. Lossy compression algorithms such as JPEG might severely narrow the channel by totally

disregarding large regions of the frequency spectrum of the image. However, using convolutional coding techniques can greatly improve the robustness of the embedded data against compression and standard digital image processing operations.

The basic requirements for many data hiding applications are imperceptibility, robustness against moderate compression and processing, and the ability to hide many bits. The traditional way to handle these contradictory requirements is to target at a specific capacity-robustness pair. Some approaches choose to robustly embed just one or a few bits (Barni et al., 1999), (Cox et al., 2000), (Langelaar et al., 2000), while others choose to embed a lot of bits but to tolerate little or no distortion (Stego). Our scheme is different from the above mentioned, firstly, the embedding is implemented in the wavelet domain and not in the spatial domain, secondly, we propose to use new schemes adapted from the area of deep-space communications such as convolutional coding.

In this paper we investigate the performance of the algorithm with the use of convolutional coding in hiding the signature data. Maximum likelihood decoding using Viterbi algorithm is implemented in extracting the hidden information.

In the next sections a detailed encounter of data embedding and extraction using convolutional coding will be provided.

## GENERAL EMBEDDING SYSTEM

The model we are following for data embedding is represented in Figure 1. The original host image is first transformed to the wavelet domain. Suppose that we want to embed  $N$  bits of information. The  $N$  bits of information could represent a compressed version of the signature data where vector quantization (VQ) has been implemented on the data. Now let the information sequence arranged in a vector of size  $n$  be hidden into the elements of the host. The embedding could be direct or it could depend on a secret key that is known only to copyright owner and to authorized recipient. The information symbols are added to

the host coefficients after scaling by a factor  $\mathbf{a}$  to produce a watermarked image that conceals the secret information. At the recipient side, the objective will be to extract this information with the highest possible fidelity.

In order to guarantee invisibility, the watermark sequence bits could be multiplied by a perceptual mask and not a constant scale factor  $\mathbf{a}$ . The perceptual mask can be obtained after analyzing the original image with a psychovisual model, which takes into account the different sensitivity of the human visual system (HVS) to alterations in the different elements of the host. The perceptual mask for an image is different depending on the domain chosen for embedding and on the specific properties of the HVS that are being taken into account. The details on how  $\mathbf{a}$  is evaluated in different domains can be found in (Girod 1989). We simply mention that the use of the coding and the methodology of embedding proposed here can be readily extended to other data embedding schemes. This is an important advantage of providing a general framework.

### Implementation of Convolutional Coding

Once we have built a basic scheme for reliable information embedding, its performance can be improved by means of coding. The main advantage of using convolutional codes is their error correction power together with the availability of efficient algorithms that perform soft decoding.

Given a rate  $R = k/n$  convolutional code, the  $N$  information bits are divided into groups of  $N/k$  blocks that are sequentially introduced in the convolutional encoder. The latter evolves through its state diagram and produces an output in groups of  $n$  bits, thus resulting a total of  $M = (N/n)/k$  symbols, that are transmitted through the hidden channel exactly. The values for  $k$  and  $n$  are usually small (in the order of few bits). The output bits depends not only on current set of  $k$  input bits, but also on past input bits.

### Decoding Techniques of Convolutional Coding

Regarding watermark extraction, the schematic block diagram of the decoder is shown in Figure 2. Soft and hard-decision decoding implementing the maximum likelihood (ML) decoding were used.

There are three decoding techniques available for convolutional codes: sequential decoding, Viterbi decoding, and majority logic decoding. Sequential decoding involves making trial hypotheses about the data and performs a tree search, abandoning a branch when its hypothesis disagrees too much from what is received. It is an effective technique for use with long constraint length codes. Viterbi decoding is a technique that has gained a broad application, it has the advantage that decoding complexity is deterministic. It also works

with soft decisions, and is optimum in some sense for the code. Its disadvantage is that it is limited to short constraint length codes because memory goes up exponentially with constraint length. Majority logic decoding for convolutional codes is a direct extension of the technique for block codes. They are very simple, but are generally designed to work only after hard decision. In this work, only Viterbi decoding has been implemented in the decoding of convolutional codes.

The Viterbi algorithm is a clever way of implementing maximum Likelihood decoding. It can be used for either hard or soft decision decoding. We consider hard decision decoding initially. In this case the algorithm chooses the code sequence through the trellis which has the smallest Hamming distance to the received sequence.

The exact same algorithm can be used for soft-decision decoding of convolutional codes, but in this case the minimum Euclidean distance between signals are calculated rather than Hamming distance. Although theoretically an infinite decoding delay is required to obtain optimal performance, the performance degradation is negligible provided a decoding delay of at least five times the memory of the encoder is used.

### Evaluating Error Probability Using Soft-Decision Decoding

In this section we consider the performance achieved by the Viterbi decoding algorithm. The Viterbi decoder performs either hard decision or soft decision decoding. A hard decision decoder assign either a 0 or 1 to each received bit. Soft decision decoding involves using outputs that are proportional to the log likelihood of the received bit being either 0 or 1. The log likelihood of the received bit being, for example, 0 if the received bit was 1. Soft decision decoding generally produces better performance at the cost of increasing complexity. The next section discuss the probability of error for soft decision decoding.

For soft-decision decoding we begin by determining the first event error probability. The all-zero path is assumed to be transmitted. For soft-decision decoding, the received codeword can be any real number. Because of channel fading, noise, interference, and other factors. A signal transmitted as a 0, for example, may be received as nonzero real value. From a priori knowledge of the channel, we can estimate the probabilities of the received value falling in certain regions. A similar analysis holds for the transmitting of a signal 1. In general, for AWGN channels and unquantized received signals, it can be shown that the probability of choosing a path at a distance  $d$  from the correct path is (Haykin 1994)

$$P_d = Q\left(\sqrt{\frac{2d(k/n)E_b}{N_o}}\right) \quad (1)$$

and for uncoded case the probability of error is

$$P_d = Q\left(\sqrt{\frac{E_b}{N_o}}\right) \quad (2)$$

where  $Q$  is the  $Q$ -function defined by:

$$Q(\mathbf{a}) \equiv \frac{1}{\sqrt{2\pi}} \int_{\mathbf{a}}^{\infty} e^{-b^2/2} d\mathbf{b} \quad (3)$$

and  $E_b/N_o$  is the average ratio of bit energy to noise power spectral density.

The above expression gives the first-event error probability for a path of distance  $d$  from the all-zero path.

## IMPLEMENTATION AND EXPERIMENTAL RESULTS

In order to demonstrate the effectiveness of channel coding in the performance of data embedding we have watermarked several images with different types of signature data using the above mentioned codes. We specifically make use of the Haar wavelet transform for all simulations. The implementation is as specified in Section II with a scaling parameter  $\alpha = 10$ . The method involves adding the watermark sequence to the low frequency DWT coefficients of the host image. The watermarked image is formed by taking the inverse DWT of the modified coefficients. The signature data comprises of images of size 1/4th the host image are embedded within the host image.

We perform two classes of tests. We first demonstrate the performance of the proposed method in embedding and recovering watermarks when the watermarked image undergoes distortions. The resulting watermarked signal is distorted (corrupted) using two of the most common distortions (JPEG compression and noise addition separately). The watermark is then extracted and compared with the original watermark sequence to measure robustness and extraction capability of the technique.

In the next set of tests we demonstrate the improved performance of using channel coded watermarks over uncoded watermarks. In this test, the watermark sequence is left uncoded before embedding it in the host coefficients.

The following is a more detail counterpart of the simulation work.

In this simulation the host image of *Lena* of size 256 x 256 was watermarked with a signature image of size 128 x 128 using the method described in this paper. The signature image was first compressed by a factor of 4 using vector quantization. The indices obtained were

coded using convolutional code of rate 2/3. The watermarked host image was degraded by applying JPEG compression and noise addition. To simulate the JPEG compression attack, the watermarked host image is subjected to JPEG compression with different quality factors (compression ratios). The signature data was then extracted using Viterbi algorithm. The decoded bits are then source decoded to obtain the watermark image.

For comparison, uncoded signature was similarly hidden in the transform domain of the host image *Lena* and subjected to the same JPEG compression. The corresponding results on bit error rates were plotted in Figure 3. It shows a marked improvement when the signature data was channel coded. At a quality factor of  $Q = 75\%$ , say, the bit error rate could reduce from 8% to zero.

However, it could be argued that the comparison just presented may not be fair. In the coded case, the number of bits hidden are one and a half as many as that for uncoded case. The fidelity of the watermarked host images are therefore not identical. This is true to a certain extent. Nevertheless, we like to remark first that coding of the signature data only resulted in a slight difference in the PSNR of the watermarked image for both cases. The PSNR of watermarked *Lena* in the case of the uncoded watermark was 37.35 dB, and 37.13 dB for coded watermark. The same scale factor  $\alpha = 10$  was used in both cases.

To take into consideration the fidelity factor, another set of experiments were planned as follows. The same host image *Lena*, the same signature data, and the same convolutional code as before were used to embed the signature data into the host DWT coefficients. Instead of JPEG compression, the watermarked host image is now degraded by direct noise addition in the spatial domain. The noise is Gaussian noise with zero mean and varying power (variance). Then the signature data was extracted from the distorted image and decoded using soft-threshold decoding as before. The decoded signature data is compared to the original signature and bit error rate was calculated. The result of bit error rate is plotted against PSNR of watermarked host image in Figure 4. PSNR is computed in standard manner where the noise power is equal to the variance of the added Gaussian noise. The result is compared to the uncoded case and the improvement in the bit error value with channel coding is quite obvious. For the same fidelity of the host image (e.g. PSNR = 35 dB), the bit error rate of coded signature is much lower than uncoded message (e.g. zero versus 8%). The comparison In this case is fairer than that shown in Figure 3, because the extra code bits have already been taken into account in the PSNR measure.

The bit error rate versus fidelity curves, as shown in Figure 4, are not the customarily depicted in digital communications. It is more common to show bit error rate against signal-noise-ratio. As signal in this case is the signature data, not the host image, the SNR must be defined accordingly and precisely. The signal strength in

this case is calculated by the amplitude of the scale factor  $\alpha$  that was used in the embedding process. The noise power is equal to the additive noise variance as before. The simulation results are plotted in Figure 5 for both coded and uncoded signatures. From this figure, an average of 5 dB SNR is needed for the uncoded case to obtain the same bit error rate. There is a great similarity between these curves and the widely used theoretical curves representing the bit error rate versus  $E_b/N_o$  (signal-noise-ratio) in digital communication. The theoretical curves representing bit error rate versus SNR are plotted in the same figure for comparison. The theoretical curves were calculated using Equations (1) and (2). The value of  $d$  used in Equation (1) was 4 and the code rate was  $2/3$ . The small discrepancy between the theoretical and the simulation ones is due to the distortion introduced by the wavelet transformation processes. The closeness of the theoretical and experimental results confirms in some way the correctness of the experimental procedure carried out above.

Finally, it must be added that the range of BER (0% to 50%) computed in Figures 3-5 are far too large for digital data transmission. It would be more appropriate to show BER in the range of  $10^{-3}$  to  $10^{-9}$  in log scale. Our main purpose in carrying out the experiments were to show the difference in performance of coded and uncoded cases. By necessity, the signature length cannot be too long, and we have chosen 4096 bits for our experiment. To compute BER of  $10^{-9}$  would have required a signal length of  $10^{10}$  or equivalent. This is impractical. As it is, the BER is high at the values of  $Q$  less than 75% in Figure 3 and PSNR below 30 dB in Figure 4. But at this level of attack, the host image would have been deteriorated to such an extent that it no longer has commercial value.

**SUMMARY**

In this paper we have given an outline of the advantages that channel coding using convolutional codes brings about in data embedding applications, while the

similarities with the detection problem in digital communication have also been pointed out. We have chosen the bit error probability as the reference quality measure. Analysis of the different coding schemes reveals superior performance of convolutional codes for a reasonable complexity. Comparison with uncoded case shows gains of up to 5 dB for simple codes. Note that a coding gain of 3 dB allows doubling the number of hidden information bits for the same probability of error. We also have given theoretical results (see Figure 5) that are presented in a way that becomes independent of the image to be watermarked.

It is worth noting here that the codes used here were chosen for simplicity of implementation and to obtain a qualitative comparison rather than extreme coding gains. Our best scheme can be further improved by using much longer codes.

**REFERENCES**

Barni, M., Bartolini, F., Cappellini, V., Lippi, A. and Piva, 1999. A., "A DWT-based algorithm for spatio-frequency masking of digital signaturs", Proceedings of SPIE, Security and Watermarking of Multimedia Contents, Electronic Imaging'99, San Jose, CA, Jan 25-27, vol. 3657.

Cox, I. J., Miller, M. L. and Bloom, J. A., 2000 "Watermarking Applications and Their Properties," Proceedings of the IEEE International Conference on In formation Technology, Coding and Computing, March 27-29, Las Vegas, Nevada.

Girod, B., 1989. "The Information theoretical significance of Spatial and Temporal Masking in Video Signals," Proceedings of SPIE, Human Vision processing and Digital Display, vol. 1077, pp. 178-187.

Haykin, S., 1994. "Communication systems", John Wiley and Sons.

Langelaar, G. C., Setyawan, I. and Lagendijk, R. L., 2000. "Watermarking Digital Image and Video Data," IEEE Signal Processing Magazine, September, pp. 20-46.

Stego Dos, Blocl Wolf's Picture Encoder v0.9B, <ftp://ftp.csua.berkely.edu/pub/cypherpunks/steganography/stegodos.zip>.

Tewfik, A. H. 2000 "Digital Watermarking," IEEE Signal processing Magazine, September.

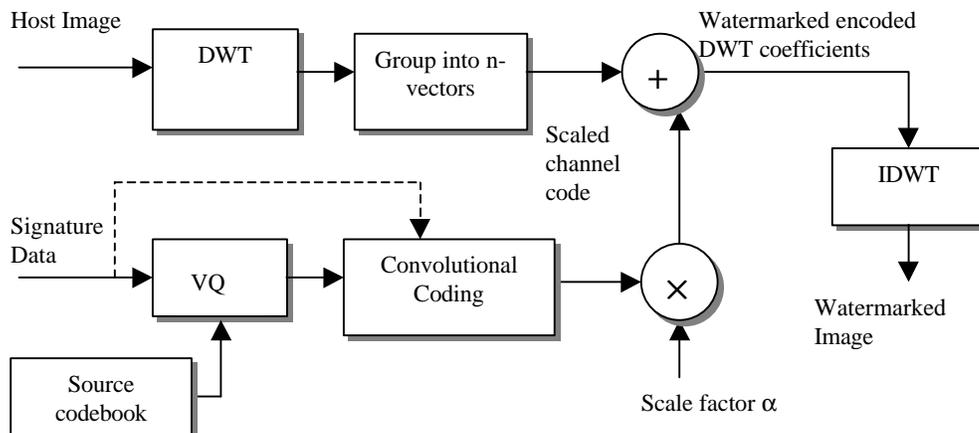


Figure 1:General Block Diagram of the Encoder.

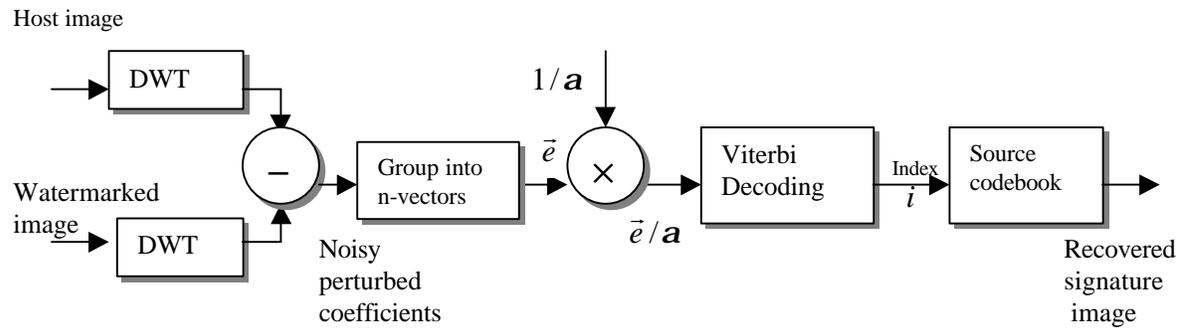


Figure 2: General Block Diagram of the Decoder.

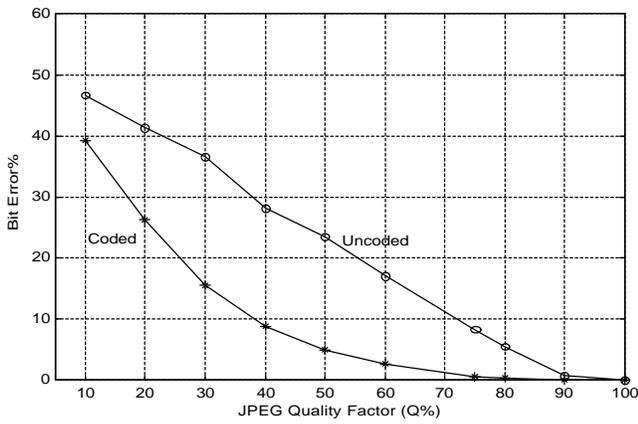


Figure 3: Bit Error Versus JPEG Compression Attack for Different Quality Factors.

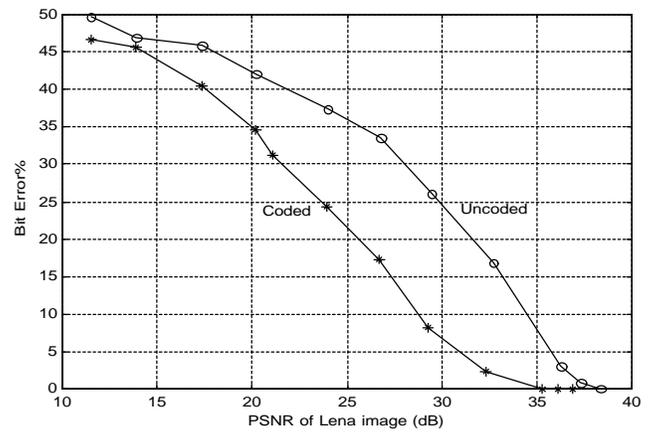


Figure 4: Bit Error Versus Noise Addition Attack for Different Quality Factors.

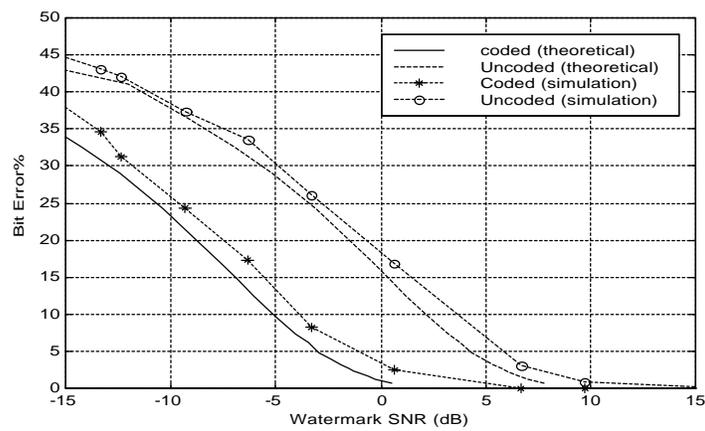


Figure 5: Bit Error Rate Versus SNR of the Watermark Signal.

# USE OF COMPUTER IMAGE WARPING TOOLS FOR SIMULATING OUTCOME OF RHINOPLASTY

Tarik Ozkul  
Computer Engineering  
American University of Sharjah,  
Sharjah, UAE  
E-mail: tozkul@aus.ac.ae

M. Haluk Ozkul, MD  
Head, ENT division  
Vakif Guraba Hospital  
Istanbul, Turkey  
E-mail: mhalukozkul@superonline.com

## KEYWORDS

Rhinoplasty, Computer simulation, Nasal tip surgery, Image warping, Rhinoplastic surgery

## ABSTRACT

Rhinoplasty is a collection of surgical procedures performed on nose for the purpose of correcting functional and shape deformities. Rhinoplasty is a surgery which requires artistic skills as well as careful surgical planning. In this paper we investigate the legitimacy of using computer image warp tools for simulating the results of rhinoplasty. The mechanical model of nose which is used for simulating the outcome of the rhinoplasty is discussed and simulation is used to predict the affect of different surgical procedures on the original patient pictures. With this we hope to find a way to aid the surgeon for planning necessary procedures during surgery and educating the patient for expected results of the rhinoplasty.

## INTRODUCTION

Rhinoplasty, which is performed for functional and aesthetic correction of nose, is the most frequently performed plastic surgery. The surgery is performed for surgical correction of nasal air passages and correcting structural defects. It is also frequently performed on perfectly functional noses to give a more aesthetic look. Whatever the reason behind the surgery, the most difficult aspect in rhinoplasty is to produce a predictable outcome which aesthetically conforms to the whole face.

Rhinoplasty is a collection of procedures performed on nose for correcting functional and aesthetical deformities. A typical rhinoplasty operation may require more than one of these procedures to be applied during the surgery. Difficulty arises from the fact that most of these surgical procedures individually affect the shape of the nose (Anderson 1969). The affect of the procedures are coupled and each one affects the shape in more than one way. Especially nasal tip positioning is approached as a different aspect of the rhinoplasty procedure because of its placement problems (Anderson 1971).

The correction of nose tip in a manner which is aesthetically pleasing is very difficult due to the fact that many coupled parameters determine the angle and the projection of the nose tip. Changing one parameter alone may adversely effect the tip position. During surgery which parameters to change and how to change them is a task that requires careful planning and artistic ability from surgeon. Objective of the procedure is to create a stable, and properly profiled tip that appears symmetric on frontal and basal views, equilateral triangular on basal view, and that flows and blends well with the rest of the face (Daniel 1992). There is no single applicable

technique for refinement of the nasal tip is available because of many anatomical variations encountered.

## PURPOSE

Motivation behind this work is as follows. Rhinoplasty is performed on structurally problematic noses as well as on structurally sound but not aesthetic looking ones. Patients who go through the rhinoplasty because of aesthetic reasons are those who are not happy with their psychological body image (Ercolani et al 1999). It is very dangerous for the patient to get wrong expectations for the surgical outcome. On the other hand it is very difficult to predict the exact outcome of rhinoplastic surgery due to anatomical differences between patients. The patient has to be given realistic expectations for the final outlook. Surgery is not a magical tool and surgeon is not a magician. It is desirable to have a tool to predict the outcome of the surgery and give an approximate idea to the patient about what to expect after the surgery. Pictorial documentation and clear understanding of surgical outcome is absolutely essential if surgeon and the patient is to avoid litigation and dissatisfaction.

Although software touch up programs exist to show the expected cosmetic changes to the patient, they are merely for photographic touch up tools and do not reflect the result of surgical procedures. A more accurate method of prediction was needed.

For simulating surgical outcome we have decided to follow a different approach. During the simulation we simulated the surgical procedure and its resultant effect on the nose. An image warp tool is used for simulation and a special warp frame is designed to accurately simulate the anatomical changes done by the rhinoplastic procedure. Using this approach, the net outcome of the simulation matches the effect of the surgical procedure almost one to one. By combining the effects of different rhinoplastic procedures on photograph of the patient, the surgeon can describe the expected outcome of the rhinoplasty quite accurately.

## ANATOMY OF THE NOSE

The nose is made up of bone and cartilage. The bony framework of the nose is made up of:

- a. Nasal bone
- b. Frontal process of the maxilla
- c. Nasal process of the frontal bone

Its cartilaginous framework is:

- a. Lower lateral (greater alar), right and left

- b. Upper lateral (lateral nasal), right and left
- c. Quadrilateral cartilage of the septum
- d. Sesamoid

The nose is attached to the forehead at the bridge of the nose. The slope of the nose starting from the bone to the tip of the nose is called dorsum. Upper part of the nose is bone, (nasal bone) and lower part is cartilage. Lower lateral cartilage has two cruras, lateral and medial. Lateral crura forms the upper side of the nostrils. Columella is supported by medial cruras of lower lateral cartilages and tip of the septum. The nostrils are located between the lip and the tip of the nose. Greater alar cartilage forms the upper side of the nostrils. Septal cartilage covers the middle portion of the nose. Lateral nasal cartilage covers the sides of the nose (Daniel et al 1988).

**SIMULATION TOOLS**

Simulation of the rhinoplasty is done on computer screen using digitized high resolution pictures of the patient. A standard IBM compatible PC with image warping software is used for simulation. Warping is an image processing process where the contours of a given object in the picture are deformed deliberately to fit contours of a different object. With this procedure a picture object can be “filled” with another picture. Warping is similar to drawing a picture on a piece of rubber. After drawing the picture one can stretch the rubber piece in any direction and distort the image.

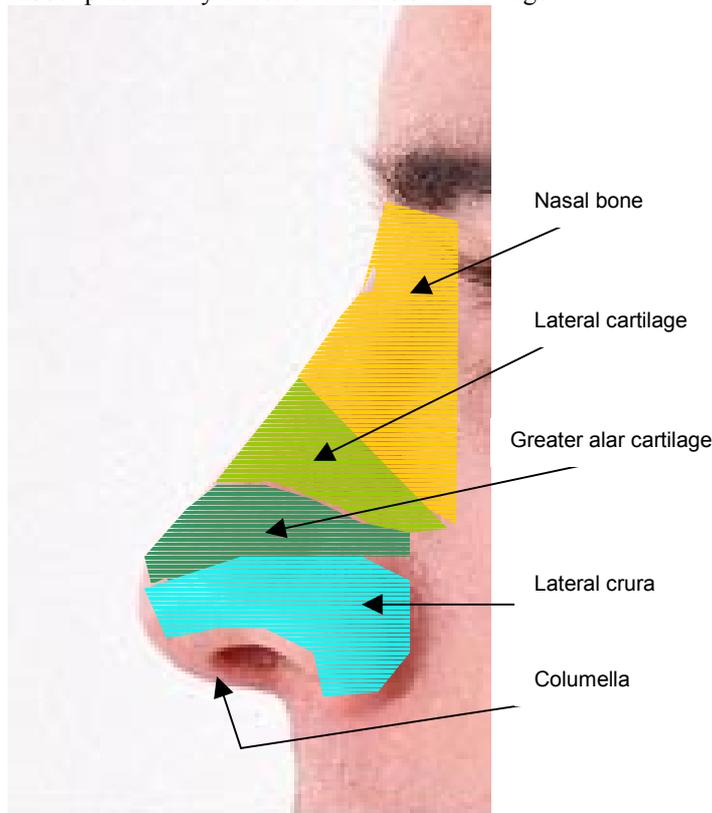


Figure 1: The anatomy of the external nose

**MAJOR RHINOPLASTY PROCEDURES**

**Removal of hump**

Picture of a patient with hump on the nose is shown in Fig. 3. Surgical correction of the nasal hump is done by filing the hump under the skin. Using endonasal or open rhinoplasty techniques, the surgeon inserts and osteotom between skin and the nasal bone and files the excessive portion of the hump. In Fig. 2. the line drawn on dorsum indicates the line above which the excess material that make up the hump needs to be removed. Material to be removed is partly bone and partly cartilage. This procedure does not affect the shape of the other parts of the nose and the outcome is quite predictable. Nevertheless it requires careful attention since grafting is quite difficult in case nasal bone is over filed (Denenberg 2002).



Figure 2. Patient with nasal hump. The line indicates the threshold above which has to be surgically removed.

*Simulation of hump removal by image warping*

Surgical removal of nasal hump is done by filing the upper portion of the nasal bone in such a way that a straight line is obtained from nasion to the point where the tip breaks. During the simulation this procedure is done on the picture by outlining the hump and then forcing the curved line indicating the hump to a straight line form.

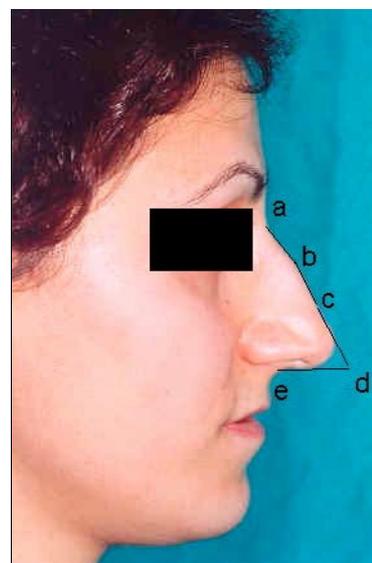


Figure 3: The triangle indicated by abc points is surgically straightened to form a straight line.

Simulation is done directly on the picture of the patient loaded on the computer screen. The Fig. 3 indicates a patient with nasal hump and important landmarks to be used during the simulation. The hump of the nose is outlined starting from nasion (point A) to the tip break point (point C). Point B in the figure indicates the tip of the hump.



Figure 4: Warp simulated nasal hump removal.

The Figure 4 indicates how the hump outline is forced to become a straight line. After selecting the warp mode of the program, the tip of the hump (point b) is forced down to be in the same line with points a and c. This will distort only part of the picture inside the inner trapezoid. This action corresponds to surgical hump removal procedure one to one. Simulation through this technique produces result that resembles the actual surgical outcome.

#### *Nose shortening and tip positioning*

Nose shortening is a complex operation of rhinoplasty where the position of the tip of the nose is changed. Changing the tip position gives the notion of changing the length of the nose. According to the tripod theory of rhinoplasty which defines the tip position of the nose, the position of the tip is defined by the two alar cartilages and the columella which is the two conjoined medial crura (Daniel 1992). Changing the length of any one of them will change the position of the tip. The tripod theory, where each cartilage is assumed to be one of the legs of a tripod is the most popular way of describing the position of the nose tip (Anderson 1971).

Tip projection refers to the posterior to anterior distance that the tip extends from the facial plane at the alar crease. Nasal tip rotation is defined as movement of the tip along a circular arc consisting of a radius centered at the nasobial angle that extends to the defining point. Lower lateral cartilages may be compared to a tripod: conjoined medial crura form one leg and the lateral crura represent the other two legs. Shortening or loss of integrity of any limb changes the spatial position of the apex (the nasal tip).

During the operation the length of the cartilages are shortened by cutting and overlapping cartilage pieces. Depending on the requirements, only the lateral cartilage or both the lateral cartilage and columella may be changed.

Simulation of the actual operation is done by simulating the length of the cartilages and changing their length with the

same constraints physically imposed on them. Tripod theory of tip rotation is simulated by forming a triangle at the exact location of lateral cartilage and the columella. The tip of the triangle is then moved toward the nasal bone simulating the shortening of lateral cartilages. The Fig. 3. the points indicated by cde indicates the position of the triangle that simulates tripod theory.



Figure 5: Tip rotation simulation.

Figure 5 indicates tip rotation simulation which is done according to tripod theory.

## CONCLUSION

The study indicated that warp tools can be used to simulate the affect of rhinoplastic procedures very effectively. Simulated pictures of patients using this approach compared to the actual pictures of patients who received rhinoplasty to see the fidelity of the procedure. No retouching or alterations performed on the images other than the simulated rhinoplastic operations. The results indicated strong resemblance to the actual surgery outputs. The intention during simulation has been to simulate the effect of rhinoplastic procedures rather than making the image look like after image of the patient. We have found that using the “procedure simulation approach” the final simulation picture can be acquired with very little effort with good fidelity. The warp simulation procedure is applied to almost all rhinoplasty procedures with success. Not all rhinoplastic procedure simulations are discussed in the paper due to the space considerations.

## REFERENCES

- Anderson, J.R. 1969. The dynamics of rhinoplasty. Proceedings of the 9th International Congress of Otorhinolaryngology. *Excerpta Medica*. Vol 206.
- Anderson, J.R. 1971. New approach to rhinoplasty. A five-year reappraisal. *Arch Otolaryngol*. Mar; 93(3): 284-91.
- Daniel, R.K. 1992. The nasal tip: anatomy and aesthetics. *Plast. Reconstr. Surg.* Feb; 89:(2): 216-24.
- Ercolani, M.; Baldaro, B.; Rossi, N.; Trombini, E.; Trombini, G. 1999. Short term outcome of rhinoplasty for medical or cosmetic indication. *J. of Physomatic Research* Vol. 47: No. 3: pp 277-281.
- Daniel, R.K.; Letourneau, A. 1988. Rhinoplasty: nasal anatomy. *Ann. Plast. Surg.* Jan: 20(1): 5-13.
- Denenberg, S. M. 2002. Case studies, Rhinoplasty tutorial, [www.facialsurgery.com](http://www.facialsurgery.com).



**SIMULATION  
IN  
TELECOM  
NETWORKS**



# CLASSIFICATION OF MPEG VIDEO CELLS IN ATM NETWORKS

Ali M. Dawood  
Etisalat College of Engineering  
Emirates Telecommunication Corporation (ETISALAT)  
Sharjah, P. O. Box: 980  
United Arab Emirates  
E-mail: dawood@ece.ac.ae

## KEYWORDS

Classification, MPEG video standard, Video over ATM networks.

## ABSTRACT

ATM networks are designed to carry different types of traffics where it provides two cell priority levels. In this paper the priority levels are extended to three levels since in MPEG video traffic more than two priorities are needed. The ATM switch was made to be able to classify the ATM cells into three categories, i.e. I, P or B picture types. These decisions are made without processing the payload of the ATM cells. Both VBR and CBR video signals were used in testing the system.

## INTRODUCTION

In a congested ATM network excessive cells are discarded because of the limited size of the buffer. To maintain a good quality of service, the source normally defines the cell loss priority. e.g. in Two-Layer coding, data of enhancement layer are given a lower priority to that of the base layer, which can then be discarded if congestion arises[1]. A wider choice is available in MPEG where B cells can be discarded prior to P and I cells [2]. In both cases, when the data is being packetized into ATM cells, the source decides which type of cells should be tagged as a high or low priority.

Cell priority tagging is straightforward in the case of the Two-Layer coding since ATM supports only two levels of priorities [3]. But in MPEG, both I and P type cells are given the high priority and B type cells are tagged as low priority. This is because I and P pictures are used in predicting B pictures, where B pictures are never been used for further predictions. However, the assigned priority tag may be altered later by the Leaky Bucket, which is used in ATM network policing [4]. If the Leaky Bucket foresees buffer overflow, high priority cells may be changed to a low priority ones regardless of whether they were of I or P type. Also in MPEG, even within the high priority I and P cells, it might be desired to discard P type cells prior to I ones; because I pictures are intra-frame coded which prevent error propagation through the frames. Therefore, an intelligent switch is needed to discriminate all I, P and B cells, irrespective of how they have been dictated by the source.

This paper presents a method of distinguishing I, P and B picture type cells of MPEG coded bit stream without processing the payload of the cells. This method adds the intelligence to the ATM switch in deciding which cells should be discarded when congestion occurs in the network. Both variable bit rate (VBR) and constant bit rate (CBR)

video signals were considered in testing the ATM cell classification system.

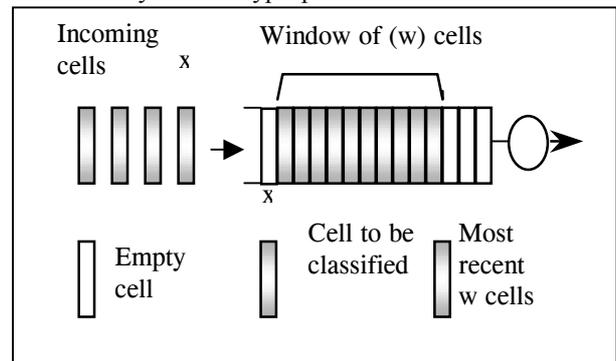
## CELL-TYPE DETECTION

MPEG video bitstream has three levels of bitrates produced by I, P and B pictures. Usually the bitrates of I pictures are higher than P pictures, and bitrates of P pictures are higher than B pictures. This feature can be employed in discriminating these picture type cells without inspecting the cell payload. In other words, the picture type can be estimated by knowing the number of cells generated from each picture in a given time. The method presented in this paper exploits this fact by introducing two thresholds values, *high* and *low* within a given window.

The window is opened in the switching buffer to contain  $w$  most recent cells, where  $w$  is the window size. As illustrated in figure 1, cell type is identified by the following steps:

- Identify the Virtual Channel Identifier (VCI) of the cell entering the window.
- Count the number of cells in the window of the same VCI with the new cell.
- Compare this number with the *low* and *high* threshold levels.
- If it is less than *low* threshold, it is of B type, above the *high* threshold it is of I type, otherwise it is of P type cell.

The accuracy of cell type prediction is then defined as the



proportion of the correctly identified cells.

Figure 1: Decision system

## SIMULATION WITH VBR VIDEO SOURCE

An image sequence was variable bitrate coded (VBR) using MPEG encoder operating at frame rate of 25 Hz with a

quantizer triplet indices of ( $q_I=3$ ,  $q_P=3$ ,  $q_B=4$ ), resulting in an average bitrate of 1 Mbit/sec. Every 44 byte data was packetized into an ATM cell and a list of inter-arrivals was produced. The list was used as a linked list to simulate 8 video sources, starting randomly at different positions in the list. Two levels of traffic smoothing were used in the simulation. The first level is *one-frame* smoothing and the second one is *three-slice* smoothing. Three-slice was chosen because its 6.6 msec period is close to one speech cell coded at 64kbits.

The window was slid cell by cell along over the multiplexed cells in the buffer. As the window size increases, possible number of cells from each VCI in the window also increase, hence, increasing the reliability of estimation. Larger window sizes also need larger *low* and *high* threshold levels. In addition to the window size itself, proper setting of threshold levels also increases the degree of prediction certainty. Table 1 shows the optimum threshold levels for various window sizes where the overall certainty of cell type prediction is maximized.

Table 1: Optimum Threshold levels for different Window Sizes

Window Size	Thresh. Low	Thresh High
20	3	4
30	4	6
40	5	8
80	9	19
160	18	38
220	25	55

Two types of window arrangements were tested. In one method (Scheme-1), the window was considered only when all the cells emanated from a particular source are of one particular picture type. Note that in MPEG video with a picture order of ...B B I B B P B B P..., during the frame transition, the window contains cells from both B picture and I or P picture for a VCI, which reduces the reliability of estimation. These transitions in the window were excluded from results in this scheme. In the second method (Scheme-2), which is more practical, the window is allowed to slide over the picture type boundaries.

### One-Frame Smoothing

Figures 2, 3 and 4 show the percentage of correct estimation of I, B, P and their aggregate for different window sizes when data for each frame were smoothed over a one-frame period. For each window size, the optimum thresholds have been chosen, where the aggregate certainty is maximized. It can be seen that in Scheme-1 the certainty increases as the window size increases. The certainty of prediction using this scheme was 99.6%, 99%, 84% and 93% for I, B, P and aggregate detection certainty respectively. However in Scheme-2, larger window sizes reduce the certainty, and it is optimum for a moderate window size of 80 cells for the sequence and bitrate used in this simulation. With this

Scheme, I, B and P were detected correctly with 95%, 95% and 82% certainty respectively and the aggregate detection certainty was 90%.

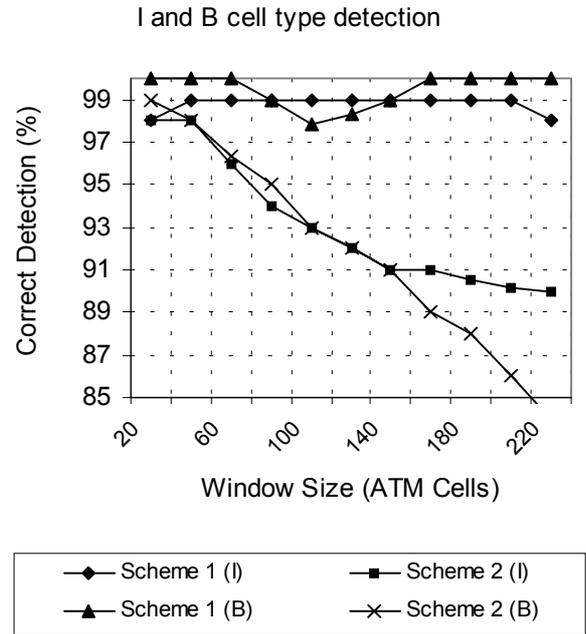


Figure 2: Correct estimation of I and B cells for different window sizes (1 frame smoothing)

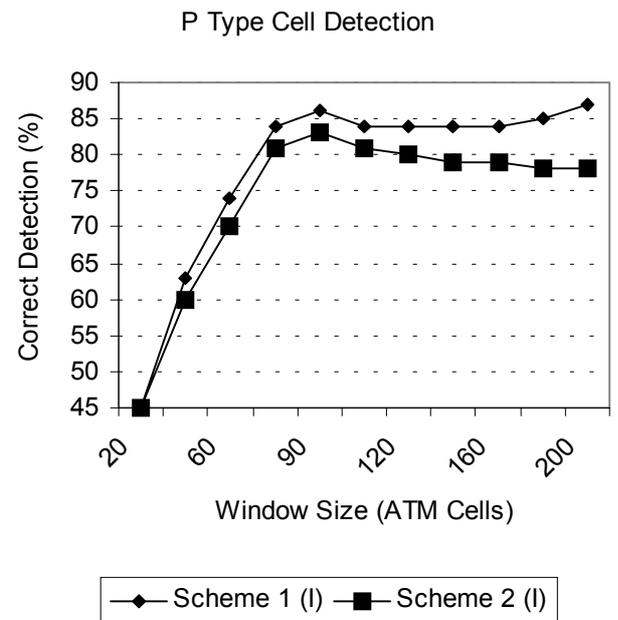


Figure 3: Correct estimation of P cells for different window sizes (1 frame smoothing)

Aggregate detection of I,P and B Type Cell Detection

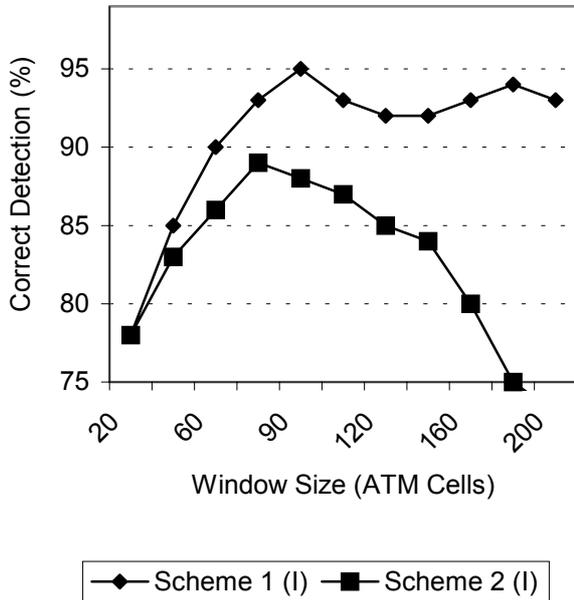


Figure 4: Aggregate correct estimation of I, P and B cells for different window sizes (1 frame smoothing)

Aggregate Detection - Different Smoothing Levels

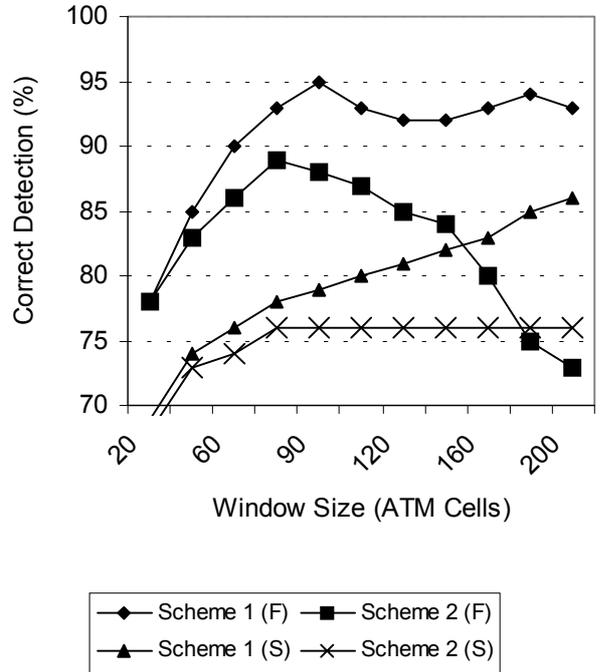


Figure 5: Aggregate estimation using 1-frame and 3 slice smoothing.

**Three-Slice Smoothing**

Figure 5 compares the aggregate percentage of correct estimation with three-slice smoothing with its counterpart using one-frame smoothing. It can be seen that at a window size of 80 cells in Scheme-1, the certainty of aggregate prediction was 78.5% and it dropped down to 75.8% for Scheme-2. These are lower than the one-frame smoothing predictions which are 93% and 90%, respectively. It can be seen from figure 5 that the optimum window size is 80 cells, which is the same as the one-frame smoothing. Since certainty curve of Scheme-2 does not improve beyond window size of 80 cells, and minimum size windows are preferred due to their reduced processing delay, the window size of 80 cells was selected as an optimum one.

**SIMULATION WITH CBR VIDEO SOURCE**

In CBR video coding mode the bit rate is made constant by allowing the quantizer triplet indices ( $q_I$ ,  $q_P$ , and  $q_B$ ) to vary according to the nature of the recorded video scene. In this case the traffic of individual I, P and B traffic is almost constant. Simulation results show that the performance of the classification system on CBR video signals is much better than the one reported for the VBR signals. Figure 6 illustrates the difference between the performance of the classification system when VBR and CBR video sources are used.

Aggregate Detection: VBR and CBR

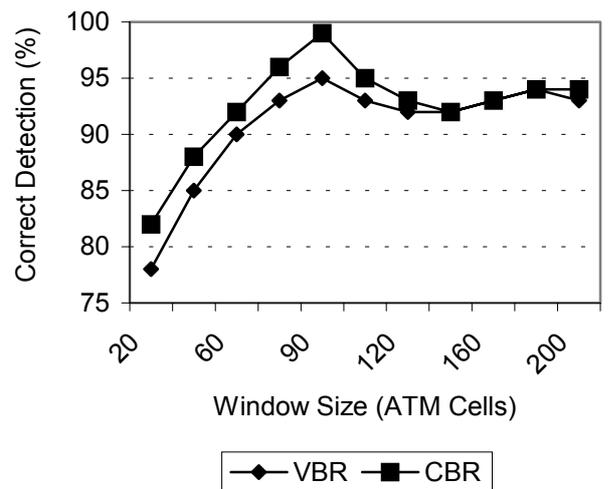


Figure 6: Aggregate estimation for VBR and CBR video.

## DISCUSSION AND CONCLUSION

### Discussion

In MPEG video, the bitrate variation of P and B pictures are highly dependent on the scene activities. The results obtained from the one-frame smoothing show that during a high scene activity some B and P type cells may be detected as I type and adversely, in a low activity scene, P cells may be detected as B. Although these reduce the level of certainty, they have little effect on the intelligence of the switch. This is because high active B and P type cells carry significant video

information, which are better not to be discarded, and also discarding low active P cells does not impair the picture quality significantly.

When three-slice smoothing is used the level of fluctuation in the I, B and P pictures is increased. This will reduce the level of certainty, as for a three-slice period some parts of B pictures may generate a high data rate, and conversely, some parts of I pictures may not have texture and hence the bitrate is reduced. As figure 5 shows the level of certainty compared to one-frame smoothing is reduced by 15%.

Finally, the classification system performed very well when constant bitrate (CBR) video was used, this was due to the improved degree of certainty available in the rate controlled traffic.

### Conclusion

A method of identifying the type of ATM cells carrying MPEG coded video without processing the content of cells has been presented. For VBR video, it is possible to predict the cell type with more than 90% confidence using a one-frame bitrate smoothing. However, the efficiency of the prediction is better than the 90% figure of the simulation, because some important P and B cells are desirable to be detected as high priority cells. Bitrate smoothing of three-slice intervals showed lower performance than one frame smoothing. Therefore, one-frame smoothing interval, or larger, is required to maintain a high prediction certainty. For CBR video, due to the rate control, the degree of certainty is much better than the VBR video with one-frame smoothing. The prediction confidence was close to 98%.

## REFERENCES

- [1] M. Ghanbari, "Two-layer coding of video signals for VBR networks", IEEE Journal of selected areas of communications, vol. 7, pp. 771-781, June 1989.
- [2] ISO/IEC, "Generic coding and moving pictures and associated audio", Recommendation H.262, March 1994.
- [3] J. Pittes and J. Schormans, "Introduction to ATM design and performance", John Wiley and sons Ltd., 1996.

- [4] J. Bae and T. Suda, "Survey of traffic control scheme and protocols in ATM networks", In proc. Of IEEE, vol. 79, pp. 170-189, February 1991.

## ACKNOWLEDGMENT

The author would like to acknowledge the financial support from Etisalat College of Engineering, a division of Emirates Telecommunications Corporation - ETISALAT, UAE.

## BIOGRAPHY

Ali M Dawood was born in Sharjah, United Arab Emirates (UAE), in 1971. He received the BTEC diploma and the B.Eng degree in Communications in 1991 and 1994, respectively, from Etisalat College of Engineering, a division of Emirates Telecommunications Corporation - ETISALAT, UAE. The M.Sc. and Ph.D. were obtained in 1995 and 1999 respectively from the University of Essex, Colchester, UK. He is currently working as an Assistant Professor in Etisalat College of Engineering. His research interests include MPEG video traffic source modelling, networks performance evaluation and high-level video processing for Multimedia applications.

# **SIMULATION IN ENGINEERING**



# SYSTEM DYNAMIC SIMULATING MODELLING OF DRIVING SYSTEM “ANCHOR WINDLASS DRIVEN BY ASYNCHRONOUS MOTOR” (BSVPAM)

Ante Munitić, Mario Oršulić, Maja Krčum, Joško Dvornik  
College of Maritime Studies  
University of Split  
Zrinsko-frankopanska 38,  
21000 Split,  
Croatia  
e-mail: [munitic@pfst.hr](mailto:munitic@pfst.hr)

## KEYWORDS

System Dynamics, Modelling, Asynchronous Engine, Windlass, Continuous and Discrete Simulation

## ABSTRACT

System dynamic simulating modelling is one of the most appropriate and successful scientific dynamics modelling methods of the complex, non-linear i.e. natural, technical and organisational systems. Investigation of behaviour dynamics of the ship's propulsion system as a typical example of complex, dynamic technical systems requires application of the most efficient modelling methods. The aim of this essay is to present the efficiency of application of the system-dynamic simulating modelling in investigation of behaviour dynamics of the BSVPAM propulsion system. The anchor windlass and its driving asynchronous motor shall be presented by mental - verbal, structural and mathematical computing models. The System Dynamics Models are, in essence, continuous models because the realities are presented by the set of non-linear differential equations, i.e. "equations of state". They are at the same time discrete models, because they used basic time step for counting i.e. discrete sampling DT, which value is determined in total accordance with "SAMPLING THEOREM" (Shannon and Kotelnikov). With the choice of basic time step DT it is possible to do computer modelling of continuous simulation models on digital computer, which is very suitable for education of the marine students and engineers, because they can study complex dynamics behaviour of marine systems and process.

## 1. INTRODUCTION

The System Dynamics Modelling is in essence special, i.e. "holistic" approach to the simulation of the dynamics behaviour of natural, technical and organisation systems, and it contains quantitative and qualitative Simulation Modelling of various natured realities. The concept of optimisation in System Dynamics is based on belief that the manual and iterative procedure, i.e. optimisation by the method "retry and error" can be successfully executed using heuristic optimisation algorithm, with the help of digital computer, and in complete coordination with System Dynamics Simulation Methodology. This simulation model BSVPAM is small part of scientifically macro project called: Intelligent Computer Simulation of the Model of Marine Processes.

## 2. SYSTEM- DYNAMIC SIMULATING SYSTEM MODELS “ANCHOR WINDLASS DRIVEN BY ASYNCHRONOUS MOTOR”

### 2.1. System dynamic model of the anchor windlass

Anchoring is an operation by which the ship is fixed to the ship's bottom. This is performed by the anchor arrangement consisting of: anchor, anchor chains, stoppers, chain locker and windlass. Some elements of the anchor arrangement are also used for the mooring of a ship. All ships are provided with the bow anchor arrangement and some of them with the stern anchor arrangement. Ships of less size may be provided with anchor windlasses driven by the internal combustion engines while on tankers where such drive may cause explosion windlasses are driven by steam.

Windlasses on other ships are mostly driven by electric motors and recently are also hydraulically driven. Electric drive of windlasses shall be performed by: - the alternative three phase electric motor, in Leonard's connection, alternative three phase electric motor directly coupled with overlapping of two or three pairs of pole.

Anchor windlass consists of driving electric motor with stopper, reduction gear and main shaft unit laid in solid bearings. Reduction gear where safety-sliding coupling is located includes a few pairs of front gears with associated shafts and bearings. High-speed rotating shafts are laid in rolling bearings while the main shaft is fitted in sliding bearings. Chain locker is situated in the main shaft having the belt brake and tightening drum. Claw clutch is located between chain locker and reduction gear enabling the independent operation of the tightening drum from the chain locker.

Reduction gear is lubricated by oil sump while main shaft and other sliding surfaces are grease lubricated. The basic equations of anchor windlass are:

$$\frac{ds}{dt} = \omega r_0 \quad (1)$$

$$F_t = (G_1 S + G_s) g \quad (2)$$

$$F_{uz} = \varphi g A S \quad (3)$$

$$F_u = F_t - F_{uz} \quad (4)$$

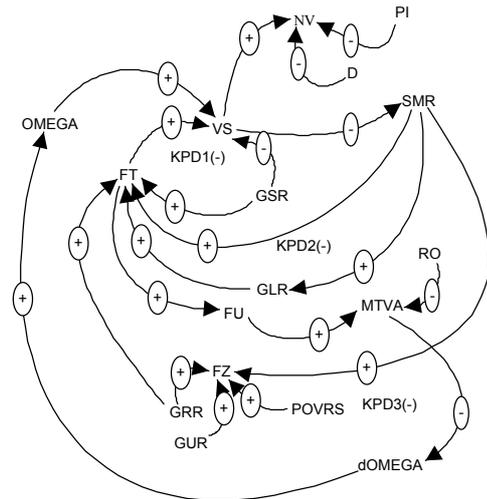
$$M_v = F_u r_0 \quad (5)$$

$$S = S_0 - \int v_s dt \quad (6)$$

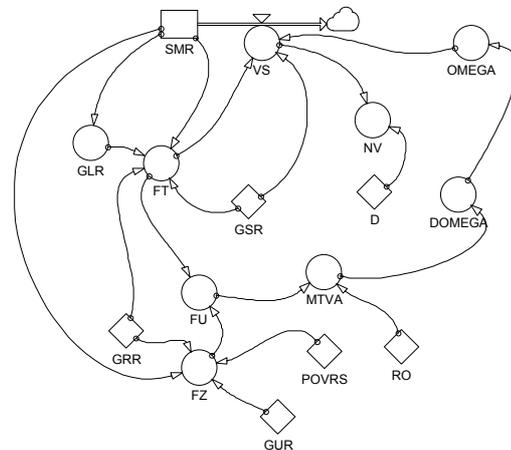
According to the basic equations the mental verbal model of ship anchor windlass may be developed or structural and anchor windlass course diagrams, respectively.

Table 1: Marks and mode of records in Dynamo language

Marks	Description	Dynamo language
M <sub>v</sub>	Winch torque	MTVA
G <sub>s</sub>	Anchor weight	GS
G <sub>1</sub>	Chain weight	GL
S	Chain length	SM
φ	Seawater density	GU
G	Gravity	GR
A	Anchor and chain area	POVRSR
S	Chain length	S
V <sub>s</sub>	Anchor lifting speed	VS
F <sub>T</sub>	Loading force	FU
F <sub>uz</sub>	Buoyancy force	FZ



Figures 1: Structural simulation model of the anchor windlass



Figures 2: Structural simulation model in DYNAMO symbolic-flow diagram

Three feedback loops are present in the concerned anchor arrangement system (KPD). KPD1 (-): VS=> (-) SMR=>(+)FT=>(+)VS; which has self-regulating dynamic character (-) because the sum of negative signs is odd number.

KPD2(-): VS=>(-)SMR=>(+)GLR=>(+)FT=>(+)VS; which also has self-regulating dynamic character .

KPD3(-)SMR=>(+)GLR=>(+)FT=>(+)FU=>(+)MTVA=>(+)dOMEGA=>(+)OMEGA=>(+)VS; which has also self-regulating dynamic character.

Within KPD a few cause and effect relations are acting (UPV) for which the following dynamic relations are valid:

“ If the anchor and chain VS lifting speed is increasing, chain length SMR is reducing what results in the negative sign of cause-effect relation”; by increasing chain and anchor VS lifting speed, the number of revolutions of

shaft NV is also increased resulting in positive sign of UPV.”

By increasing the relative anchor mass GSR, chain and anchor VS lifting speed is reduced resulting in negative sign of UPV”. “By increasing of relative chain length SMR loading force is increasing as well as the total chain weight GL and also buoyancy force FU resulting in positive sign UPV.”

“By increasing the loading force –FT as well as speed of rotation of asynchronous motor, the anchor-VS and chain lifting speed is increased resulting in positive sign of the observed UPV.” “By increasing gravity-G the loading force-FT and buoyancy force-FU are increased and accordingly observed UPV has positive sign.” “By increasing loading force-FT total force is increased and consequently by increasing total force winch torque-MTVA is increased and thus the observed UPV has a positive sign.”

“By increasing buoyancy force-FU total force is decreasing resulting in a negative sign UPV.” “By increasing chain and anchor-A area as well as seawater density buoyancy force is increased resulting in positive sign of the observed UPV.”

## 2.2. System dynamic model of asynchronous motor

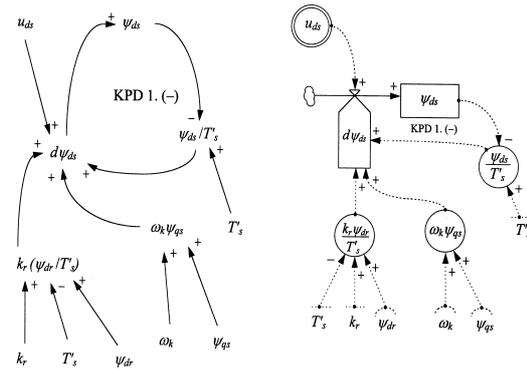
The following cause and effect relation is applicable to the first equation:

$$\frac{d\psi_{ds}}{dt} = u_{ds} - \frac{1}{T'_s} \psi_{ds} + \frac{k_r}{T'_s} \psi_{dr} + \omega_k \psi_{qs} \quad (7)$$

Variation speed of system - d condition is decreasing what results in negative sign of the observed cause and effect relation.” “By increasing rotor linkage factor –Kr, system condition variation is also increased resulting in positive sign of the observed UPV.” “By the increase of stator - Ts transient time constant, system condition variation is reduced resulting in negative sign of the observed UPV.” “If product is increasing and if stator voltage variation in axis d - Uds is increasing then system condition variation speed is also increasing resulting in a positive sign of the observed UPV.”

On the basis of the specified model given in the form of cause and effect relation of system elements, the mental verbal model of equation of asynchronous motor condition may be determined and thus a structural model and continuity diagram of the mentioned equation may be elaborated.

In this short paper, it is impossible to give a complete model (27 equations) of the asynchronous motor, complete model has been



Figures 3: Structural diagram and continuity diagram of the first differential equation of the asynchronous motor condition

presented in IASTED 1998., Pittsburg, USA.

## 2.3. Computer simulating model BSV PAM

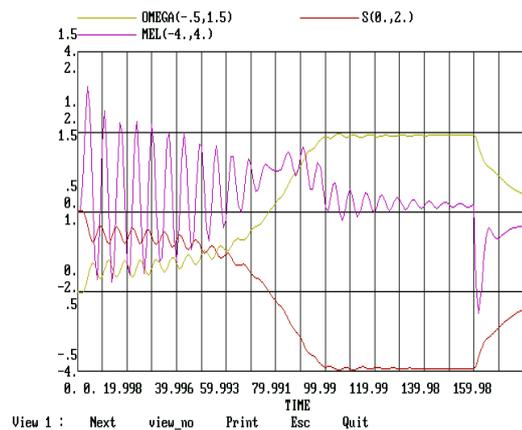
PARAMETERS OF SHIP'S ANCHOR WINDLASS:	
C G=7000	TOTAL NOMINAL WEIGHT OF ANCHOR WITH CHAIN (kg)
C GS=3000	NOMINAL ANCHOR WEIGHT (kg)
K GSR=GS/G	RELATIVE VALUE OF ANCHOR WEIGHT
K GL=4000/SM	NOMINAL CHAIN WEIGHT PER LENGTH UNIT (kg/m)
K GLR=GL/G	RELATIVE VALUE OF CHAIN WEIGHT PER LENGTH UNIT
C SM=100	CHAIN LENGTH (m)
C GR=9.81	GRAVITY (M/s*s)
K GRR=GR/SM	RELATIVE GRAVITY
C GU=1.25	SEAWATER DENSITY (kg/m*m*m)
K GUR=GU/SM	RELATIVE SEAWATER DENSITY
C VOLUM=1000	CHAIN AND ANCHOR VOLUME ( m*m*m)
K POVRSR=VOLUM/SM	CHAIN AND ANCHOR AREA
R VS.KL=CLIP(STEP(OMEGA.K,0),0,FT.K, GSR)	ANCHOR LIFTING SPEED
C D=1	
A NV.K=VS.KL/(D*3.14)	SHAFT NUMBE OF ROTATION
L SMR.K=SMR.J+DT*(-VS.JK)	
N SMR=100	
A MTVA.K=FU.K*RO	WINDLASS TORQUE
C RO=0.1	CHAIN LOCKER DIAMETER
A FT.K=CLIP(STEP((GLR*SMR.K+GSR)*GR,0),0,GLR*SMR.K*GR,GSR)	LOADING FORCE
A FZ.K=CLIP(STEP(GUR*GRR*POVRSR*SMR.K,0),0,SMR.K,15)	BUOYANCY FORCE
A FU.K=FT.K-FZ.K	TOTAL FORCE
SAVE SMR,MTVA,FU,FT,FZ,VS	
PARAMETERS OF ASYNCHRONOUS MOTOR:	
C Rs=0.0141	STATOR TRANSFORMED OPERATING RESISTANCE
C Rr=0.0934	ROTOR TRANSFORMED OPERATING RESISTANCE + 5Rr
C Lcs= 0.286	STATOR TRANSFORMED INDUCTANCE
C lcr= 0.1	ROTOR TRANSFORMED INDUCTANCE
C Lm=3.32	TRANSFORMED MUTUAL INDUCTANCE
C TCS=20.3	STATOR TRANSIENT TIME CONSTANT
C Tcr=3.11	ROTOR WITH 5R TRANSIENT TIME CONSTANT
C Ks=0,965	STATOR LINKAGE FACTOR

$C_{Kr}=0.95$  ROTOR LINKAGE FACTOR  
 $C_{Lsigs}=0.12$  STATOR LEAKAGE INDUCTANCE  
 $C_{Lsigr}=0.175$  ROTOR LEAKAGE INDUCTANCE  
 $C_{SIGMA}=0.083$  LEAKAGE FACTOR ( $SIGMA=1-K_s*K_r$ )  
 $CH=57.6$  INERTIA CONSTANT  
 I DIFFERENTIAL EQUATION OF CONDITION:  
 $R_{dPSIds.KL}=U_{ds.K}-(PSIds.K/Tcs)+$   
 $OMEGA.K*PSIqs.K+(Kr*PSIdr.K)/Tcs$   
 $DPSIds=$  VARIATION SPEED OF LINKAGE FLUX  
 $PSIds$  (Wb/s)  
 $U_{ds}=$  STATOR VOLTAGE IN AXIS d (V)  
 $Tcs=$  STATOR TRANSIENT TIME CONSTANT  
 $OMEGA.K=$  OMP STATOR ROTATION  
 $OMEGA.K$  SYNCHRONOUS SPEED (rad/s)  
 $PSIqs=$  STATOR LINKAGE MAGNETIC FLUX IN  
 AXIS q (Wb/s)  
 $Kr=$  ROTOR LINKAGE FACTOR  
 $PSIdr=$  ROTOR LINKAGE MAGNETIC FLUX IN  
 AXIS d (Wb)  
 $L_{PSIds.K}=PSIds.J+DT*(dPSIds.JK)$   
 $N_{PSIds}=0$   
 $PSIds=$  STATOR LINKAGE FLUX IN AXIS d (Wb)  
 $DPSIds=$  VARIATION SPEED OF STATOR  
 LINKAGE FLUX IN AXIS d (Wb/s)  
 $A_{Uds.K}=STEP(1,0)+$   
 $CLIP(1,0,FT.K,GSR+1e-20)+STEP(-1,0)$   
 $A_{OMEGA.K.K}=1$   
 $OMEGA.K=$  OMP STATOR ROTATION  
 $OMEGA.K$  SYNCHRONOUS SPEED (rad/s)  
 II DIFFERENTIAL EQUATION OF CONDITION:  
 $R_{dPSIqs.KL}=U_{qs.K}-(PSIqs.K/Tcs)-$   
 $OMEGA.K*PSIds.K+(Kr*PSIqr.K)/Tcs$   
 $DPSIqs=$  VARIATION SPEED OF STATOR  
 LINKAGE MAGNETIC FLUX IN AXIS q (Wb/s)  
 $U_{qs}=$  STATOR VOLTAGE IN AXIS q (V)  
 $PSIqs=$  STATOR LINKAGE MAGNETIC FLUX IN  
 AXIS q (Wb/s)  
 $Tcs=$  STATOR TRANSIENT TIME CONSTANT  
 $OMEGA.K=$  OMP STATOR ROTATION  
 $OMEGA.K$  SYNCHRONOUS SPEED (rad/s)  
 $PSIds=$  STATOR LINKAGE FLUX IN AXIS d (Wb)  
 $Kr=$  ROTOR LINKAGE FACTOR  
 $PSIqr=$  ROTOR LINKAGE MAGNETIC  
 FLUX IN AXIS q (Wb)  
 $L_{PSIqs}=0$   
 $PSIqs=$  STATOR LINKAGE MAGNETIC FLUX IN  
 AXIS q (Wb/s)  
 $DPSIqs=$  VARIABLE VARIATION SPEED  $PSIqs$   
 (Wb/s)  
 $A_{Uqs.K}=0$   
 $U_{sq}=$  STATOR VOLTAGE IN AXIS q (V)  
 $A_{Uas.K}=SQRT(U_{ds.K}*U_{ds.K}+U_{qs.K}*U_{qs.K})$   
 $U_{as}=$  VECTOR SUM OF VOLTAGE  
 COMPONENTS IN AXES q AND d  
 III DIFFERENTIAL EQUATION OF CONDITION:  
 $R_{DpsiDR.kl}=U_{dr.K}-(PSIdr.K/Tcr)+$   
 $(OMEGA.K-OMEGA.K)*PSIqr.K+K_s*PSIds.K/Tcr$   
 $A_{Udr.K}=0$   
 $DPSIdr=$  ROTOR VARIATION SPEED OF  
 LINKAGE MAGNETIC FLUX IN AXIS d (Wb/s)  
 $Tcr=$  ROTOR SA 5R TRANSIENT TIME  
 CONSTANT  
 $K_s=$  STATOR LINKAGE FACTOR  
 $OMAGAK=$  OMP STATOR SYNCHRONOUS  
 $OMAGAK$  ROTATION SPEED (rad/s)  
 $L_{PSIdr.K}=PSIdr.J+DT*(dPSIdr.JK)$   
 $PSIdr=$  ROTOR LINKAGE MAGNETIC FLUX IN  
 AXIS d (Wb)  
 $N_{PSIdr}=0$   
 IV DIFFERENTIAL EQUATION OF CONDITION:  
 $R_{dPSIqr.KL}=U_{qr.K}-(PSIqr.K/Tcr)-$   
 $(OMEGA.K-OMEGA.K)*PSIdr.K+K_s*PSIas.K/Tcr$   
 $A_{Uqr.K}=0$   
 $DPSIqr=$  ROTOR VARIATION SPEED OF  
 LINKAGE MAGNETIC FLUX IN AXIS q (Wb/s)

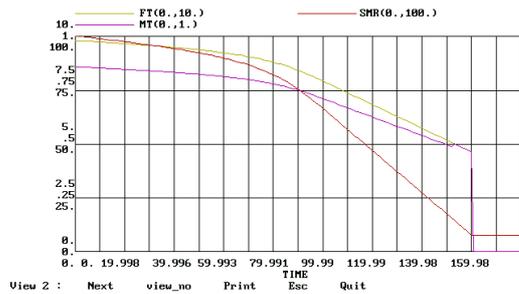
$Tcr=$  ROTOR SA 5R TRANSIENT TIME CONSTANT  
 $K_s=$  STATOR LINKAGE FACTOR  
 $OMEGA.K=$  OMP STATOR SYNCHRONOUS  
 $OMEGA.K$  ROTATION SPEED (rad/s)  
 $LPSIqr.K=PSIqr.J+DT*(dPSIqr.JK)$   
 $PSIqr=$  ROTOR LINKAGE MAGNETIC FLUX IN  
 AXIS q (Wb)  
 $NPSIqr=0$   
 V DIFFERENTIAL EQUATION OF CONDITION:  
 $RdOMEGA.KL=(1/(2*h))*(Ks/Lcr)*(PSIqs.K*PSIdr.K-$   
 $PSIds.K*PSIqr.K)-(1/(2*H))*MT.K$   
 $DOMEGA=$  VARIATION SPEED OF ANGLE SPEED  
 $DOMEGA$  (rad/s(s))  
 $H=$  INERTIA CONSTANT  
 $K_s=$  STATOR LINKAGE FACTOR  
 $Lcr=$  ROTOR TRANSFORMED INDUCTANCE  
 $L_{OMEGA.K}=OMEGA.J+DT*(dOMEGA.JK)$   
 $OMEGA=$  ANGLE SPEED (rad/s)  
 $N_{OMEGA}=0$   
 VI EQUATION OF ELECTROMAGNETIC TORQUE:  
 $A_{Mel.K}=PSIds.K*Iqs.K-PSIqs.K*Ids.K$   
 VII EQUATION OF LOADING TORQUE:  
 $A_{MT.K}=STEP(MTVA.K*KOPT,0)$   
 $C_{KOPT}=1$   
 VIII ADDITIONAL CURRENTS EQUATIONS:  
 $A_{IDS.K}=(1/Lcs)*(PSIds.K-Kr*PSIdr.K)$   
 $A_{Iqs.K}=(1/Lcs)*(PSIqs.K-Kr*PSIqr.K)$   
 $A_{Iqs.K}=(1/Lcs)*(PSIqs.K-Kr*PSIqr.K)$   
 $A_{Ias.K}=SQRT(Ids.K*Ids.K+Iqs.K*Iqs.K)$   
 $A_{ldr.K}=(1/Lcr)*(PSIdr.K-Ks*PSIds.K)$   
 $A_{lqr.K}=(1/Lcr)*(PSIqr.K-Ks*PSIqs.K)$   
 $A_{lar.K}=SQRT(ldr.K*ldr.K+lqr.K*lqr.K)$   
 IX ADDITIONAL SLIP AND NUMBER OF  
 ASYNCHRONOUS MOTOR REVOLUTION  
 EQUATIONS:  
 $S_{S:K}=(OMEGA.K-OMEGA.K)/OMEGA.K$   
 $N_{S}=1$   
 SAVE  $dPSIds,PSIds$   
 SAVE  $dPSIqs,PSIqs$   
 SAVE  $dPSIdr,PSIdr$   
 SAVE  $dPSIqr,PSIqr$   
 SAVE  $dOMEGA,OMEGA,OMEGA.K$   
 SAVE  $U_{ds},U_{qs},U_{as},I_{ds},I_{qs},I_{as}$   
 SAVE  $I_{dr},I_{qr},I_{ar}$   
 SAVE  $Mel,MT,S$   
 $SPEC\ DT=.01,LENGTH=180,SAVPER=1$

## 2.4. The Results of Simulation

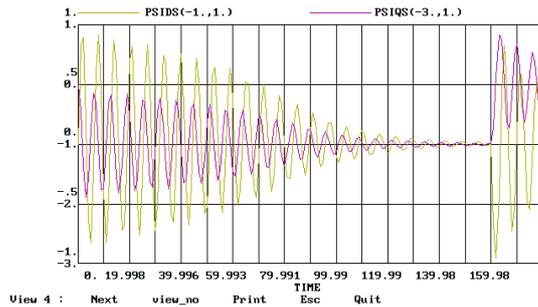
Graphical figure of the simulation results of the BSVMP:



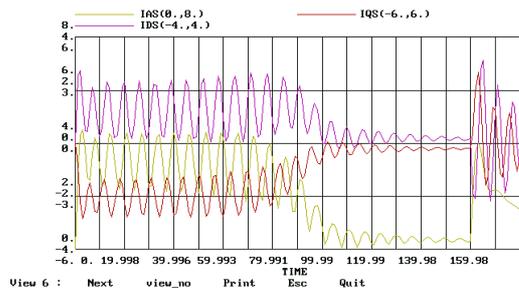
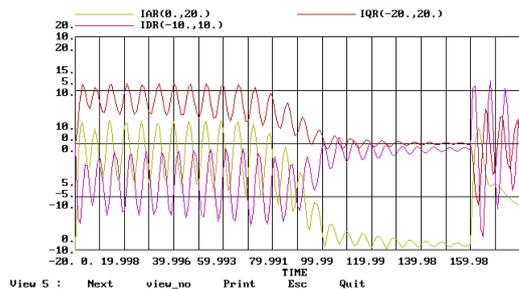
Figures 4: Diagram of loading torque, electric torque and slipping



Figures 5: Diagram of loading force, buoyancy, chain length and speed



Figures 6: Stator magnetic fluxes



Figures 7: Stator and rotor currents

### 3. CONCLUSION

System Dynamics is such scientific methodology that provides the simulation of the most complex systems. In the shown example the methodology evidently indicates to the high quality of the simulations of the complex dynamical systems and it gives the opportunity to every student or engineer interested to by the same methodology modulates, optimises and simulates any scenario of the existing realities. Furthermore,

the users which use this simulation methodology of the continuous models on a digital computer, create a possibility to themselves of the newest knowledge's in the behaviour of the dynamical systems. The Methodology is also significant because it doesn't contain only a computer type of modelling, but it clearly determinates the metal, structural and mathematical modelling of the same system realities. Based on our long-term experience in the application of the dynamical methodology of simulating and in this short presentation we provide every expert in need with the possibility to acquire additional knowledge about the same system in a quick scientifically based way of exploring the complex systems.

### 4. REFERENCES

- Forrester, Jay W. 1973/1971. "Principles of Systems", MIT Press, Cambridge Massachusetts, USA.
- Jadric, M. and B. Francic. 1996. "Dinamikaelektričnih strojeva."(in Croatian), Manualia Universitatis Studiorum Spalatiensis, Graphics, Zagreb, Croatia.
- Munitic, A.; L. Milic and M. Milikovic. 1997. "System Dynamics computer Simulation Model of the Marine Diesel – Drive Generation Set 1997. Automatic Control System." IMACS World Congress on Scientific Computation, Modelling and Applied Matematichs, vol.5, Wiissenschaft & Technik Verlag, Berlin.
- Munitic, A.; I. Kuzmunic and M. Krčum. 1998. "System Dynamic Simulation Modelling of the Marine Synchronous Generator", IASTED, Pittsburg. p.p.372.-375
- Munitic, A. 1989. "Application Possibilities of System Dynamics Modelling." System Dynamics, Edited by Susan Spencer and George Richardson, Proceedings of the SCS Western Multiconference, SanDiego, California, A Society for Computer Simulation International, San Diego, USA.
- Munitić A.; L. Milić and M. Bupić. Oct.18-20, 2001. "System Dynamics Simulation Modeling and Heuristic Optimization of The Induction Motor", Simulation Symposium, Marseille.
- Richardson, George P. and Pugh III Aleksander L. 1981. "Introduction to System Dymanics Modelling with Dynamo", MIT Press, Cambridge, Massachusetts, USA.

# MODELING AND SIMULATION OF GAS TURBINE ENGINE PERFORMANCE

M. Montazeri-Gh  
S.Mojallal-A  
System Simulation and Control Laboratory,  
Iran University Of Science and Technology ,Tehran, Iran  
Montazeri@iust.ac.ir

## KEYWORDS

Modeling, Engine performance, Gas turbine, Simulation

## ABSTRACT

Modeling and simulation of gas turbine engine plays a key role in preliminary stages of engine design. In this paper modeling and simulation of gas turbine engine performance are presented in steady state and transient modes. Compressor characteristics that are used in this modeling obtained from two different methods including generalized and designed methods. Variation of specific heat ratio ( $\gamma$ ) and compressor isentropic efficiency ( $\eta_c$ ) has also been considered. Finally, steady state and transient performance of a turbojet engine have been achieved by numerical simulation using Simulink ToolBox and the results are checked by experimental results. It was seen that two methods to obtain compressor characteristics would lead to the same results for prediction of steady state performance and acceleration time. However, the variation of gamma and efficiency has considerable effects on the results especially at low speeds.

## INTRODUCTION

Efficiency and reliability of gas turbine engines have been considerably improved over the past two decades. This fact caused rapid improvement of mathematical models of off-Design performance, numerical simulation and dynamic behavior of gas turbines (Crosa et al. 1998). Gas turbine engines are often required to rapidly follow changes in power demand. The prediction of their dynamic behavior is very important both to help the design of plant components, control systems and also performance optimization. Since gas turbine engines play a key role in modern airplane fly and one of the most available dynamic analysis methods are mathematical models, the importance of the subject reveal more and more (Doug Garrard 1996). These models are powerful tools in research, development, mounting, and optimization (Visser 2000).

Nowadays control techniques using computer simulations have demonstrated the possibility of improving the engine performance (Lombardo 1996). The development of a suitable control system requires a deep understanding of the

steady state and transient behavior of the gas turbine engine (Cohen et al. 1996).

The essential role of aircraft engine control systems is to supply fuel as prime energy in accordance with demand and engine operation. Several dynamic simulation models are described in literatures some of them in the frequency domain and others in time domain (Yamane et al. 1995). However, in pervious studies the values of gamma and compressor efficiency have been taken constant (Ferrari et al. 1995).

This paper presents the mathematical modeling for gas turbine engine in both steady state and transient modes. In this paper, the variation of gamma and compressor isentropic efficiency and their effects on engine performance have been considered. In addition, compressor characteristics, used in this modeling, have been extracted from two different methods; generalized and designed methods. Finally, the steady state and transient performance for a turbojet engine have been simulated using simulink, and the results are compared with the test data.

## MATHEMATICAL MODELING

The heart of a gas turbine engine is gas generator. A schematic diagram of a gas generator is shown in fig(1). The compressor, combustion chamber and turbine are the major components of a gas generator, that is common to the turbojet, turbofan, turboprop and turboshaft engines. The purpose of a gas generator is to supply high pressure air at compressor outlet and high temperature gas at combustion exit. As mentioned before identification of controllable parameters and mathematical modeling is the first step to design an engine control system. Mathematical model can be obtained by using dynamic and thermodynamic governing equations for engine components.

### Engine Steady State Mode

Steady state modeling starts from design point, as the engine parameters are known at this point. In steady state mode, the turbine power is equal to compressor power. Also, the turbine

flow is equal to summation of compressor flow and fuel mass flow. The calculation of off-Design performance specifications requires component characteristics. The variation of pressure ratio and compressor efficiency with respect to air mass flow are obtained from compressor characteristics. The turbine is assumed to be choked and the turbine corrected air mass flow is constant with respect to variations of pressure ratio (Glone.1998).

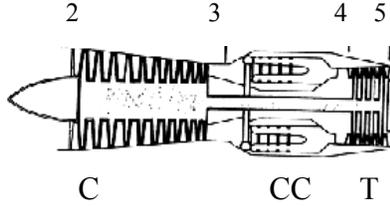


Figure 1: Gas generator

The pressure losses at inlet ducting, combustion chamber and exhaust ducting are essentially secondary effects, and may be considered constant at off-design conditions. As mentioned before the turbine power is equal to compressor power at steady state mode therefore, to calculate a steady state operating point, a point must be found so that the net power is equal to zero. Then the equilibrium running line is obtained by joining up these points.

A constant speed is selected from compressor characteristics and then a pressure ratio is guessed. Using this speed and pressure ratio and this fact that the compressor power is equal to turbine power, the compressor outlet temperature and fuel mass flow are calculated. The compressor pressure ratio is also calculated from turbine characteristic again and compared with the pressure ratio, that is guessed. If these two pressure ratios are equal, the steady state point is found.

## Engine Transient Mode

Transient performance deals with the operating regime where engine parameters change with respect to the time. If a single spool turbojet engine is operating in a steady state point and the fuel flow is increased by the control system, due to the increased temperature, the turbine power output increases. This extra power drives the compressor to a higher rotational speed. In other words the unbalanced power produces spool acceleration. This acceleration continues until the steady state condition corresponding to the new fuel flow is achieved. Conversely, for a deceleration, the unbalanced power is negative and the spool speed reduces accordingly. During transient operation a gas turbine can be considered to satisfy compatibility of flow, not the work, and the net power applied to the rotor can be used to calculate its acceleration or deceleration. The acceleration of the compressor rotor and the excess torque  $\Delta G$  are related by the following Newton's second law of motion:

$$\Delta G = J\dot{\omega} \quad (1)$$

Where  $J$  is the polar moment of inertia of the rotor and the  $\dot{\omega}$  is its angular acceleration. The excess torque  $\Delta G$  is given by:

$$\Delta G = G_t - G_c$$

where  $G_t$  is the turbine torque and  $G_c$  is the compressor torque. The net power can be written as below:

$$P_{net} = (P_t - P_c) = \Delta G \cdot \omega, \quad \omega = \frac{2\pi N}{60}$$

$$\Delta G = \frac{60 \times (p_t - p_c)}{2\pi N} \quad (2)$$

where  $p_t$  is the turbine power,  $p_c$  is the compressor power and  $N$  is the rotor speed. Finally by use of these equations the rotor acceleration is given by:

$$\dot{N} = \left( \frac{60}{2\pi} \right)^2 \frac{(p_t - p_c)}{N \cdot J} \quad (3)$$

## Compressor

Energy storage law is used to calculate the compressor power

$$P_c = w_3 \cdot C_{pc} \cdot \Delta T_{23} \quad (4)$$

where  $w_3$  is the compressor outlet air mass flow and  $\Delta T_{23}$  is the compressor temperature rise. Relationship between air mass flow, pressure ratio and speed is given by:

$$w_2 = w_3 = f(\pi_c, N) \quad (5)$$

where  $w_2$  is the compressor inlet air mass flow,  $\pi_c$  is the pressure ratio and  $N$  is the rotational speed. Compressor characteristics play the essential role in gas turbine engine performance specifications. Thus the use of accurate compressor characteristics will give accurate calculations.

In this paper the compressor characteristic curves used for modeling are extracted from two different methods including generalized and designed method. Using compressor pressure ratio, the compressor temperature rise is given by:

$$\Delta T_{23} = \frac{T_2}{\eta_c} \cdot ((\pi_c)^\gamma - 1) \quad (6)$$

Where  $T_2$  is the compressor inlet temperature,  $\gamma$  is the ratio of specific heats and  $\eta_c$  is the compressor isentropic efficiency. As mentioned before, in this paper, variations of  $\gamma$  and  $\eta_c$  are considered and then the results have been compared with constant parameters.

## Combustion Chamber

The effect of combustion chamber volume has been considered in calculations. Combustion chamber mass storage can be modelled as follows:

$$\frac{d\rho_4}{dt} = (w_3 - w_4 + \dot{m}_f) / v_{comb} \quad (7)$$

Combustion chamber energy storage can also be written as follows:

$$\frac{dU}{dt} = C_{pc} \cdot T_{t3} \cdot w_3 - C_{pt} \cdot T_{t4} \cdot w_4 + \dot{m}_f \cdot h_{PR} \quad (8)$$

Assuming the variation of energy at inside the combustion chamber is equal to the variation of energy at outlet of combustion chamber, the value of energy is given by

$$U = U_{comb} = U_4 = \rho_4 \cdot v_{comb} \cdot C_{pt} \cdot T_{t4} \quad (9)$$

Substituting equation (8) into equation (9) gives:

$$\frac{dT_{t4}}{dt} + \frac{C_{pt} \cdot w_4 + C_{pt}(w_3 + \dot{m}_f - w_4)}{v_{comb} C_{pt} \rho_4} T_{t4} = \frac{C_{pc} \tau_c T_{t2} w_2 + \dot{m}_f h_{PR}}{v_{comb} C_{pt} \rho_4} \quad (10)$$

where  $v_{comb}$  is the combustion chamber volume,  $\rho_4$  is the density,  $C_{pt}$  is the specific heat and  $T_t$  is the total temperature.

### Turbine

The turbine power is obtained from energy storage by the following equation:

$$P_t = \eta_t w_4 C_{pt} \Delta T_{45} \quad (11)$$

Where  $\eta_t$  is the turbine efficiency,  $\Delta T_{45}$  is the turbine temperature drop and  $w_4$  is the turbine inlet air mass flow. In the mathematical model it is assumed that the turbine is choked and thus, the turbine characteristic curve can be considered constant.

Using equation (3) and with the help of equation (4) and equation (11) the rotor acceleration in transient mode can be calculated using the following equations:

$$\dot{N} = \frac{1}{\left(\frac{2\pi}{60}\right)^2 JN} \{C_{pc} w_3 T_{t2} [(1 + \frac{\dot{m}_f}{w_3}) \frac{C_{pt}}{C_{pc}} \frac{T_{t4}}{T_{t2}} \eta_t (1 - \tau_t) - (\tau_c - 1)]\} \quad (12)$$

Where  $\tau_t$  is the turbine temperature ratio.

The flow chart of the transient performance calculations is presented in fig. (3).

## SIMULATION RESULTS ANALYSIS COMPARED WITH EXPERIMENTAL DATA

The modeling presented above has been used to simulate the turbojet engine performance using SIMULINK software as shown in fig. (2) and fig. (4).

The results from simulation of engine performance are shown in figures (5) to (10). Fig. (6) shows the steady state performance results using both generalised and designed compressor characteristics. As it can be seen in this figure both methods predict the steady state performance very close to each other. On the other hand, fig. (7) shows the effects of variations of  $\gamma$  and  $\eta_c$  on engine performance. This figure reveals that the variation of  $\gamma$  and  $\eta_c$  have considerable effects on steady state engine performance prediction. Fig. (5) compares the simulation results with experimental data showing good agreement.

The results obtained from transient simulation are shown in figures (8) to (10). Fig. (8) depicts the transient performance for acceleration and deceleration modes. The effects of variation of  $\gamma$  and  $\eta_c$  on transient performance are shown in fig. (9). As it can be seen in this figure, variations of  $\gamma$  and  $\eta_c$  have considerable effects on the variation of engine performance during transient modes. Furthermore, fig. (10) depicts the effects of compressor characteristics modeling methods on the transient behavior. As shown in this figure, the acceleration time predicted by both methods is the same. However, the prediction of surge margin during transient mode is quite different especially at lower speed.

## CONCLUSION

In this paper modeling and simulation of steady state and transient performance of a gas turbine engine was presented. Two methods (generalized and designed) were utilized for compressor characteristics modeling. In addition, effects of  $\gamma$  and  $\eta_c$  variations were considered. Simulation program was developed using Simulink and the results were compared with test results for a turbojet engine. These results showed that it is appropriate to use characteristic curves from generalized method to predict the steady state performance. The same trend was also seen for accelerating time. However,  $\gamma$  and  $\eta_c$  if considered constant would lead to error in steady state performance especially at low speeds.

## REFERENCES

- Crosa, G, F. Pittaluga, A. Trucco, F. Beltrami A. To relli, F.Traverso. 1998. "Heavy duty gas turbine plant aerothermodynamic simulation using simulink." ASME 550/ vol. 120, (Jul)
- Doug Garrard. 1996 "ATEC: The aerodynamic turbine engine code for the analysis of transient and dynamic gas turbine engine system operations " ASME 96-GT-193
- Visser W P. J. 2000 "Gas turbine engine simulation at NLR.", National Aerospace Laboratory NLR, Amsterdam, the Netherlands
- Lombardo,G. 1996 "Adaptive control of gas Turbin engine for axial compressor faults.", ASME, 96-GT-445,



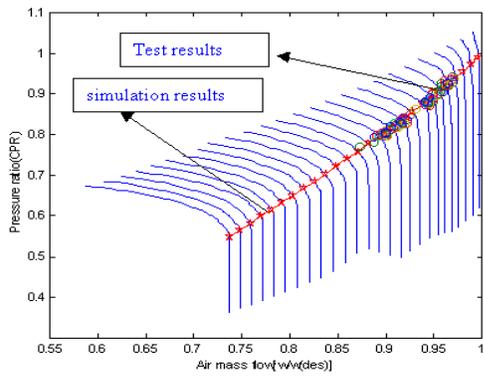


Figure 5: Comparison between simulation and test results at steady state condition

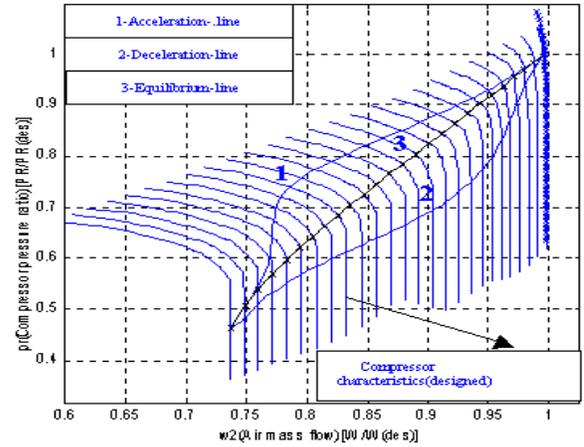


Figure 8: Transient performance

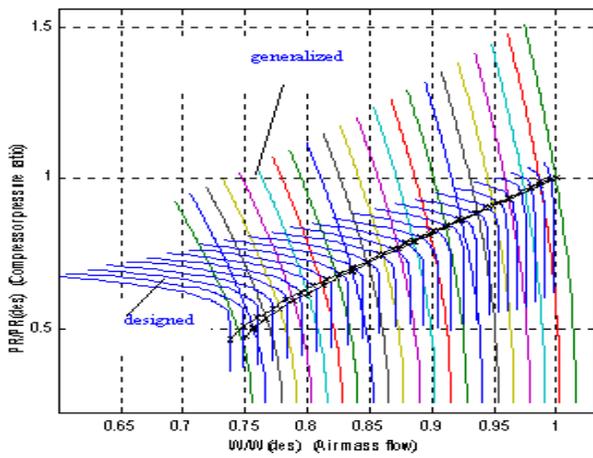


Figure 6: Equilibrium running line using two compressor characteristics methods

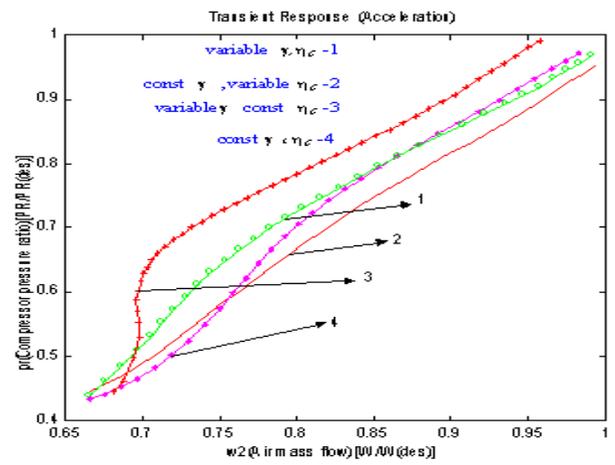


Figure 9: The effects of variation of  $\gamma, \eta_c$  on transient performance

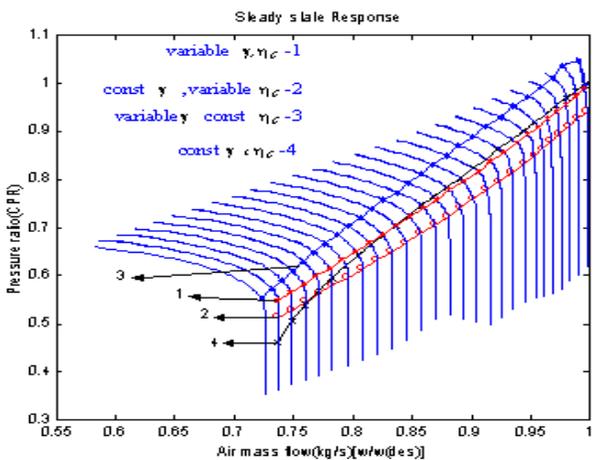


Figure 7: The effects of variation of  $\gamma, \eta_c$  on steady state performance

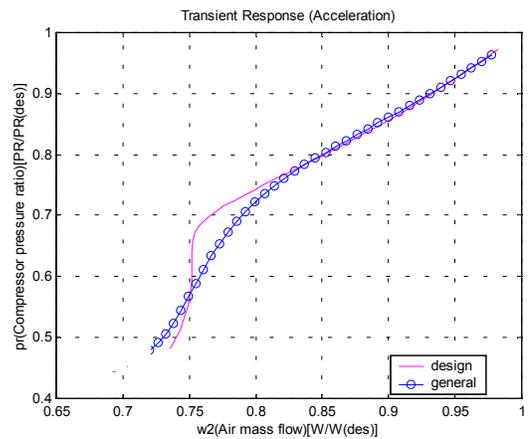


Figure 10: Engine acceleration line using two compressor characteristics methods

# FORECAST OF THE DURABILITY OF THE CONTACTS OF A CONTACTOR WITH THE TURNING ARC

Allagui Hatem, Ouni Abdel Aziz and Javoronkov Mikhail  
Laboratoire des Systèmes Electriques  
Ecole Nationale d'Ingénieurs de Tunis,  
BP 37- le Belvédère, 1002, Tunis,  
Tunisia  
E-mail: Hatem.allegui@enit.rnu.tn

## KEYWORDS

Durability, Arc in Displacement, Displacement Speed, Erosion.

## ABSTRACT

The main goal of this paper is to forecast the contact's durability when the latter is subjected to the action of an electric turning arc such as that generated in a circuit-breaker. We are interested in the range of currents, which is not represented in the curve of durability presented by the manufacturer of the apparatus. To have a comprehensive study, we carried out experiments, which showed that the arc in displacement is very advantageous and that the durability forecast is something delicate.

## INTRODUCTION

In order to create the extinction's conditions of the electric arc in low voltage circuit-breakers and also in some high voltage apparatuses, we often use the magnetic blowing technique. The displacement of the arc is caused by the interaction between the magnetic field and the arc current.

Some manufacturers of this equipment sought to develop devices with turning arc where the arc moves between two circular tracks under the magnetic field action.

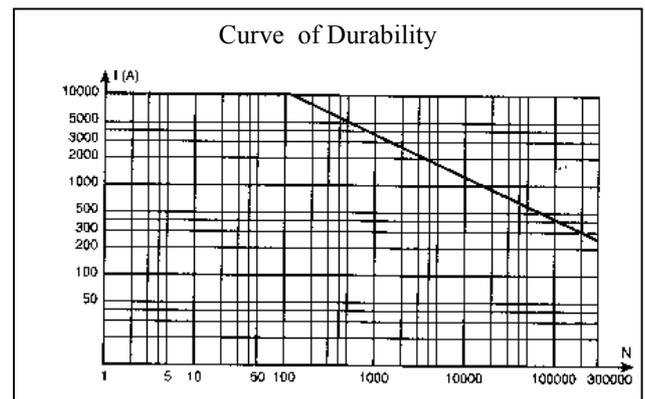
This paper is completely devoted to the forecast of the durability of the contacts of Merlin-Gérin Rollarc contactor under the action of an electric turning arc. It is based on the results of previously published work that has already gone into mathematical modelling of thermal phenomena leading to the wear of contacts (Deveautour 1992; Javoronkov, Boutkevitch, Belkine and Vedechenkov 1978)

## PROBLEM OF ELECTRIC DURABILITY OF THE CONTACTS

The electric durability that is the number of the cycles of manoeuvres (closing-opening) that an apparatus is able to make without harmful wear of the contacts was left by international standardisation to the choice of the manufacturers. In the case of the contactors and because of

high rate of manoeuvre, durability is a very significant size, which depends mainly on the break current, on the employment category and the assigned tension of employment.

We give in what follows the durability curve of this contactor (Figures 1), according to the break current in AC3 or AC4 category, extract of a technical catalogue (Merlin Gérin catalogue 1996).



Figures 1: Durability of the Rollarc Contactor according to the Break current

Within sight of this durability curve, we notice that it is represented only for the break current superior than 250 A, while we seek to determine the corresponding durability to a 40 A break current corresponding to the typical load of use of Rollarc contactor in Tunisia.

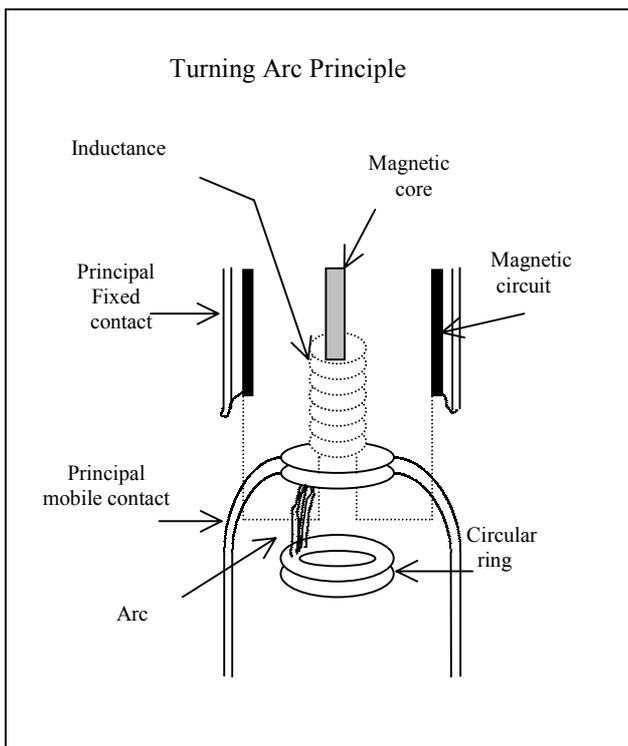
A question is raised: can we prolong the previous curve in a zone not represented by the manufacturer and with the assistance of the extrapolation method, to deduce the durability corresponding to a break current of about 40 A?

In fact, the extrapolation shows that the durability corresponding to this current is about 10.000.000 of cycles. However, this durability appears exaggerated even in mechanical point of view. In order to check this assumption, it is necessary to study the turning arc problems.

## OPERATION PRINCIPLE OF TURNING ARC DEVICES

The arc is put in motion between two circular tracks (figures 2). When the principal contacts separate, the current to stop crosses a solenoid thus creating a magnetic field. Since the separation of the circular contacts, and the apparition of the arc between them, the arc subjected to transverse induction, should then move in the direction indicated by the Laplace law. In fact, the combination of the magnetic field and the current creates a force exerted on the arc, involving this one in rotation between the two circular tracks.

The force is proportional to square current, it results a natural adaptation of this break technique to the current to stop. With high current the speed is high, the cooling is intense, right before the zero of current the speed is still sufficient to turn the arc and favours thus the dielectric rigidity recovery, the wear of the arc contacts is very weak compared to the systems with the immobile arc. With low current, the turning speed is slow, this characteristic involves a very soft extinction of arc, without over-voltage, which makes it possible with this technique to equalise the performance of breaking in air universally appreciate.

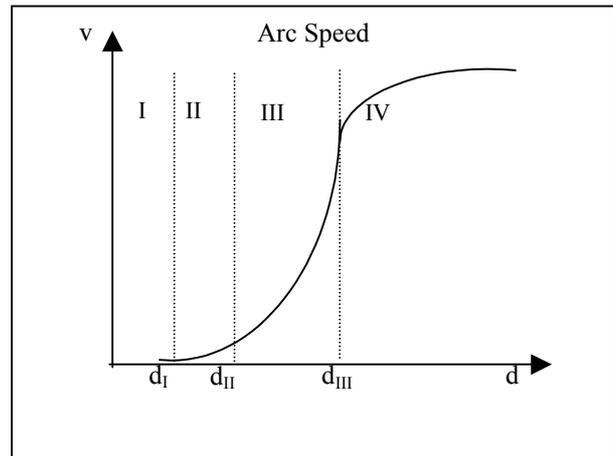


Figures 2: Schematic diagram of the electric Turning Arc

The studies of the arc in displacement were very numerous, but they present some dispersion due to the very diverse essays conditions. However, we can classify the previously published work in two categories. The first concerns the problems of electric arc displacement and the second concerns the contacts erosion according to the speed arc displacement. In the first category, it was noted that the arc

is put in motion after its ignition only if the current and the inter electrodes distance exceeded minimal values which depend on the experimental conditions (Schröder 1967; Poeffel 1980; Salge 1964).

Indeed, the speed- inter electrodes distance characteristic which was established by Schröder (Schröder 1967) is divided in several zones according to the inter-electrodes distance (figures 3).



Figures 3: Speed of the Arc Displacement according to the inter- electrodes distance

- I: zone of the arc stagnation.
- II: zone of displacement at low speed.
- III: zone of speed in growth.
- IV: zone of saturation speed.

For zone III, this author presented the following expression of the arc displacement speed  $v$  noted  $v$ :

$$v = k I^{0,61} B^{1,4} d^{2,22} \quad (1)$$

- $k$ : coefficient, which depends on material used;
- $B$ : magnetic field induction;
- $I$ : arc current;
- $d$ : inter-electrodes distance.

In addition, these studies show that electric arc displacement depends on all the parameters determining the behaviour of the electric arc in question: nature and the shape of the electrodes, nature and pressure of plasma-producing gas in which the arc develops, operating conditions to which is submitted the arc and mainly the break current and magnetic induction. In the second category of work (Salge 1964; Erk and Schröder 1968), it is noted that the erosion caused by an arc in displacement at low speed is appreciably much more than that caused by the arc, which moves at high speed (up to 30 times).

It is thus clear that the arc stagnation or the low displacement speed of this arc causing high erosion presents the problems for which it is necessary to seek solutions. In what follows, we will present our studies in this field.

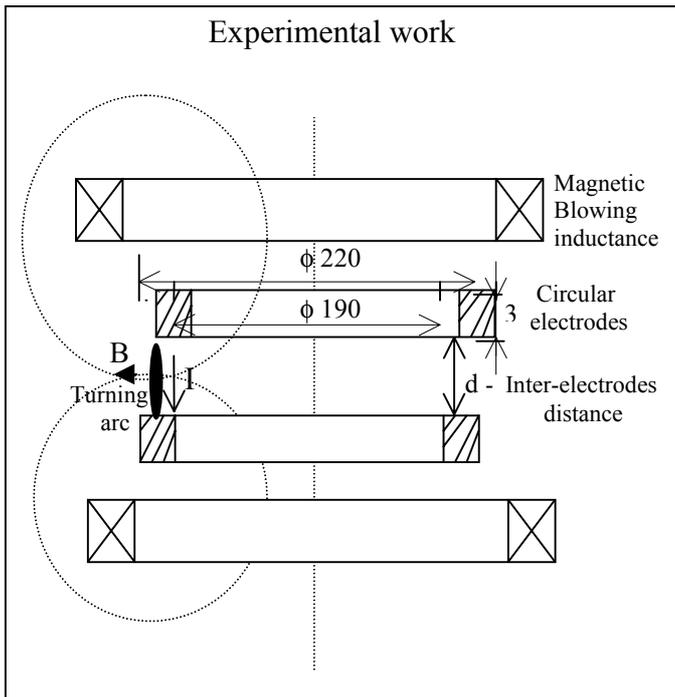
## EXPERIMENTAL WORK

### Device description

In order to carry out experiments relating to the arc displacement speed and to study the influence of many parameters on this size, we designed a turning arc device whose principle of operation is similar to that of Merlin Génin except that the circular contacts which are manufactured out of copper can be dismantled and that the arc is struck by the wire fusible technique.

This device makes it possible to fix the various parameters of the arc and its environment (break current, arc duration, magnetic field, inter electrodes distance) and to record the rotation average speed arc and erosion.

The power circuit is supplied by rectified current variable between 0 and 35 A with a tension of 550V. The excitation circuit, ensuring magnetic blowing, is independent of the circuit of the principal current, it is traversed by a direct current delivered by an auxiliary source varying from 0 to 4 A. Some sensors of measurement are used and connected to an acquisition data system.

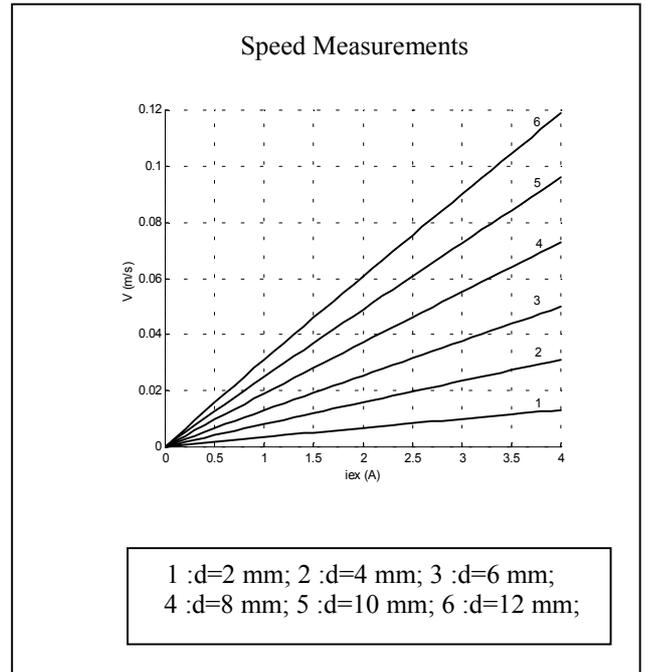


Figures 4: Device of Turning Arc

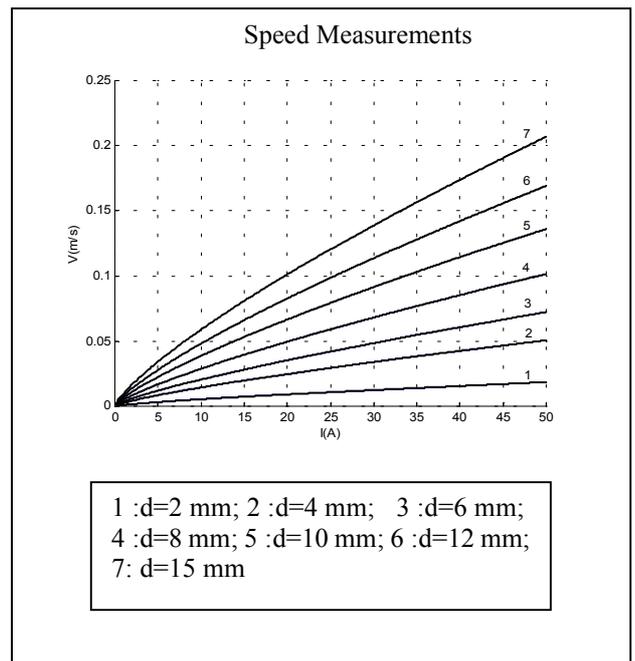
The dimensions are given in mm.

### Measurements of speed

Using this device, we raised the speed characteristics according to the operate current, the inter electrodes distance and the break current. Figures 5 and 6 present two examples of measurement.



Figures 5: Speed according to the Excitation current  $V=f(i_{ex})$  with  $I = 18$  A and  $d = \text{constant}$



Figures 6: Speed according to the Excitation current  $V=f(i_{ex})$  with  $I = 2,5$  A and  $d = \text{constant}$

Based on speed experiments, we have developed the following expression, which takes account of all the forgoing cited influence factors:

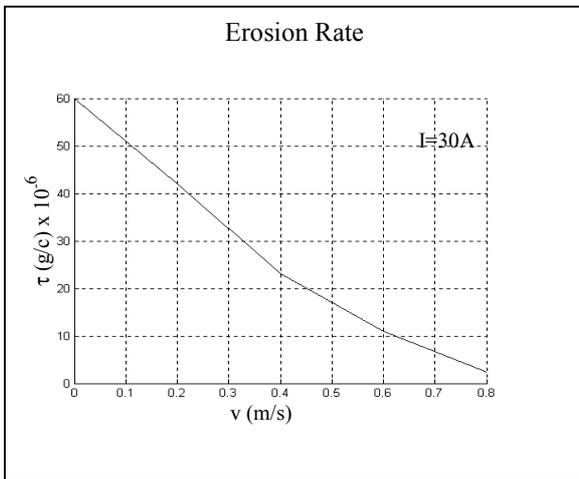
$$v = k I^{0,78} i_{ex}^{0,97} d^{1,24} \quad (2)$$

where  $i_{ex}$  is the excitation current.

**Measurement of erosion**

The sizes to be measured in this experiment are the arc current, the existence time arc, the displacement arc speed and the erosion. Since, the current and existence time of the arc are the principal influence factors, it becomes interesting to present erosion in a ratio form of the mass lost by the product of the current and time. This size, which is noted  $\tau$  allows establishing a simplified calculation of erosion. Indeed, the lost mass is measured using a precision electronic balance.

The first experiment consists in applying an electric arc current of about 30 A during 60 seconds, while varying speed by the variation of the excitation current. We have applied a relatively long arc duration in order to seize the loss of mass well. The results can be presented in figure 7.



Figures 7: Erosion Rate according to the Arc Displacement Speed

The second experiment in this part consists in applying electric arcs of various current intensities during 60 seconds as well, while varying the displacement arc speed. Some results of this experiment are presented in the following tables (table 1 and 2):

Table 1: Rate of Erosion for I=30 A

Iex (A)	0	2,5
V(m/s)	0	0,8
$\tau$ (g/c)	$6 \cdot 10^{-5}$	$2,310^{-6}$

Table 2: Rate of Erosion for I=10 A

Iex (A)	0,5
V(m/s)	0,2
$\tau$ (g/c)	$22 \cdot 10^{-5}$

Based on these experiments, we notice that when V decreases, the erosion increases until reaching the one produced by the immobile arc and that erosion Rate is not important if the arc turning speed is high.

**RESULTS ANALYSIS**

In this part, we seek to exploit the results of our experiments and those of the previous published work in order to answer the question already raised: will the durability of the Rollarc contractor increase or not if one exploits it under-load rating? For the traditional contactors, the answer is obvious: yes.

Indeed, many experimental studies of the erosion by the immobile arc made it possible to establish the empirical formula of the mass erosion:

$$M = k I^\alpha t^\beta \tag{3}$$

K,  $\alpha$  and  $\beta$  are parameters dependent on the conditions of the experiment and nature of materials of contacts.

For the contactors, H.W. Turner and its collaborators (Turner H, Turner. C and Frey 1966) received the following formula:

$$M = k I^{1,6} t \tag{4}$$

The formula (3) and (4) present erosion as thermal phenomena leading to evaporation, fusion and evacuation of the evaporated and melted matter (and sometimes solid matter). It is obvious that with the reduction of the current or time, the heat introduced by the electric arc into the contacts decreases and the volume or the mass of the evaporated matter decreases too. But for the systems with turning arc, the time intervenes through the displacement arc speed. However, arc in displacement can be modelled by a motionless model arc which moves by jumps (Allagui, Javoronkov and Chaabane 2000), the time of stagnation of the arc before the following jump is determined by the average speed of displacement. In this case, the reduction speed must then cause increase of the time of arc stagnation.

In case of the Rollarc contactor, if the speed displacement of the arc satisfies the expression (1) and since the break current is the source of magnetic blowing ( $I=B$ ), the speed displacement arc can be given by:

$$v = k I^2 d^{2,22} \tag{5}$$

According to this expression, when the break current decreases from 315 A to 40 A (315 A: nominal current contactor), that is to say approximately 8 times, the average speed displacement of the arc decreases by 64 times and this way there is the risk that the speed displacement arc passes to function in a low speed zone (zone II).

The experiments, which we carried out, were precisely in this zone (0 to 0,8 m/s) and it was noted that the speed displacement arc as an erosion factor becomes dominating compared to the current. According to tables 2 and 3, despite the decrease of the current by 3 times (30 A to 10 A), the erosion increases by 10 times when speed decreases by 4 times.

Therefore, when the Rollarc contactor is exploited in a low speed displacement arc zone, even if the current decreases, the erosion augments. The extrapolation in this zone of the curve of figures 1 is not thus valid.

## CONCLUSION

Being given that contacts erosion generally determines the lifetime of whole apparatus, we practically studied the erosion of the copper contacts according to the current and the time which intervenes intermediary by speed the displacement arc. We found that the current reduction will not necessarily cause the reduction of the erosion in the turning arc devices.

## REFERENCES

- Allagui, H, Javoronkov, M and Chaabane, A. 2000. "Erosion Produced by the Electric Arc in Displacement". Tunisian Days of Electrical Engineering and Automatic, Tunisia (Mar).
- Devautour, J. 1992. "Contribution to the study of the arc- electrodes interactions. Influence of metallurgical structure on the mechanisms of erosion of the circuit-breakers" Thesis of doctorate of the Paris 6 University (Jul).
- Erk In, Schröder K. H. 1968. "Über Material verlust homogener und Heterogener Kontaktwerkstoffe für Schaltgerät put Magnetischer Licht-Bogenbelastung". Elektrotechn. Ztschr, 373-377.
- Javoronkov, M, Boutkevitch, G, Belkine, G and Vedechenkov, N. 1978. "Electric Erosion of the Contacts and the Electrodes in Fort Currents" Energy Edition, Moscow.
- Merlin Gerin Catalogues. 1996. Distribution HT/MT.
- Poeffel, K. 1980. "Influences of the Copper Electrode Surfaces on Movement Initial Arc" IEEE Transactions on Plasma Science, Vol, PS-8, n°4, 443, (Dec).
- Salge, J. 1964 "Über die Wanderung Von Kochstromlichtbögen in Engen Spalten EIB Unterdruck", ETZ- A, data base 85, 14, 417-425
- Schröder K.H. 1967. "Über das Wanderung Sverhalten Magnetischer Ablenkter Lichtbogen Fusspunkte auf Unterschiedliche Kontakt Werkstoffe". Kontakte in der Elektrotechnik Dresden 39-47.
- Turner H. W, C. Turner, C and F. Frey. 1966. "The Relation between Wear of Copper Contacts and Arc Current". ERA Report ref. 5119.

# SENSOR INFORMATION FUSION FOR THE NEEDS OF FAULT DIAGNOSIS IN MARINE DIESEL ENGINE PROPULSION PLANT

Radovan Antonić

Ante Munitić

Split College of Maritime Studies

Zrinsko Frankopanska 38 Split, Croatia

E-mail: [antonic@pfst.hr](mailto:antonic@pfst.hr) ; [munitic@pfst.hr](mailto:munitic@pfst.hr)

Zoran Vukić

Faculty of Electrical Engineering and

Computing, Zagreb, Unska 3, Croatia

E-mail: [zoran.vukic@fer.hr](mailto:zoran.vukic@fer.hr)

## KEYWORDS

Diesel engine propulsion plant, structural description, redundant information, sensor fusion, neural network

## ABSTRACT

Complex industrial processes like Marine Diesel Propulsion Plant (MDPP) have complex interrelations and interdependencies between variables and parameters. This characteristic could be used in estimating unknown or unmeasured variables from the information gathered by other measurements and sources using information fusion by means of a soft computing methods. In the paper, a structural analysis approach to identifying most relevant variable interrelations, components or subsystems of MDPP with inherent redundant information has been proposed. Sensor information fusion method was chosen to be using artificial neural networks (ANN). The paper presents proposed ANN with structure and learning algorithms. Simulation have been carried out in Matlab-Simulink environment with engine speed estimation example.

## 1. INTRODUCTION

Diagnostic and control systems of marine diesel propulsion plant require a large number of different sensors with different measuring types and locations at various critical points on the propulsion engine and its subsystems (temperatures, pressures, flow rates, levels, metal content of the lubricating oil, water content in the fuel oil and more). The data from sensors are collected and transmitted to the processing units.

The main purpose of most signal processing is to yield knowledge of a situation so that proper decisions can be done. Many of these signals should be combined in some way to enable decisions of such conditions as emergency states, when to change oil, time to repair or replace parts, engine efficiency etc.

In some specific situations human intuition, heuristic knowledge and experience have to be fused together with sensor data for good plant estimation (overall engine efficiency, degradation of oil condition, fault conditions,..). One effective approach in such case is information fusion that will be discussed in the paper.

In the cases when a sensor fails to operate or operates with faults, sensor information fusion methods are needed to reconstruct the lost signals - information.

Aiming to use sensor information fusion with existing sensors the need for exploring possible redundancies inherent to the system structure is evident. One suitable method is structural description or analysis of the system decomposing it into functional dependent and related components or subsystems. This approach for MDPP will be presented in the paper. Sensor information fusion method was chosen to be using artificial neural networks which are very suitable in the case of on-line gathered data.

Simulation example will be given for marine diesel engine speed estimation using redundant relations and data.

## 2. STRUCTURAL DESCRIPTION OF A SYSTEM - GENERAL APPROACH

One can consider a system  $S$  like union of its functional components  $\bigcup_{i=1}^n C_i$ , each of them

establishing some relations or constraints  $f_i$  between a set of variables and parameters (known or unknown)  $z_j$  of the system, i.e. :

$$f_i(z_1, z_2, \dots, z_p), \quad 1 < p \leq m$$

where  $f_i$  can represent dynamic, static, linear or non linear relation, crisp or fuzzy rules, empirical or any other relation-constraint.

Structural model of the system can then be represented with a set of constraints:  $F = \{f_1, f_2, \dots, f_n\}$  and a set of variables and parameters  $Z = \{z_1, z_2, \dots, z_m\} = Z^K \cup Z^X$  to which constraints are valid.  $Z^K = U \cup Y \cup C$  is a set of

known variables and parameters, where  $U$  represents a set of control variables,  $Y$  is a set of measured outputs and  $C$  is a set of known constant parameters.  $Z^X$  is a set of unknown variables and parameters of the system.

Now, the structural model of the system can be represented by directed graph with nodes and connecting arcs  $G(F, Z, A)$ . The elements of a set of arcs in such graph  $A \subset (F \times Z)$  are defined with the following mapping scheme - binary relations:

$$\begin{aligned} A &: F \times Z \quad \{0,1\} \\ A^K &: F \times Z^K \quad \{0,1\} \\ A^X &: Z^X \times F \quad \{0,1\} \end{aligned} \quad (1)$$

For more details see (Izadi-Zamanabadi 1999).

### 3. STRUCTURAL ANALYSIS OF MARINE DIESEL PROPULSION PLANT – REDUNDANT DATA AND RELATIONS

A structural analysis model to identify most relevant variable interrelations, components or subsystems of MDPP with inherent redundant information which could be used in fusion process will be explored.

#### 3.1 Structural description of MDPP

The main purpose of the structural description of MDPP here is to explore some inherent redundant relations which can be used in calculating unknown or unmeasured variables using sensor information fusion method.

Figure 1 shows the structure of MDPP with its main structural components:  $C_1$  - diesel engine dynamics,  $C_3$  - engine shaft dynamics,  $C_5$  and  $C_6$  - propeller shaft dynamics,  $C_8$  - ship speed dynamics,  $C_{10}$  - hull dynamics, and corresponding sensors: fuel index sensor  $C_2$ , engine speed sensor  $C_4$ , pitch propeller sensor  $C_7$  and ship speed sensor  $C_9$ .

Relations and constraints between variables and parameters can be obtained in various ways: by mathematical modelling, by simulation, using experimental data, eliciting expert's and operator's knowledge, etc. For details see (Antonić and Radica 1991; AntoniĆ et al. 2000; AntoniĆ and Vukić 2002; Vukić et al. 1998; Izadi-Zamanabadi 1999).

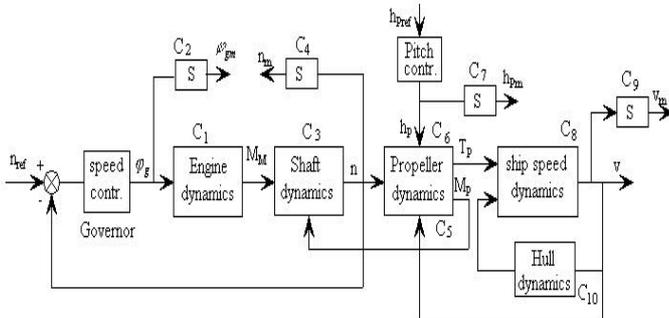


Figure 1: The structural diagram of diesel propulsion plant

where:  $n_{ref}$  - engine reference speed (set value);  $n$  - engine speed;  $\phi_g$  - fuel link position;  $h_{pref}$  - propeller pitch set value;  $h_p$  - propeller pitch;  $v$  - ship speed;  $K_M$ ,  $T_M$  - engine gain and time constant;  $M_P$ ,  $T_P$  - propeller torque and thrust;  $v_a$  - advance propeller speed;  $R_u$  - total hull resistance

The structure of MDPP in Figure 1 can be represented as union of its components :  $\bigcup_{i=1}^{10} C_i$ .

A set of constraints / structural relations is:

$$F = \{f_1, f_2, \dots, f_{10}\} \quad (2)$$

A set of known measurable variables and parameters is:

$$Z^K = \{\phi_{gm}, n_m, h_{pm}, v_m, K_M\} \quad (3)$$

A set of unknown variables and parameters is:

$$Z^X = \{\phi_g, n, h_p, v, M_M, M_P, T_P, R_u\} \quad (4)$$

The measuring noise is here neglected so:

$$n_m = n; \phi_{gm} = \phi_g; h_{pm} = h_p; v_m = v. \quad (5)$$

Adequate structure graph of the MDPP with variable and parameter relations is shown in figure 2.

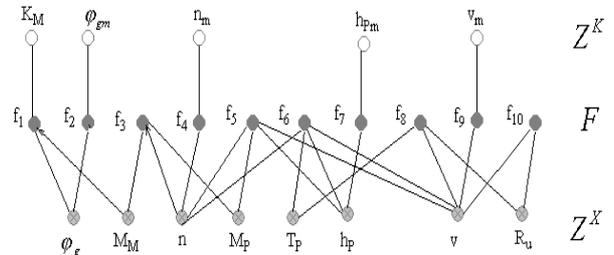


Figure 2: Structure graph of MDPP system with variable and parameter relations

#### 3.2 Redundant relations and information fusion

From the structural graph of MDPP system one can get analytical redundant relations between variables and parameters: direct relations and indirect or derived ones (with sensor information fusion and some reasoning method). For direct relations structural constraints are applied only to known - measured variables i.e. to subset  $Z^K$ , while derived relations are those to which structural constraints of unknown - unmeasured variables are applied, i.e. to subset  $Z^X$ .

Indeed, derived redundant relations are more interesting, because they result with analytical redundancy what is a key point for information fusion. These are frequently based on the human expert knowledge and operator experience.

It is evident from the structure graph model of MDPP, that there are redundant relations and information which can be used in case of faulty sensors.

For instance, in the case of engine speed sensor fault (component  $C_2$  in structural diagram) the value of the engine speed could be estimated i.e. calculated using information fusion from other sources ( $C_5$  and  $C_6$ ).

Unknown variable can be estimated by integrating several other measurements into a single robust estimator (software sensor). The fusion of data from different sensors will add new valuable information that would be otherwise unavailable. The need of data fusion arises also

from the fact the information gathered is often incomplete, uncertain, imprecise or may be from a faulty sensor. There are several possible methods for data fusion and the very effective one is artificial neural network approach.

#### 4. INFORMATION FUSION IN MDPP USING ANN APPROACH – SIMULATION EXAMPLE

The ability of ANN to learn from experience i.e. from history of data during on-line operation is making them the preferred choice for process modelling with intrinsic variable and parameter interrelations. In the above structural description of MDPP the redundant relations between variables and parameters were illustrated. Some of them will be used in the information fusion example.

##### 4.1 Engine speed estimation using information fusion Speed sensor faulty - simulation example

Engine diagnosis and control system needs speed information during normal operation and gets it continually from speed sensor.

In the case of speed sensor failure it would be desirable to have a system that could estimate engine speed (most critical variable in closed loop speed control) from various sets of inputs i.e. information sources giving redundancy in speed information and thus leading to more robust control system. That is especially important if all speed sensors (usually two) are in faulty conditions.

The required engine speed value could be estimated on-line from other variables which are related to it (see Figure 2) : propeller torque  $M_p$  or propeller thrust  $T_p$ , ship speed  $v$ , propeller pitch  $h_p$  if the propeller is controllable (CPP). Figure 3 a and b illustrate engine speed estimation from other known variables - signals measured on-line ( $M_p$ ,  $T_p$ ,  $v$ ).

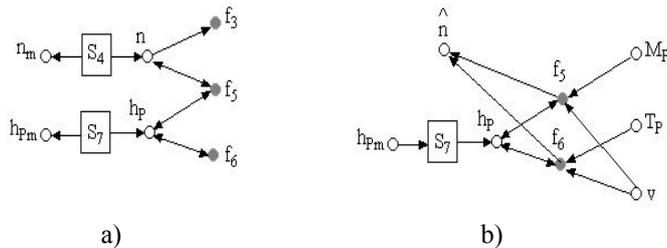


Figure 3: Engine speed estimation using information fusion

##### 4.2 Neural network structure and learning algorithms

In the engine speed estimation example three independent input signals to the ANN and one output signal which should be the best estimate of engine speed in case of faulty sensor were used.

The data from different sources are usually pre-processed (data normalization, filtering, principal component analysis, etc.) before applied to the ANN for fusion purpose.

The ANN, in this experiment, was organized in two processing stages i.e. two ANN were designed and used (Figure 4).

The first stage consists of estimation ANN and is for feature extraction from input signals. The second stage consists of ANN for information fusing i.e. decision making and selecting the best estimate from the first ANN.

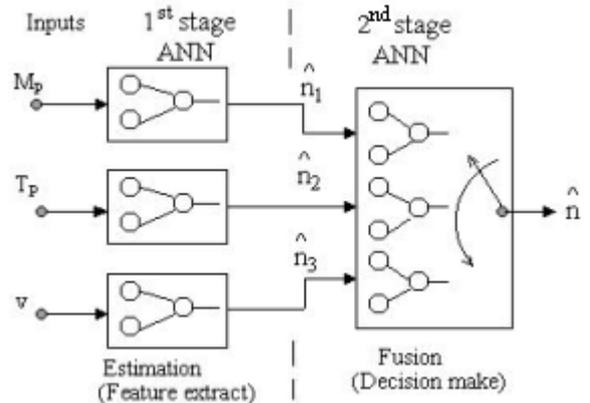


Figure 4: Concept of ANN for engine speed fusion

The first stage consists of three identical feed forward NN (in Figure 5a shown only one for input variable  $M_p$ ) each with one hidden layer with log-sigmoid transfer function and one output layer with linear transfer function. The second stage consists of self-organising NN with one competitive layer with three inputs (these are outputs from the first stage) and one output ADALINE stage (in figure 5 b). There are three neurones in competitive layer and only one is a winner in a time. Euclidean distance measure (see AntoniĆ and Vukić 2002) in decision making i.e. choosing the best estimate in each time step was used. In the estimation stage of NN, 3 inputs are fed (propeller torque  $M_p$ , propeller thrust  $T_p$  and ship speed  $v$ ) to estimate engine speed  $n$ .

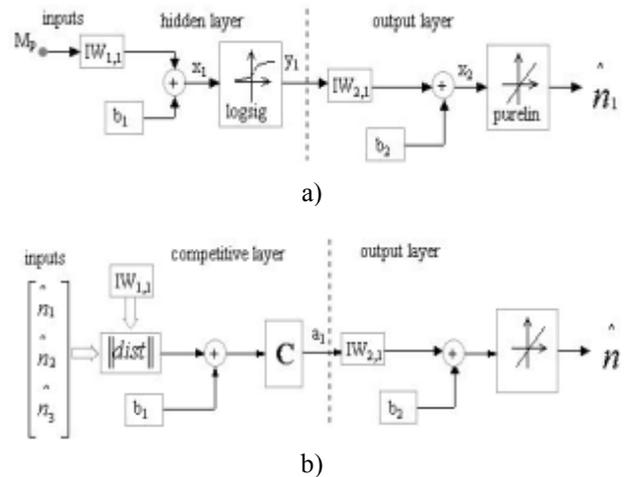


Figure 5: Structure of ANN for engine speed estimation

The mse (the mean squared error between the target i.e. expected values and the network outputs – estimated values) performance function is chosen as a criterion.

$$mse = \frac{1}{N} \sum_{k=1}^N (n(k) - \hat{n}(k))^2 \quad (6)$$

Performance goal was set to  $mse = 0.01 \text{ rads}^{-1}$ .

In minimising performance function the gradient descent back-propagation learning algorithm for updating network weights and biases with adaptive learning rate was used.

$$x(k+1) = x(k) - \alpha(k)g(k) \quad (7)$$

where  $x(k)$  is a vector of current weights and biases,  $g(k)$  is the current gradient and  $\alpha(k)$  is the learning rate.

For comparison purpose we used two learning algorithms:

- Quasi-Newton (BFGS) learning algorithm,

$$x(k+1) = x(k) - H^{-1}(k)g(k) \quad (8)$$

$H(k)$  is the Hessian matrix (second derivatives) of performance function at the current values of the weights and biases.

- Levenberg-Marquardt learning algorithm

$$x(k+1) = x(k) - [J^T J + \mu I]^{-1} J^T e \quad (9)$$

where  $J$  is the Jacobian matrix which contains first derivatives of the network errors with respect to the weights and biases,  $e$  is a vector of network errors,  $\mu$  is a scalar.

### 4.3 Simulation results in engine speed estimation

The training set used for the proposed ANN is obtained from the real diesel engine propulsion plant simulator PPS2000 (Norcontrol) with propulsion diesel engine MAN B&W type 5L90MC with maximum power of 18.000 kW installed on the very large crude carrier, (fully loaded). We've got training set values with diesel engine working in four basic operating regimes – modes (table 1): Full ahead (with engine power of 100 %), Half (engine power of 75 %), Slow (engine power of 50 %) and Dead slow (engine power of 25 %).

Table 1: Simulated engine data for training ANN

Engine regime	Engine power (%)	Engine speed - n (rad/s)	Mp (Nm) x10 <sup>6</sup>	Tp (N) x10 <sup>6</sup>	Ship speed v (m/s)
Full Ahead	100	7.74	2.20	1.46	7.71
Half	75	7.02	1.90	1.21	7.06
Slow	50	5.14	1.05	0.66	5.11
Dead slow	25	3.10	0.41	0.26	3.10

The second part of the simulation was carried out by using Matlab/Simulink environment.

After training the ANN given in Figure 5 using training data set from table 1, we've got very good results for engine speed estimates in four operating points (Full Ahead, Half, Slow, Dead slow). These are presented in table 2 and Figure 6. The differences between speed estimates are very small (with mse:  $3.3 \cdot 10^{-3}$  with Mp data,  $9.98 \cdot 10^{-4}$  with Tp and  $6.16 \cdot 10^{-4}$  with v data set).

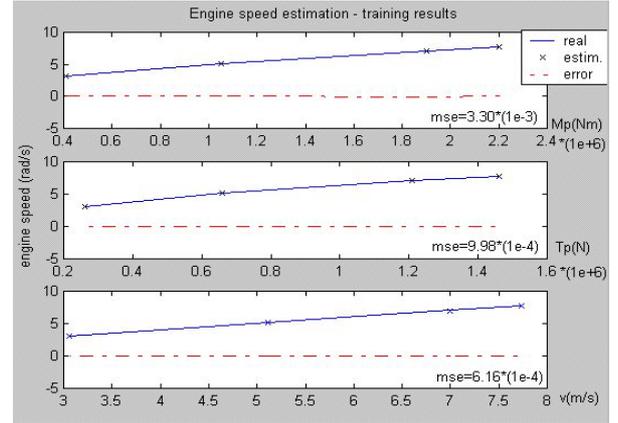


Figure 6: Engine speed estimation with training data set from Mp, Tp, v

Table 2: Estimated engine speed n from Mp, Tp, v

Engine speed (target) n (rad/s)	Estimated speed from other signals (with training data)		
	Mp	Tp	v
7.740	7.682	7.712	7.692
7.020	7.114	7.065	7.073
5.140	5.097	5.114	5.124
3.100	3.114	3.113	3.118

Comparing results obtained during training session of NN with two different learning algorithms: Levenberg-Marquardt (LM) and Quasi-Newton we've noticed very little difference (table 3).

Nevertheless, we prefer LM learning algorithm because the estimation error (mse) and training period (epochs) were a bit lesser.

Table 3: Comparison results of two learning algorithms in training NN

	Levenberg-Marquardt			Quasi-Newton		
	Mp	Tp	v	Mp	Tp	v
mse	$3.30 \cdot 10^{-3}$	$9.98 \cdot 10^{-4}$	$6.16 \cdot 10^{-4}$	$3.30 \cdot 10^{-3}$	$9.79 \cdot 10^{-4}$	$9.61 \cdot 10^{-4}$
epoch	>500	115	26	>500	118	35

Performance goal (mse = 0.01) for the best engine speed estimate (with Tp data set) was reached in very short time (4.17 s) i.e. after only 115 epochs of training.

Applying testing data set to ANN concurrently for three inputs: Mp, Tp and v, less accurate results were obtained (Figure 7) but nevertheless useful for practical use, except those estimated from ship speed data where the mean squared error was 11.55 %. The best results were obtained from propeller thrust measurement Tp (mse = 1.73 %).

The largest discrepancy between training and testing results were obtained for ship speed signal, maybe because of small training set.

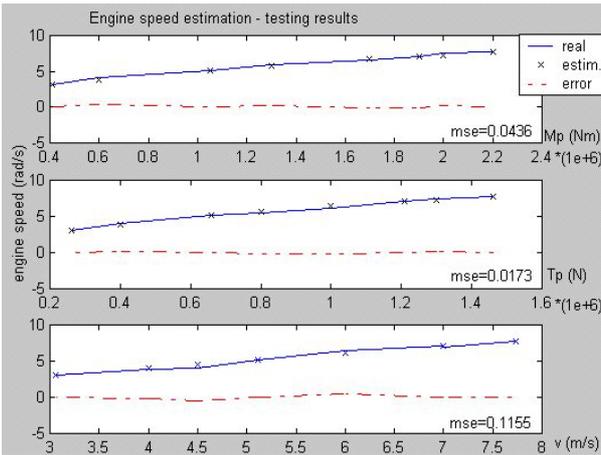


Figure 7: Engine speed estimation with testing data set from  $M_p$ ,  $T_p$ ,  $v$ .

In each time step, the designed ANN chooses the best estimate on its output so the final results were acceptable. Testing example with engine power of 100 % and expected real value of  $n = 7.74$  rad/s: the best speed estimate, was with  $T_p$  data:  $n_e = 7.712$  rad/s (see Figure 8a). We also tested ANN output in the case of lost one or even two of three input signals and have got good engine speed estimate. Figure 8b illustrates situation with two input signals missed (sensor faults). The ANN output was  $n_e = 7.785$  rad/s (Figure 8b).

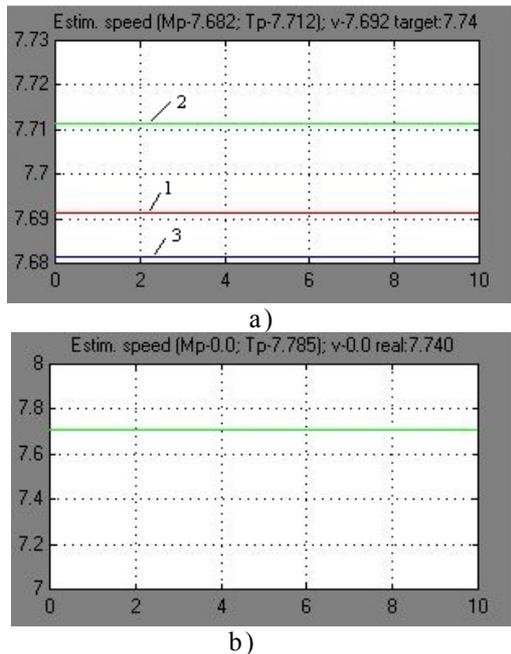


Figure 8: Engine speed estimate (best ANN output)

Applying testing data within all operating regions is illustrated in Figure 9. Test results for engine speed estimate are fairly good for  $M_p$  and  $T_p$ .

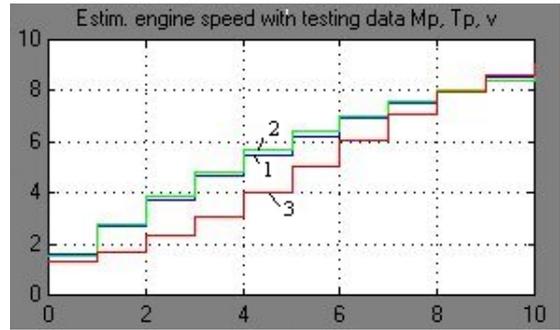


Figure 9: Estimating engine speed with testing data  $M_p$ ,  $T_p$ ,  $v$  within the operating region

Data fusion of three signals with expert modification of contribution coefficients on engine speed with  $K_{M_p} = 0.34$ ,  $K_{T_p} = 0.36$ ,  $K_v = 0.30$  had given quite good estimate (see Figure 10).

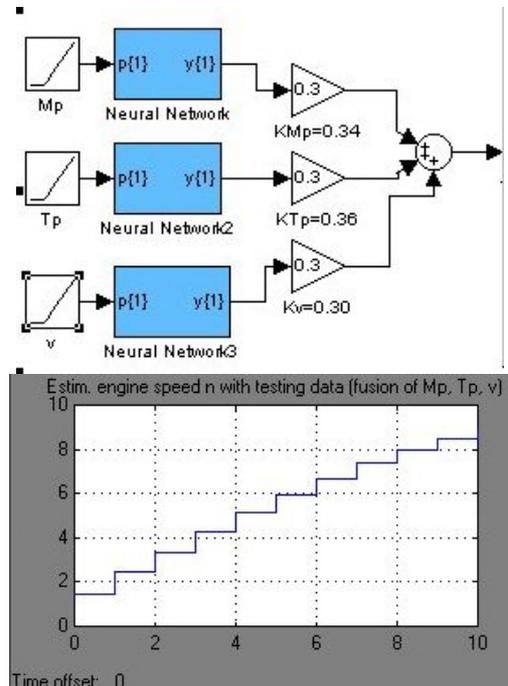


Figure 10: Estimating engine speed with data fusion of  $M_p$ ,  $T_p$ ,  $v$

Finally, three testing cases with  $M_p$  as input signal to the ANN has been studied in parallel and the output (speed estimate) was recorded in the diagram (see Figure 11): The first case was with  $M_p$  as only input signal. Input signal in the second case was  $M_p$  with the added noise (zero mean Gaussian with variance of 0.02). In the third case, the disturbance signal (sine wave of amplitude of 0.1 and frequency of 1 rad/s) was added to  $M_p$ . We could conclude that proposed fusion scheme is rather robust to noise and disturbance in input signals.

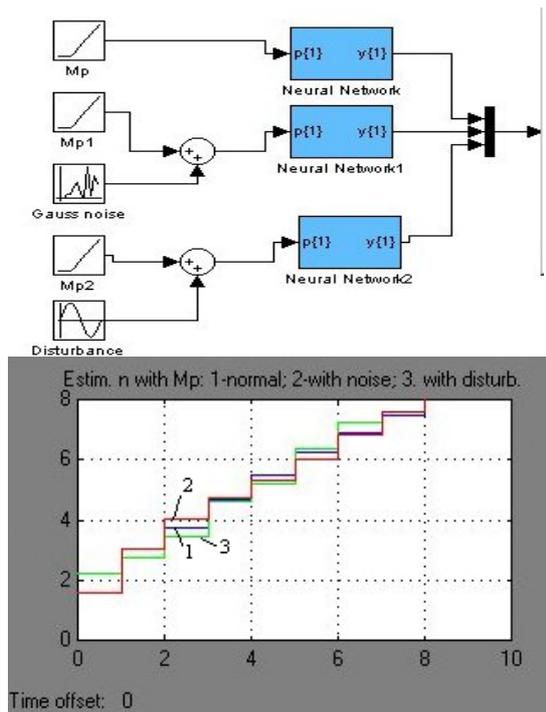


Figure 11: Estimating engine speed with  $M_p$ : 1 – normal ; 2-with Gauss noise; 3-with sine disturbance

## CONCLUSION

Sensor information fusion concept is becoming more and more attractive especially in the area of diagnostics and control systems. Some important advantages of using information fusion in combination with soft computing technologies like artificial neural networks, fuzzy logic, genetic programming could give more robustness, reliability, fault tolerance and intelligence to control systems.

The structural approach is presented and applied to the marine diesel engine propulsion plant as an effective method to identify the subsystems with inherent redundant information. Based on that analysis we proposed ANN for information fusion process which consists of two stages: The first stage is an estimation ANN for feature extraction from input signals and the second stage is for information fusing i.e. decision making and selecting the best estimate from the first stage ANN. We tested it with the simulation example. Diesel engine speed was estimated on the basis of three other signals: propeller torque  $M_p$ , propeller thrust  $T_p$  and ship speed  $v$ . It was shown that good speed estimation could be obtained using other available information in the case of faulty speed sensor. Only a part of the obtained results was presented in the paper.

The proposed fusion scheme was also tested with noise and disturbance signals added to the  $M_p$  input signal and concluded fairly good scheme robustness.

Better results would probably be obtained if larger sets of training and testing data were used. The generalisation scheme in the sensor information fusion within MDPP will be of our interest in the near future.

## REFERENCES

- Antonić, R. and G. Radica. 1991. "Diesel Engine Diagnostic Simulation Model based on Trend Curves", Proceedings of 33'rd Symposium ETAN in Marine, Zadar, 29-31.
- Antonić, R., Z. Vukić, and I. Kuzmanić. 2000. "Basic Principles and Techniques of Expert Knowledge Elicitation for the Needs of Technical Systems Diagnostics", Brodogradnja 48,4, 330-337.
- Antonić, R and Z. Vukić. 2002. "Knowledge Representation in Diagnosis and Control of Marine Diesel Engine", Brodogradnja 50,1, 57-66.
- Izadi-Zamanabadi, R. 1999. Fault Tolerant Supervisory Control - System Analysis and Logic Design, PhD Thesis, Department of Control Eng. , Aalborg Univ. , Denmark.
- Samarasooriya, V.N.S. and P.K. Varshney. 2000. "A fuzzy modelling approach to decision fusion under uncertainty", Fuzzy sets and systems 114, 59-69.
- Vukić, Z. H. Ožbolt and M. Blanke. 1998. "Analytical Model-based Fault Detection and Isolation in Control Systems", Brodogradnja 46, 3, 253-263
- Zilouchian, A. and Mo. Jamshidi. 2001. Intelligent Control Systems Using Soft Computing Methodologies. CRC Press, 2001

## BIOGRAPHY

**Radovan Antonić** received B.Sc. degree in electrical engineering from University of Split in 1980, and M.Sc. degree in electrical engineering - automation from Zagreb University in 1986. He finishes PhD thesis at Faculty of electrical engineering and computing of Zagreb University. He is IEEE member and research interest include control systems, fault tolerant control, expert systems, specially in marine applications. Currently, he is a higher lecturer at Split College of Maritime Studies.

**Zoran Vukić** received his B.Sc, M.Sc. and Ph.D. in electrical engineering from University of Zagreb in 1972, 1977 and 1989 respectively. He is professor at the University of Zagreb, Department of Control and Computer Engineering in Automation. His research interest is in fault tolerant control, adaptive & robust control, nonlinear control, guidance and control of marine vehicles. He is member of IEEE Control Sys. Soc., IEEE Oceanic Eng. Soc. and IFAC Technical Committees on Marine systems and Adaptive and learning control. Author or co-author of more than 100 bibliographical units (conference and journal papers, research projects, referee reports, textbooks etc.).

**Ante Munitić** received his first B.Sc. in Electric and Energetic Engineering in 1968, and his second in 1974, his M.Sc. degree Organisation System and Cybernetics Science (Operational Research) in 1978, and his Ph.D. of Organisation Science (System Dynamics), in 1983. He is currently a full Profesor of Computer and Informatic Science at the University of Split. He has published over 80 papers on system dynamics modeling and simulation, operational research, marine automatic control system and The Theory of Chaos. He has published two books: "Computer Simulation with help of System Dynamics" and "Basic Electric Energetic and Electronics Engineering".

# FRACTIONAL INFINITE IMPULSE RESPONSE FILTERS

R. El-Khazali, M. Al-Mualla, H. Al-Ahmad  
Etisalat College of Engineering  
Sharjah-UAE

[khazali@ece.ac.ae](mailto:khazali@ece.ac.ae), [almualla@ece.ac.ae](mailto:almualla@ece.ac.ae), [alahmad@ece.ac.ae](mailto:alahmad@ece.ac.ae)

## KEYWORDS

Fractional System, Fractional Filters, IIR Filters

## ABSTRACT

This paper introduces a finite dimensional discrete model for a fractional infinite impulse response filters (FrIIR). The fractional filters represent an infinite dimensional system which obeys a non-Newtonian behavior. The FrIIR filters are approximated by finite but large dimensional digital filters which can be implemented. The frequency response and the stability of the digital FrIIR filter are discussed. The main ideas of this work are illustrated by simple numerical example.

## 1. INTRODUCTION

A growing interest in fractional systems is evolved lately. The fractional systems describe system dynamics that obey non-Newtonian motions. Such systems are described by differential equations of non-integer order. Many researchers investigated the system performance using the theory of fractional calculus [1-9]. Necessary and sufficient conditions for fractional systems stability are discussed in [5,6]. The stability of continuous fractional system are also discussed in the sense of Lyapunov [2]. New attempts to thoroughly understand the behavior of circuits, oscillators and passive filters have also been discussed by [10]. In their approach, the bandwidth of the passive low pass continuous filters was found to be proportional to the order of the fractional capacitor.

This paper is organized as follows: some background discussion is introduced next. Section 3 defines a discrete model for the fractional filters, while a finite realization model for the infinite dimensional system is introduced in section 4. The basic idea of this development is illustrated via simple examples.

## 2. BACKGROUND

Consider first a single-variable fractional system in the continuous time described by:

$$D_t^\alpha x(t) = \lambda x(t) + bu(t), \quad x(t_0) = x_0 \quad (1)$$

where  $x \in \mathfrak{R}$ ,  $u \in \mathfrak{R}$ ,  $D_t^\alpha x(t) \equiv \frac{d^\alpha x(t)}{dt^\alpha}$ ,  $b$  is a real constant while  $\lambda \in \mathcal{C}$  is a complex number and  $0 < \alpha \leq 1$ .

The stability of the continuous fractional system given by (1) is discussed in [5, 6] and stated for completeness in the following theorem.

**Theorem 1 [6]:** Consider system (1). Let  $\lambda = re^{i\theta}$ . The homogeneous part of system (1) is:

- 1- *Asymptotically stable iff  $|\theta| > \frac{\alpha\pi}{2}$ .*
- 2- *Stable iff either it is asymptotically stable, or the critically stable eigenvalues which satisfy  $|\theta| = \alpha\pi/2$  have geometric multiplicity one.*

Theorem 1 defines the stability region outside the closed angular sector  $|\theta| \leq \alpha\pi/2$ . Generalized stability conditions of multi-fractional LTI systems are introduced in [6], while sufficient stability conditions for fractional systems using the second method of Lyapunov are discussed in [2, 4].

Notice that the autonomous part of system (1) decays in the order of  $t^{-\alpha}$  as  $t \rightarrow \infty$  which is of a slower rate than any exponential decay [5].

In order to derive a discrete model of the Fractional filter given by (1), one may consider the Grünwald-Letnikov approximation of a fractional derivative of order  $\alpha$  [10].

$$\frac{d^\alpha x(t)}{dt^\alpha} \triangleq h^{-\alpha} \sum_{j=0}^m C_j^\alpha x(m-j) \quad (2)$$

where  $h$  is the integration step size and

$$C_j^\alpha \equiv (-1)^j \binom{\alpha}{j} \quad (3)$$

Clearly, when  $\alpha = 1$  equation (2) yields the backward-Euler approximation of the first order derivative, i.e.;  $\frac{dx(t)}{dt} = \frac{1}{h} (x(m) - x(m-1))$ . While for  $\alpha = 2$ , equation (2) yields  $\frac{dx(t)}{dt} = \frac{[x(m) - 2x(m-1) + x(m-2)]}{h^2}$ .

## 3. DISCRETE MODEL AND FILTER TRANSFER FUNCTION

Substituting from (2) into (1) yields a discrete difference equation for the fractional filter of the form:

$$h^{-\alpha} \sum_{j=0}^m C_j^\alpha x(m-j) = \lambda x(m) + bu(m) \quad (4)$$

It is very clear that as  $m \rightarrow \infty$  equation (4) gives the best approximation of the fractional system. Hence, the

dimension of the single-variable fractional system is infinite. However, since  $0 < \alpha < 1$  one may realize that as  $m$  becomes large (i.e.,  $m > 15$ ), the values of the coefficients,  $C_j^\alpha$ , reduce to zero. This leads to approximate the infinite dimensional single-variable fractional system with a finite dimensional one.

In order to understand the discrete system behavior, consider the homogenous part of system (4). The solution of the first samples of  $x(m)$  indicates a continuously increasing system response given by:

$$x(1) = \frac{C_1^\alpha}{(1-h^\alpha \lambda)} x(0)$$

$$x(2) = \left( \frac{(C_1^\alpha)^2}{(1-h^\alpha \lambda)^2} - \frac{C_2^\alpha}{(1-h^\alpha \lambda)} \right) x(0)$$

$$x(3) = \left( \frac{(C_1^\alpha)^3}{(1-h^\alpha \lambda)^3} - \frac{2C_1^\alpha C_2^\alpha}{(1-h^\alpha \lambda)^2} + \frac{C_3^\alpha}{(1-h^\alpha \lambda)} \right) x(0)$$

$$x(4) = \left( \frac{(C_1^\alpha)^4}{(1-h^\alpha \lambda)^4} - \frac{3(C_1^\alpha)^2 C_2^\alpha}{(1-h^\alpha \lambda)^3} + \frac{2C_1^\alpha C_3^\alpha}{(1-h^\alpha \lambda)^2} + \frac{(C_2^\alpha)^2}{(1-h^\alpha \lambda)^2} - \frac{C_4^\alpha}{(1-h^\alpha \lambda)} \right) x(0)$$

$$\vdots$$

Clearly, for system stability, the homogenous system performance must converge to a finite value. However, it is impractical to investigate the convergence of each individual sample. Thus, there is a need to develop a system transfer function for the IIR digital filter which enables one to investigate the stability and the frequency response of such systems.

Taking the z-transform of both sides for equation (4) and assuming that the system is at rest, the transfer function of the system is given by:

$$H(z) = \frac{X(z)}{U(z)} = \frac{h^\alpha b}{\left( \sum_{j=0}^m C_j^\alpha z^{-j} - h^\alpha \lambda \right)} \quad (5)$$

$$H(z) = \frac{X(z)}{U(z)} = \frac{h^\alpha b z^m}{\left( (1-h^\alpha \lambda)z^m + C_1^\alpha z^{m-1} + C_2^\alpha z^{m-2} + \dots + C_{m-1}^\alpha z + C_m^\alpha \right)}$$

Figure 1 represents the block diagram of a  $m^{\text{th}}$ -dimensional digital signal processor of a FrIIR filter of order  $\alpha$  where  $K_i = \frac{C_i^\alpha}{(1-h^\alpha \lambda)}$ ,  $i = 1, 2, \dots, m$ .

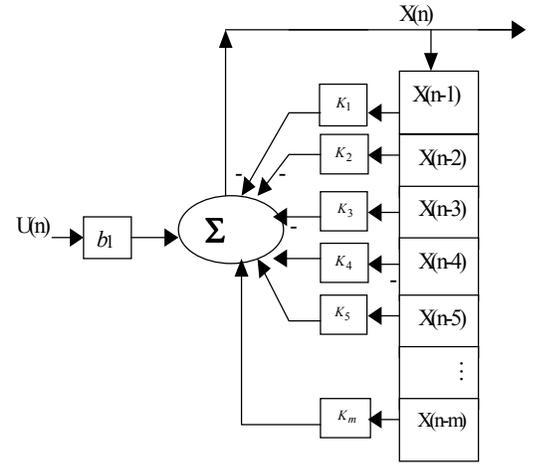


Figure 1: Linear digital signal processor of a one-variable FrIIR filter

**Theorem 1:** Consider the discrete system given by (4). The system is stable if and only if all roots of the characteristic equation,

$$\sum_{j=0}^m C_j^\alpha z^{-j} - h^\alpha \lambda = 0 \quad (6)$$

lie inside the unit circle.

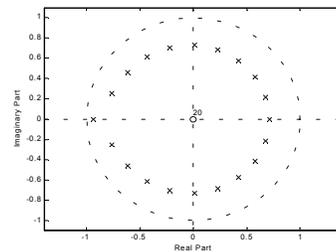
The stability of fractional continuous systems is discussed in [2,5]. The stability of the discrete fractional filters as a function of the fractional derivative is still open for further discussion.

#### 4. FINITE DIMENSIONAL MODEL OF FrIIR FILTERS

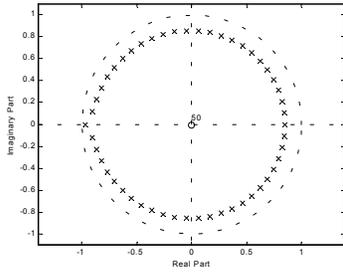
The following example illustrates the basic concepts of the finite dimensional modeling of Fractional infinite impulse response filters. The order of the model is subject to an error analysis that could be tolerated.

**Example 1:** Consider a single variable fractional filter described in (4) where  $\lambda = -1$  and  $\alpha = 1/2$ .

Figure 2 shows the poles and zeros of the characteristic equation (6) for two values of  $m$ . Clearly all the poles lie inside the unit circle for the given value of  $\lambda$  which represents a stable system.



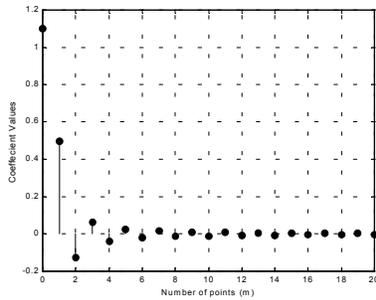
(a)



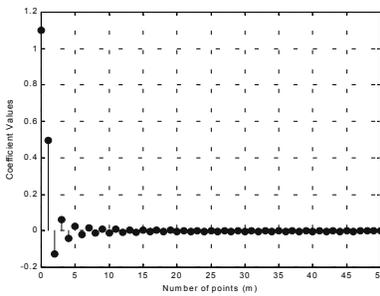
(b)

Figure 2: The root location of the fractional filter for a)  $m = 20$  points and b)  $m = 50$

Observe that the roots of the  $50^{th}$ -order model are of bigger magnitude than that of the  $20^{th}$ -order model. This implies that larger FrIIR model require more time to settle down than the smaller ones. On the other hand, the coefficient values of the characteristic equation for  $m = 20$  and  $m = 50$  points are shown in figure 3.



(a)



(b)

Figure 3: The coefficient values of a FrIIR filter with a)  $m = 20$  and b)  $m = 50$ .

Clearly the sum of the coefficient values of the characteristic equation (4) for  $m = 20$  is 1.51266718598854, while for  $m = 50$  it is equal to 1.51381765478557. The difference between the two sums is 0.00115046879703. Hence, if one considers only a  $10^{th}$ -order approximation, the difference in this case becomes 0.00388647192424 which indicates that as  $m$  gets larger,

the weighting coefficients due to adding more degrees of freedom to the model is of less effect. Therefore, one may approximate a single variable FrIIR filter with  $20^{th}$ -order IIR filter. If one accepts larger error, a  $10^{th}$ -order IIR may be sufficient.

The impulse response of a  $10^{th}$ -order and a  $50^{th}$ -order digital filters is shown in Figure 4. It shows that the impulse response is identical at the  $10^{th}$ -sample with 0% error as expected. The largest error occurs as we move to the right. For example, the impulse response error at the  $50^{th}$ -sample is 95.48%. A better comparison is between the impulse response of the two discrete models and the exact continuous model.

One should also realize that the size of the finite approximation for the FrIIR filter reduces with increasing  $\alpha$ , i.e; when  $\alpha = 1$ , a one-dimensional filter using the backward Euler approximation is sufficient in this case.

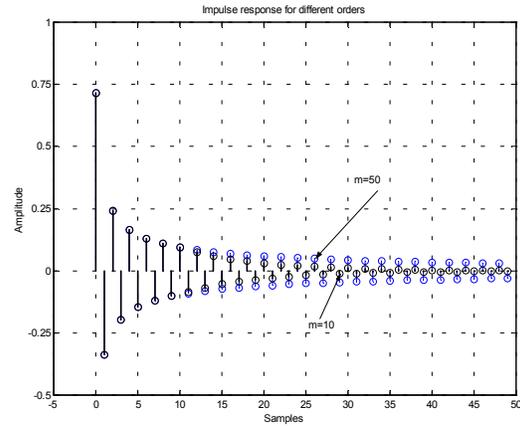


Figure 4: Impulse response of a finite dimensional FrIIF of orders a)  $m=10$  and b)  $m=50$ , respectively.

## 5. CONCLUSIONS

A finite model of a fractional infinite impulse response FrIIR filter is proposed. The same methodology will be applied to fractional filters of  $n$  variables. The stability results of theorem 1 derived in [5] can be extended for digital FrIIF. In fact the well known stability disk for digital IIR is deformed for the case of the digital FrIIF. It depends mainly on  $\alpha$  and left for further discussion. The real challenge is the implementation of such FrIIF with the knowledge of the high cost since each single variable fractional filter is approximated with a digital filter of order at least  $m > 20$ .

## 6. REFERENCES

- [1] Carpinteri, A. and Mainardi, F.(1997). Fractals and Fractional Calculus in Cotinuum Mechanics, *Spinger-Verlag*, Wein and New York.
- [2] El-Khazali, R. and Momani, S. (2001). Stability Analysis of Composite Fractional Systems, *ISCAS-IEEE Conf.*, Nov., 19-22, FL., USA.
- [3] Gorenflo, R. and Mainardi, F. (1996), “ Fractional Calculus: Integral and Differential Equations of Fractional Order, *CISM Course Scaling Laws and Fractality in Continuum Mechanics*, Lecture Notes, Udine, Italy, Sept., 23-27.
- [4] Hadid, S., Al-Shamani, J. (1986), Liapunov Stability of D.E. of non-integer order. *Arab J. Maths.*, No. 5, (Vol 1 & 2), pp. 5-17.
- [5] Matignon, D. (1996). Stability Results for Fractional Differential Equations with Applications to Control Processing. *Lille, France, IMACS, IEEE-SMC*, pp. 963-968.
- [6] Matignon, D. (1996), “Recent results in fractional differential systems theory,” Tech. Rep. 96 C 004, *Ecole National Superieure des Telecommunications*.
- [7] Miller, K.S., and Ross, B. (1993). *An introduction to the fractional calculus and fractional differential equations*, John Wiley & Sons.
- [8] Oldham, K.B. and Spanier, J.(1974). *The fractional calculus*, Acad. Press.
- [9] Podlubny, I. (1999). *Fractional Differential Equations*, Acad. Press.
- [10] W. Ahmad, R. El-Khazali and A.S. Elwakil, ‘ Fractional-Order Wien-Bridge Oscillator’ *IEE Electronic letters*, Vol., 37, No. 18, pp. 110-112.

# PERFORMANCE OF H-TERNARY LINE CODE IN THE PRESENCE OF NOISE: ANALYSIS AND COMPARISON

*Abdullatif Glass*  
*Department of Planning and Training*  
*Technical Studies Institute, P O Box 26833*  
*Abu-Dhabi, UAE*  
Tel: +971 2 5045937, Fax: +971 2 5854282  
e-mail: [aglass@ieee.org](mailto:aglass@ieee.org)

*Eesa Bastaki*  
*Department of Electrical Engineering*  
*UAE University, P.O. Box 17555*  
*Al-Ain, UAE*  
Tel: +971 3 7621703, Fax: +971 3 7616996  
e-mail: [eesa@uaeu.ac.ae](mailto:eesa@uaeu.ac.ae)

*Nidhal Abdulaziz*  
*Department of Electrical and Computer Systems Engineering*  
*Monash University, Clayton*  
*Victoria 3168, Australia*  
Tel: +61 3 99055350, Fax: +61 3 99055358  
e-mail: [nidhal.abdulaziz@eng.monash.edu.au](mailto:nidhal.abdulaziz@eng.monash.edu.au)

## KEYWORDS

Line codes, Data transmission, Probability of error, Noise performance modelling, Telecommunications.

## ABSTRACT

The performance of H-ternary line code is evaluated based on average probability of error computation for the received code when corrupted by additive white Gaussian noise. The calculation is performed using accurate statistical analysis of binary to ternary transition probabilities and encoder output symbol probabilities. The results show that the H-ternary line code performs better than the dicode however it slightly degrades in comparison with pseudo-ternary (AMI) code. Comparison between these line codes is also made based on some other desirable features.

## INTRODUCTION

The recent revolution in information technology requires the transmission of huge amount of data across telecommunication networks in the form of text, digitised speech and video, multimedia, etc. This imposes not only having telecommunication networks that support such increase in bandwidth requirement but also an efficient way to transmit this information. Thus, various encoding schemes have been proposed to this information/data in baseband and passband modulated schemes. In both cases, the main objective is to preserve important characteristics of the shape of the signal before being injected to the transmission channel. These characteristics include spectrum occupancy, timing information, power, transparency, ability to combat noise on the channel, and finally the cost and complexity of both encoder and decoder. Different line encoding schemes have been proposed and then deployed as standards in different parts

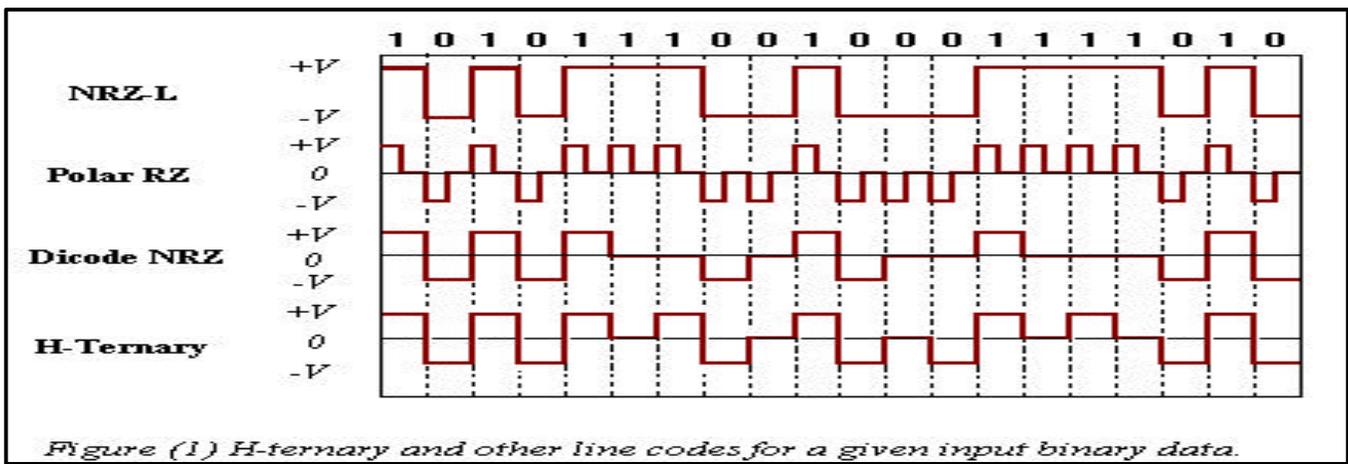
of the world. These include a generalised multilevel line code, where binary, ternary and quaternary line codes are considered as special cases where two, three and four levels are used for signal representation. Initial survey of these research works has been treated in (Bylanski 1976; Cattermole 1983; Lathi 1998), and the most recent surveys of these line codes are summarised in (Couch 2001; Dutton 1995; Takasaki 1991; Xiong 2000).

H-ternary line code is proposed based on exploiting the advantage of its predecessor such as dicode, polar RZ and NRZ-L codes. An introductory paper, which highlights the principle operation of the proposed line code, is covered in (Glass and Bastaki 2001). Encoder and decoder operation together with their simulation models has been treated in (Glass et al. 2001 and 2002). Signal transmission over band-limited and noisy channel leads to deformation of the received signal. This deformation is a result of many factors that exist on the channel such as inter-symbol interference, different sources of noise, fading, and channel bandwidth limitation. Calculation of signal to noise performance is a classical routine and may be found in the literatures. In our analysis only additive white Gaussian noise is considered in the performance evaluation of the line code under consideration.

In this paper, a brief introductory part is given in the next section which summaries the H-Ternary line code operation that we have proposed for data transmission. Section three provides mathematical analysis of the line code performance in the presence of noise. Comparison between the results that were obtained from the analysis with other line codes is given in section four. Finally, our discussion and conclusion are summarised in section five.

## ENCODING AND DECODING PRINCIPLES

The H-ternary line code operates on a hybrid principle that combines the binary NRZ-L, the ternary dicode and the polar RZ codes and thus it is called hybrid-ternary. The H-ternary code has three output levels for signal



representation; these are positive (+), zero (0), and negative (-). Figure 1 shows the waveforms of the H-ternary line code together, with other line codes that is derived from a given input binary data. The following subsections give description of the procedures for the encoding and decoding principles.

### Encoder Operation

The states table shown in Figure 2 depicts the encoding procedure. The H-ternary code has three output levels for signal representation; these are positive (+), zero (0), and negative (-). These three levels are represented by three states. The state of the line code could be in any of these three states. A transition takes place to the next state as a result of a binary input 1 or 0 and the encoder output present state. The encoding procedure is as follows:

- (1) The encoder produces + level when the input is a binary 1 and whether the encoder output present state is at 0 or - level.
- (2) The encoder produces - level when the input is a binary 0 and whether the encoder output present state is at 0 or + level.
- (3) The encoder produces 0 level when the input is binary 1 and the encoder present state is + level or when the input is binary 0 and the encoder present state is - level.
- (4) Initially, the encoder output present state is assumed at 0 level when the first binary bit arrives at the encoder input.

Input Binary	Output Ternary	
	Present State	Next State
1	0	+
1	-	+
1	+	0
0	-	0
0	0	-
0	+	-

Figure 2: Encoder output states table

The procedure gives the reader a sufficient information about the operation of this new line code scheme. Further details and comparison together with design and modelling of the encoder can be sought from references (Glass and

Bastaki 2001; Glass et al. 2001). The variation of this new line code is that it violates the encoding rule of NRZ-L and dicode when a sequence of 1s or 0s arrives. In the latter case, it operates on the same encoding rule of polar RZ but with full period pulse occupancy.

### Decoder Operation

Figure 3 shows the states table of the H-ternary decoding procedure. It is a reverse process of the encoding operation given in the previous subsection. The decoder has only two output states (binary) whereas the input is three states ternary. The decoding procedure is as follows (Glass and Bastaki 2001; Glass et al. 2002).

- (1) The decoder produces an output binary 1 when the input ternary is at + level and whether the decoder output present state is a binary 1 or 0.
- (2) The decoder also produces an output binary 1 when the input ternary is at 0 level and the decoder output present state is at a binary 1.
- (3) Similarly, the decoder produces an output binary 0 when the input ternary is at - level and whether the decoder output present state is a binary 0 or 1.
- (4) Finally, the decoder produces an output binary 0 when the input ternary is at 0 level and the decoder output present state is a binary 0.

It is clear that the decoding process at the receiver is quite similar to that of the NRZ-L code when the + and - levels are received. The difference arises when level 0 is received. In which case, the decision is made depending on the decoder output present state.

Input Ternary	Output Binary	
	Present State	Next State
+	1	1
+	0	1
0	1	1
0	0	0
-	0	0
-	1	0

Figure 3: Decoder output states table

### H-TERNARY PERFORMANCE EVALUATION

Performance evaluation for line codes can be categorised into many parameters. These parameters are mentioned in

the introduction and in many other literatures. In this section, the performance of the H-ternary line code in the presence of additive white Gaussian noise (AWGN) will be considered while other parameters will be discussed later in section five. This implies that the channel frequency response is a flat response with an infinite bandwidth. This assumption is not accurate for many practical channels however; it is reasonably accurate model as long as the transmitted signal bandwidth is much narrower than that of the channel.

A block diagram shown in Figure 4 represents the system under consideration, where the first and last blocks represent the encoder and decoder. The received signal at the detector is a three-level signal with AWGN added to it on the transmission channel. The first step in calculation of the average probability of error due to transmission channel noise is to calculate the probability of the transmitted signal levels (i.e. +, -, and 0) and their average energy.

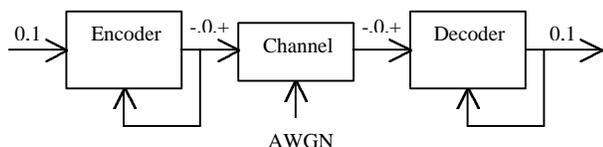
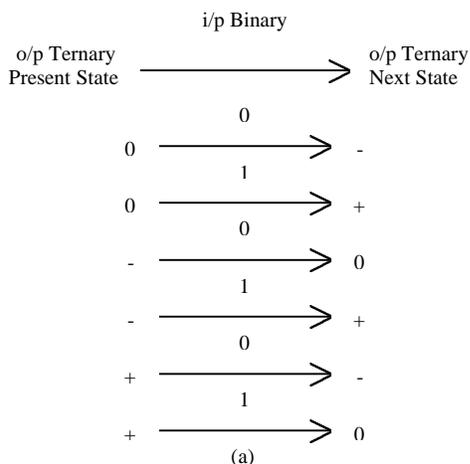


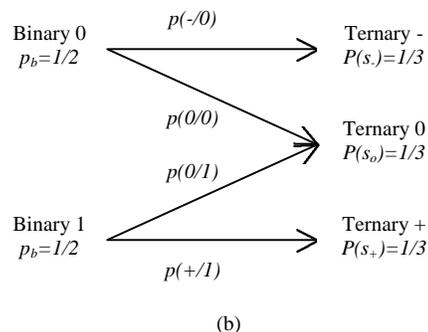
Figure 4: System block diagram

### Probability and Average Energy of H-Ternary Signal Levels

To develop a general formula for the average probability of error or bit error rate (BER) for the detected H-ternary signal, we need to consider the signal at the transmitter. The a priori probability of the transmitted signal levels (+, 0, -) depends upon the probability of binary to ternary signal encoding (conversion). Figure 5 shows the encoding principles at the transmitter, which has been derived from Figure 2.



(a)



(b)

Figure 5: Encoder input and output symbols and probabilities

The encoding process and the equally likely a priori inputs actually result in the calculation of the H-ternary signal probabilities. The encoding process demonstrated in Figure 5a has produced the transition probabilities of Figure 5b.

The following encoding process demonstrates the method of evaluating the transition probabilities and the final H-ternary signal levels probability. This shows that when the input is binary 0, the output possibilities of the (-) level is twice the possibilities of the (0) level. This means that transition probabilities are  $p(-/0) = \frac{2}{3}$  and  $p(0/0) = \frac{1}{3}$ . The same principle applies for the binary input 1, which results in output ternary level (+) twice as many as level (0). Thus,  $p(+/1) = \frac{2}{3}$  and  $p(0/1) = \frac{1}{3}$ . Thus, the final encoding probabilities from binary to H-ternary line code symbols shown in Figure 5b can be found through the following calculation.

$$P(s_+) = \frac{2}{3} p_b(1) = \frac{2}{3} (\frac{1}{2}) = \frac{1}{3}. \quad (1)$$

$$P(s_0) = \frac{1}{3} p_b(1) + \frac{1}{3} p_b(0) = \frac{1}{3} (\frac{1}{2}) + \frac{1}{3} (\frac{1}{2}) = \frac{1}{3}. \quad (2)$$

$$P(s_-) = \frac{2}{3} p_b(0) = \frac{2}{3} (\frac{1}{2}) = \frac{1}{3}. \quad (3)$$

Where  $p_b(\cdot)$  is the probability of binary input 0 or 1.

To calculate the average energy, consider the signal constellation given in Figure 6. The three transmitted signals are the +, 0, and - levels when  $s_+(t) = \sqrt{E}$ ,  $s_0(t) = 0$ , and  $s_-(t) = -\sqrt{E}$  respectively, where  $E$  is the transmitted symbol energy.

Thus, the average energy of the transmitted signal  $E_{avg}$ , given the probabilities of each transmitted signal level, is as follows.

$$E_{avg} = \sum_{i=1}^3 E_i P_i = EP(s_+) + 0 + EP(s_-) = \frac{2}{3} E. \quad (4)$$

Where  $E_i$  is the transmitted symbol energy given above and  $P_i$  is its probability given in equations (1-3).

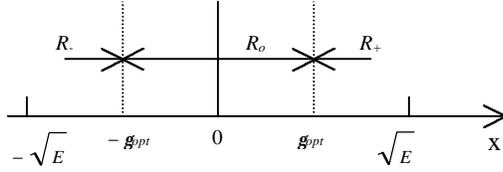


Figure 6: Signal constellation of H-ternary

### Probability of Error

The fidelity of the line code in combating channel noise is conveniently measured in terms of the bit error rate or probability of error. To calculate the probability of error we need to make use of the average transmitted signal energy and the probability of each level being transmitted. Figure 6 shows the three signal regions and their optimum decision levels. The optimum threshold level can be derived based on the principle that it should minimise the average error probability.

Since the transmitted signals are equally likely and the signal constellations are symmetric, the optimum threshold level is given as  $g_{opt} = \frac{\sqrt{E}}{2}$ .

The details of the average probability of error,  $P_e$  calculation is given in Appendix (A). The relation between the probability of error and correct decision,  $P_c$  is given by  $P_e = 1 - P_c$ . (5)

$$P_e = \frac{4}{3} Q\left(\sqrt{\frac{3}{4} \frac{E_{avg}}{N_o}}\right) \quad (6)$$

Where  $Q(\cdot)$  is the error function and  $N_o$  is the double-sided noise spectral density.

### RESULTS

The results calculated in section III are compared with other peer codes under consideration. These codes include the NRZ-L, dicode and AMI codes where their probabilities of error are given by (Couch 2001; Lathi 1998; Xiong 2000).

- (1) The NRZ-L code. This code is the ideal code where all other codes are referenced to and has the same probability of error as that of the binary phase shift keying (BPSK).

$$P_e = Q\left(\sqrt{\frac{2E_{avg}}{N_o}}\right), \quad (7)$$

where  $E_{avg} = E$ .

- (2) The AMI code.

$$P_e = \frac{1}{2} Q\left(\frac{2E_{avg}/N_o - \ln 2}{2\sqrt{E_{avg}/N_o}}\right) + Q\left(\frac{2E_{avg}/N_o + \ln 2}{2\sqrt{E_{avg}/N_o}}\right) \quad (8)$$

where in this case,  $E_{avg} = \frac{1}{2} E$ .

- (3) The dicode code has twice the probability of error of AMI.

$$P_e = 2P_e(AMI), \quad (9)$$

and  $E_{avg}$  is also the same as in AMI case.

Figure 7 shows the performance of the H-ternary line code in comparison with the peer line codes. It is evident that NRZ-L is performing the best among all codes and considered as reference for comparison. The H-ternary line code is performing better than dicode up to certain average signal to noise ratio. It is however shown that it degrades in performance in comparison to AMI code.

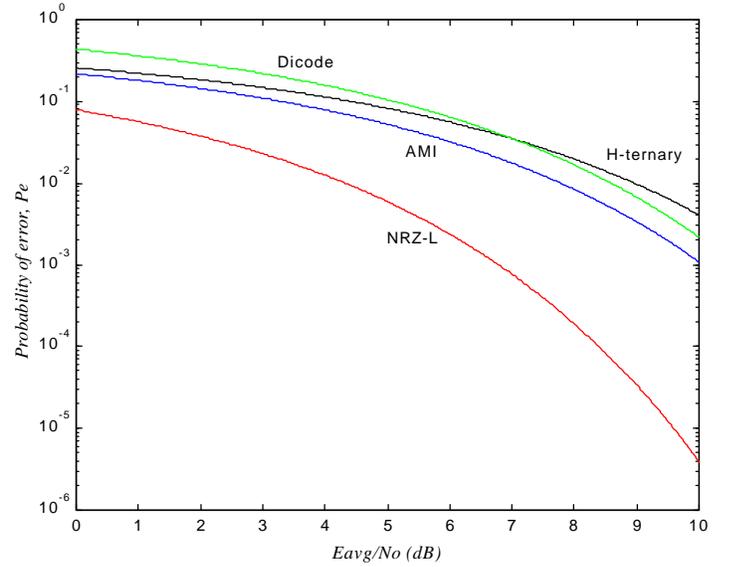


Figure 7: Performance comparison of line codes

### DISCUSSION AND CONCLUSION

Line codes are proposed and deployed into many applications to reshape binary pulses. The applications of line codes are not limited to telecommunication network transmission but they extend to digital recording systems and optical fibre applications.

The H-ternary line code performance in the presence of noise is determined using statistical analysis. The probability of error performance is calculated based on accurate analysis of the transition probabilities, the H-ternary symbol probabilities and the optimum signal level for detection. The H-ternary line code has shown improvement in performance over the dicode code for an average signal to noise ratio of less than 7 dB. It is also shown that AMI is performing better than H-ternary code. The H-ternary line code however has many other desirable characteristics when compared with these peer line codes under consideration. These are:

- (1) The H-ternary line code is a type of correlative codes. This property enables the code to detect single error at receiver.
- (2) It has better timing information content.

(3) Further its power spectral density concentrates around half the bit-rate transmission with no zero frequency (dc) component as with other line codes.

In conclusion, in this paper, H-ternary line code performance has been evaluated in the presence of AWGN. Its probability of error has been derived based on statistical analysis of the transmitted signal levels and its energy. The results obtained from the analysis are compared with other line codes under consideration. Applications and some other desirable features of these line codes are also discussed.

## REFERENCES

- Bylanski, P; and D. Ingram. 1976. *Digital transmission systems*. Peter Peregrinus.
- Cattermole, W. 1983. "Principles of digital line coding." *Int. J. Electron.*, Vol. 55, No. 1, pp. 3-33.
- Couch, L. W. 2001. *Digital and analog communication systems (3<sup>rd</sup> Ed)*. Pearson Education. pp.472-546.
- Dutton, H; and P. Lenhard. 1995. *High-speed networking technology: an introductory survey*. Prentice Hall PTR. pp. 2/1-2/51.
- Glass, A.M. and E. Bastaki. 2001. "H-Ternary line code for data transmission". *International Conference on Communications, Computer and Power (ICCCP'01)* (Sultan Qaboos University, Feb.12-14, Muscat, Oman) pp. 107-110.
- Glass, A., B. Ali, and E. Bastaki. 2001. "Design and modelling of H-Ternary line encoder for digital data transmission". *International Conference on Info-Tech & Info-Net (ICII2001)* (Beijing, China. Oct.29–Nov.1). pp.503-507.
- Glass, A., B. Ali, and N. Abdulaziz. 2002. "H-Ternary line decoder for digital data transmission: circuit design and modelling". *6<sup>th</sup> International Symposium on Digital Signal Processing for Communication Systems (DSPCS2002)* ( Sydney, Australia, Jan. 28-31). pp.149-153.
- Lathi, P. 1998. *Modern digital and analog communication systems (3<sup>rd</sup> Ed)*. Oxford University Press. pp.294-353.
- Takasaki, Y. 1991. *Digital transmission design and jitter analysis*. Artech House. pp.35-60.
- Xiong, F. 2000. *Digital modulation techniques*. Artech House. pp.17-83.

## APPENDIX (A)

### Error Probability Derivation

Using the signal constellation diagram shown in Figure 6, the analysis of the H-ternary line code average probability of error can be derived. The probability of correct decision,  $P_c$  of the line code is the summation of the correct decision probabilities of the random variable,  $X$  for each of the three received signal levels/symbols (+, 0, and -) within the regions  $R_+$ ,  $R_0$ , and  $R_-$ .

$$P_c = \Pr(X \in R_+ / s_+)P(s_+) + \Pr(X \in R_0 / s_0)P(s_0) + \Pr(X \in R_- / s_-)P(s_-). \quad (A1)$$

From the symmetry of the signal around the 0 level, the first and last terms of equation (A1) give the same result, and given the probability of each transmitted signal levels in section III.1 (equations 1-3), we have,

$$P_c = 2 \Pr(X \in R_+ / s_+) \frac{1}{3} + \Pr(X \in R_0 / s_0) \frac{1}{3} = 2 \Pr(X > \mathbf{g} / \sqrt{E}) \frac{1}{3} + 2 \Pr(0 \leq X \leq \mathbf{g} / 0) \frac{1}{3}. \quad (A2)$$

The probability of correct decision of the random variable  $X$  bounded by any of the above regions and given that particular symbol is transmitted can be rewritten in a form of  $Q$ -function (Couch 2001; Lathi 1998; Xiong 2000). Thus,

$$\Pr(X > \mathbf{g} / \sqrt{E}) = Q\left(\frac{\mathbf{g} - \sqrt{E}}{\sqrt{N_o/2}}\right) \quad (A3)$$

$$\Pr(0 \leq X \leq \mathbf{g} / 0) = Q\left(\frac{0-0}{\sqrt{N_o/2}}\right) - Q\left(\frac{\mathbf{g}-0}{\sqrt{N_o/2}}\right) = \frac{1}{2} - Q\left(\frac{\mathbf{g}}{\sqrt{N_o/2}}\right) \quad (A4)$$

In above equation the  $Q$ -function is defined as,

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt.$$

Substituting equations (A3 and A4) into (A2) and  $\mathbf{g} = \mathbf{g}_{opt} = \frac{\sqrt{E}}{2}$  we get

$$P_c = \frac{2}{3} \left[ Q\left(\frac{\frac{\sqrt{E}}{2} - \sqrt{E}}{\sqrt{N_o/2}}\right) + \frac{1}{2} - Q\left(\frac{\frac{\sqrt{E}}{2}}{\sqrt{N_o/2}}\right) \right] = \frac{2}{3} \left[ 1 - Q\left(\frac{\frac{\sqrt{E}}{2}}{\sqrt{N_o/2}}\right) + \frac{1}{2} - Q\left(\frac{\frac{\sqrt{E}}{2}}{\sqrt{N_o/2}}\right) \right] = \frac{2}{3} \left[ \frac{3}{2} - 2Q\left(\frac{\frac{\sqrt{E}}{2}}{\sqrt{N_o/2}}\right) \right] = 1 - \frac{4}{3} Q\left(\frac{\frac{\sqrt{E}}{2}}{\sqrt{N_o/2}}\right) = 1 - \frac{4}{3} Q\left(\sqrt{\frac{3 E_{avg}}{4 N_o}}\right) \quad (A5)$$

In above equation,  $Q(-x) = 1 - Q(x)$  and equation (4) (i.e.  $E = \frac{3}{2} E_{avg}$ ) are substituted.

Finally, substituting equation (A5) into equation (5) produces the final result.

$$P_e = \frac{4}{3} Q\left(\sqrt{\frac{3 E_{avg}}{4 N_o}}\right). \quad (A6)$$

# MODELING COMA MULTIPROCESSOR SYSTEMS BASED ON SCI

M. AL-Rousan  
American University Of Sharjah  
Computer Engineering  
Sharjah, UAE  
Email: [malrousan@aus.ac.ae](mailto:malrousan@aus.ac.ae)

A. Shatnawi  
S. Ahmed  
Jordan University of Science and Technology  
Computer Engineering Department  
Irbid, Jordan  
Email: [{ali,saba}@just.edu.jo](mailto:{ali,saba}@just.edu.jo)

## KEYWORDS

Modeling, COMA, SCI, Multiprocessors

## ABSTRACT

The Scalable Coherent Interface (SCI) provides high bandwidth, low latency communication. Hierarchical COMA systems generally suffer from high message latencies. Implementing the SCI communication protocol in hierarchical COMA systems reduces latencies and its cache coherence protocol will reduce directory sizes. In this paper, a hierarchical COMA system is modified in such a manner as to fit SCI. A new replacement policy where replaced lines swap storage places with new ones fetched in to the attraction memory is proposed. Similar to other COMA systems, this new system still gives better performance when part of the attraction memory is left unallocated.

## INTRODUCTION

Large scale shared memory multiprocessors offer a powerful and flexible computing environment. However, such systems may not offer optimal performance at all times for several reasons. The main reason is the increase in latencies when requests are serviced by remote nodes. Therefore, to improve performance the number of requests serviced remotely should be minimized. How this is done depends on the system used. In Non-Uniform Memory Access (NUMA) systems data replication and page migration are used (Heinrich et al. 1999, Stensrom et al. 1992). However, such methods usually depend on the data access patterns of the program for efficient operation. The initial distribution of data across the system also has a significant effect on the performance of such systems.

COMA systems use hardware techniques to replicate data in the main memories of nodes that need it most. This decreases remote access to data at the cost of hardware complexity and memory overhead. In addition to the cache coherence protocol used, COMA systems also need an efficient memory replacement algorithm. This algorithm must ensure that no data is accidentally lost from the system and at the same time not cause a significant increase in the total number of messages exchanged between nodes in the system. Increasing the number of messages usually increase network and memory contention and message latencies leading to performance degradation.

NUMA multiprocessor systems based on the IEEE Scalable Coherent Interface (SCI) have been proposed and studied in many works including (AL-Rousan et al. 2001, Stensrom et al. 1992). However, the impact of SCI on COMA systems have not gained enough attention. In this work we study the performance and implementation of COMA multiprocessors based on SCI

(IEEE 1999). Attention is focused on two issues. The first is how the SCI cache coherence protocol can be used in such systems. The second issue is the proposal of a new replacement policy which may be used in COMA systems when they are built based on SCI communication and cache coherence protocol.

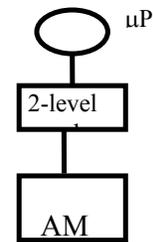


Figure 1: Processing Node Organization

Next, we discuss the general model used in this paper. Section 3 and 4 discuss the issues of implementing SCI in COMA systems and the architecture simulator. The results are discussed in Section 5 followed by the conclusions in Section 6.

## SYSTEM MODEL

In the COMA system under consideration, it is assumed that each processing node (referred to as node for simplicity) consists of a processor (CPU), a small primary cache, a larger secondary cache and a relatively large main memory called the *Attraction Memory (AM)*. A single node is shown in Figure 1. The complete system studied in this work, is constructed by interconnecting several of these nodes using either a single level or a multilevel hierarchy of rings as I will explain later in this section.

It is assumed that for every four nodes there is a directory. The four nodes along with their directory are connected to a single ring and they will be referred to as a subsystem or subtree. This directory knows what lines are stored in its subsystem and in which of the four nodes. It must be updated when a line moves from the AM of one node to another's. In addition to that, the directory is also responsible for locating a valid copy of a requested line and returning the data to the requester.

The basic interconnection topology used in this study is the ring and a hierarchy of rings may be used to connect the whole system. The number of levels in this hierarchy depends on the number of nodes and directories in the system. A system made up of four nodes and one directory is connected by one ring (see Figure 2). In larger systems (more than four nodes and several directories), rings are also used to connect the directories and enable communication between the subsystems. An eight node

system, made up of two subsystems in which the two directories are also connected by a ring, is shown in Figure 2. The interface between each node or directory and its ring is the standard SCI node proposed in the IEEE protocol (IEEE 1992).

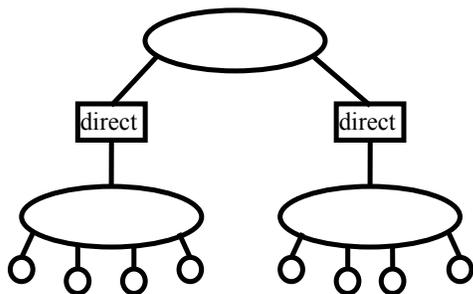


Figure 2: An 8-Node System With 2 Directories

### SCI IN COMA SYSTEMS

Both SCI communication and cache coherence protocols are implemented in the system under consideration. Implementation of the communication protocol is straightforward and no adjustments to the basic COMA system were needed. The reader may refer to the SCI standard for more details concerning this communication protocols (Al-rousan et al. 2001, IEEE 1992). However, the cache coherence protocol requires a few assumptions for efficient implementation. The details of such requirements are given in the following subsections.

### SCI Cache Coherence Protocol

SCI cache coherence protocol maintains the sharing set in the form of a double linked list which starts from the main memory (Al-rousan et al. 2001, Heinrich et al. 1999, IEEE 1992). In systems which support the concept of a home node, the sharing set of a line starts from the main memory of its home node.

Since hierarchical COMA systems do not support this concept, a node had to be chosen to start the sharing set from. This choice had to ensure that the resulting system still supports the basic concepts of hierarchical COMA systems. Therefore, this node could not be statically assigned to the line as in a Flat COMA (Stenstrom et al. 1992, Grujic et al. 1996, Dahlgren et al. 1997). So, terminology similar to that used in F-COMA is used here to name one copy of each memory line a master copy (MC) (Grujic et al. 1996). The sharing set for a line can start from the node which currently has the master copy in its AM. There are no restrictions on the storage of master copies and they can be stored in the attraction memory of any node in the system. However, the directories must always know where master copies are. So, when a master copy is moved from the AM of one node to another's, all directories responsible for both nodes are updated to reflect the new location of this master copy.

Directories in this system need only keep track of master copies of lines. They do not need to know where the other copies are located, because according to SCI data is supplied either from main memory or head of the sharing set. In this system, main memory refers to the current holder of the line's master copy. It would only waste memory space if directories stored the locations of all copies of a line. In addition to that, keeping track of multiple copies could sometimes lead to multiple

sharing sets for the same line and as a consequence requesting nodes might not always get a valid copy of the line.

### Replacement Policy in COMA

When a node receives new data in the size of a memory block, space in the AM must be allocated to store this data. Usually the replacement algorithm is invoked to select one of the existing memory blocks either to be overwritten or removed. Many replacement algorithms are discussed in (Cho et al. 1999, Torrellas and Padua 1999) which do not work straightforward for SCI. This led to the proposal of a new replacement algorithm which is used in the system to be studied. The main advantage of this algorithm over others, is that directories no longer have to search their subsystem(s) for a node to hold the removed block. Instead, they just update their information concerning the locations of master copies and forward the replacement message to its destination.

The algorithm works as follows. When a node receives a new block the AM is searched for a non master copy of any line. As soon as one is found, it is overwritten by the new block and no replacement messages are sent in this case.

If all copies stored in the AM are master copies, then a master copy with no sharing set is selected and must be sent to another node for storage. Instead of sending this MC to the directory, the removed MC swaps places with the new block. In this case the directory just receives an update message to reflect the change in the locations of master copies and is not responsible for finding a new storage place for the removed line. It then forwards this message to its target node. This message will be referred to as a *swap* message in future. In addition to source and target addresses, a swap message carries the removed data block and the address of the requested block so the target node can recognize for which of its blocks the swap message is. If an MC with a sharing set is selected for the swap operation, then another message to the head of the set is necessary. This message is to change the backward address pointer in the head's cache to point to the new holder of the line's MC. Therefore master copies with no sharing sets are selected for replacement. However, if all copies in the AM are master copies with sharing sets, then one of them must be selected for removal. The first option is to select an MC with a sharing set for which the head is the same node. In this case only a swap message is necessary and the backward address pointer is modified to hold the address of the new holder of the line's MC. This modification is performed without issuing any extra messages on the interconnection network. If the head of the sharing set is a remote node, then in addition to the swap message another message is issued on the interconnection network. This message is sent to the head to inform it that the location of the MC of the line has changed. This message is referred to as *change-MC-location*. This message just carries source and target addresses, the removed line's address and the address of the new holder of the removed line's MC. It carries no data. Upon receiving this message, the head just modifies the backward address pointer in its cache to hold the address of the node which now has the line's MC in its AM.

The replacement algorithm selects blocks for removal in the following order: (1) Non master copy is overwritten and no swapping occurs (2) Master copy with no sharing set swaps places with the new block (3) Master copy with a sharing set and the node itself is head of the set is swapped with the new block (4) Master copy with a sharing set and remote node is

head of the sharing set swaps places with the new block and the head is informed of this change.

Similar to read and write requests, a swap message might be forwarded to a node which no longer has the MC in its AM. The node responds to this in the same manner described concerning read and write requests. It sends a message to the directory to inform it that it no longer has the MC in its AM and the directory must locate the new holder of the line's MC and forward to it the swap message.

### Miss Types and Latencies

In the system under consideration, miss types are classified according to the way the system responds to them. As far as the cache memory is concerned, there are three types of misses; capacity, coherence and cold misses. A capacity miss is a miss due to the limited size of the cache and data is usually fetched from the AM. A coherence miss occurs when the data is in the cache but it is invalid and cannot be used. Coherence misses occur in systems which use a write-invalidate cache coherence protocol. Such misses cannot be serviced from the local AM. A cold miss results when a line is accessed for the first time. A cold miss may or may not be serviced from the local AM.

Misses in the cache memory may or may not be serviced from the local AM. While a miss in the AM can only be serviced by a remote node. Similar to a cache, a coherence miss in the AM occurs when data is in the AM but is invalid. A write miss or a read miss occurs if the node does not have the MC of the requested line in its AM. A write hit occurs when the node has the MC in its AM and all it needs to do is purge the line's sharing set. A read hit occurs when the node has the MC in its AM.

Since write hits, read hits, capacity misses and some of the cold misses can be serviced by the local AM, the latency they experience is just the memory access latency. While write misses, read misses, coherence misses and sometimes cold misses cannot be serviced locally and experience longer latencies. These latencies vary depending on the locations of the requester and responder in the network.

SCI defines a constant network latency of 1 cycle/ symbol ( 16 bits ) between two nodes on opposite ends of a link. This is the wiring latency or delay. The time a packet takes to reach its target depends on the number of links between its source and destination. It also depends on the packet length and the number of nodes on the way to the destination. For each node on the way, the packet experiences a 2 cycle decoding delay and a 1 cycle passing delay. The passing delay is the time taken to pass the packet to the bypass buffer. All delays used in this work are consistent with the work done in (Al-rousan et al. 2001, IEEE 1992).

In addition to the above mentioned network delays, the packet also experiences processing and queuing delays. Processing delays in the directories result from the time taken to search the directory memory for a certain line, the time taken to determine whether or not the packet's target node is within the directory's subsystem and the time needed to update a directory when master copies change locations.

### ARCHITECTURE SIMULATOR

We develop an SCI-based simulator to evaluate the performance of the COMA system under consideration. The simulated COMA system consists of a number of processing

nodes, with one processor per node. Each node has its own AM and a 2-level cache memory

All communications are based on the SCI communication protocol which is simulated in detail. Infinite lengths are assumed for the input queue, the output queue and the bypass buffer. Depending on the number of nodes, the simulator models one of the hierarchical networks shown in Figure 3.

The SCI cache coherence protocol using the typical set implementation option is considered in our simulator. Thus, a fixed memory and cache line size of 64 bytes is assumed. The sharing set for a line is implemented at the secondary cache level. This means that a processor which has a copy of the line in its secondary cache is a member of the line's sharing set irrelevant to whether or not the line is in the primary cache. A direct mapped primary cache and a 2-way set associative secondary cache with a random replacement policy are used.

Since the AM is organized as a large cache, a fully and a 4-way associative AM are simulated. The size of the memory includes 64 Kbytes AM, 16 Kbytes primary cache and a 128 Kbytes secondary cache. The AM size was chosen based on the fact that the whole system memory is large enough to hold the complete data set for each of the traces used. At the same time, a single AM unit is not large enough to hold a complete data set and AM misses will occur. The cache sizes were chosen relative to the AM size and to each other, such that primary cache misses are sometimes secondary cache hits and capacity misses still occur because neither of the caches is large enough to hold a complete data set. Other works (Stenstrom et al. 1992) use cache and memory sizes other than the ones assumed here. However, their assumption is also based on the data set size of traces they use and the fact that all types of misses should occur in the simulated system.

Five traces were used to evaluate system performance. These traces represent the addresses generated by a matrix addition, matrix multiplication, Gaussian elimination, quick sort algorithm and a binary search algorithm programs. However, these traces only specify the addresses referenced by these programs. They do not specify a read or a write operation for each address. Therefore, read and write operations were chosen randomly with a probability of 0.7 for a read operation and 0.3 for a write operation (Joe and Hennessy 1994).

The latencies associated with the processing nodes and directories in the system include cache, memory access, directory access, and update latencies which are assumed to be 1, 9, 30, 20 ns respectively. The latencies assume no contention on the local processing node busses and the bus transmission latencies are included in the given numbers. These are the latencies used by researchers in other studies on COMA systems (Joe and Hennessy 1994). The latencies due to the interconnection network and SCI nodes include wiring, decoding, passing, and switching delays which are assumed to be 2, 1, 1, 2 ns . These latencies are consistent with the latencies in (Al-rousan et al. 2001, IEEE 1992).

### RESULTS AND ANALYSIS

In this section, the results obtained for a 4-node system and an 8-node system will be discussed. These results were obtained using the five traces mentioned above. A comparison between the two systems will also be carried out. This work studies system performance based on the cache and memory hit and miss rates, execution times, and memory overhead and for 4-, 8- and 16-node systems

In the figures depicted in this section, cache hits represent read requests that hit in the primary and/or secondary caches. Capacity misses represent read requests that miss in the primary and secondary caches but hit in the local AM. Write hits represent write requests for lines whose master copies are in the local AM. These three hit and miss types represent read and write requests serviced by the local AM. Coherence misses represent read and write requests for lines found in the local AM but are invalid and must still be fetched from a remote node. AM misses represent read and write requests for lines not found in the primary cache, secondary cache and local AM and must be fetched from remote nodes. Cache hits, capacity misses and write hits represent requests serviced locally and their sum is referred to as the node hit rate. While, coherence misses and AM misses represent requests serviced by remote nodes and their sum is referred to as the node miss rate.

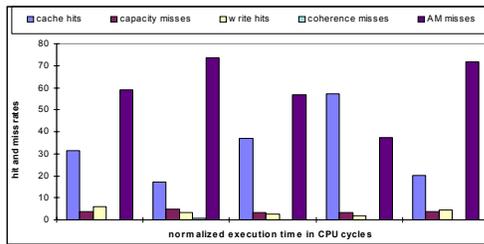


Figure 3: Hit And Miss Rates For a 4-Node System With 4-way Associative AM.

### Node Hit and Miss Rates

Figures 3 and 4 show the various hit and miss rates versus execution times for the systems for both the fully and 4-way associative AM. Cache hits depend mainly on the data access pattern of the application, the cache size and the mapping function used. In the system studied, a direct mapped primary cache and a 2-way associative secondary cache are used.

It can be seen that the node hit and miss rates are approximately equal for the matrix addition and quick sort applications in the three systems. While the other three applications experience varying hit and miss rates from one system to another. Since parameters such as the memory size and associativity are kept constant and only the number of nodes is changed, then this variation of hit and miss rates for some of the applications is caused by the change in the data distribution and the order in which nodes reference this data as their number is increased or decreased.

Some applications have their highest hit rate on an 8-node system, while others have their lowest hit rate on this system. This is also true for the 4-node system. The figures also show that the binary search application has the lowest execution time in a 4-node and an 8-node system. The matrix addition application has the lowest execution time on the 16-node system (not shown). The binary search application also experiences its highest miss rate on the 16-node system and its lowest miss rate on the 8-node system. This is due to the data distribution across the system and the order in which this data is referenced by nodes. In addition to the matrix addition and binary search applications, the matrix multiplication and gaussian elimination applications also experience this variation in execution times with respect to the other nodes for the same reasons.

Results for a fully associative AM are shown in Figures 4 and 7. The observations are not much different from those for a 4-way associative AM, except that the quick sort application is the only one with approximately equal hit and miss rates in the three systems. The other applications experience variations in their hit and miss rates as the number of nodes is increased or decreased for the same reasons explained earlier in the section.

Comparing the results in Figures 3 and 4 for a 4-node system, shows that for all applications and for a fully associative AM, the capacity miss rate, the write hit rate and the coherence miss rate are higher than those for a 4-way associative AM. While the AM miss rate is lower in a system with fully associative attraction memories than in a system with 4-way associative attraction memories. This agrees with the theoretical expectations for the hit and miss rates. This is also the case in an 8-node system in Figures 5.

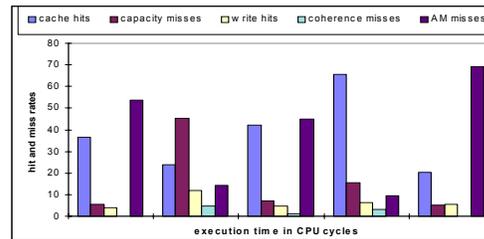


Figure 4: Hit And Miss Rates For a 4-Node System With Fully Associative AM

### Application Execution Times

The execution time of any application depends on many parameters. One of these parameters is message latencies in the system which usually depends on the network latency and contention. Execution times are studied by varying memory utilization and observing the effect of this variation on different application execution times.

Theoretically, an increase in memory utilization should lead to an increase in the application execution time. This increase results from the increase in replacement messages when utilization increases. An increase in the number of messages in the system increases network contention and message latencies leading to higher execution times for applications. These results are shown in Figures 7 and 8 for a 4-node and 8-node systems with a fully associative AM. The execution times of a 16-node system with fully associative AM behave in a manner similar to that of an 8-node system and is not shown here. Execution times for the three systems with a 4-way associative AM behave in a manner similar to that of a 4-node ring with a fully associative AM as utilization is increased and these graphs are also not shown.

Figure 7 shows the results for a 4-node system where the execution time increases as memory utilization is increased. There are cases where the execution time does not increase with utilization. This occurs at the same utilization where the replacement per miss ratio decreased instead of increasing which led to a similar decrease in execution times. Otherwise, the execution time increases with utilization. For the 8-node and 16-node systems, the execution times are approximately constant and only increase in a significant manner at high utilization (96% and above). This is also due to the replacement per miss ratio which is zero for utilization less than 100%. If there are no extra messages generated, then message latencies will remain approximately constant leading to constant

execution times. For a 4-way associative AM, the change in execution times with utilization is similar to that of a 4-node ring with fully associative AM.

### CONCLUSIONS

In this work a COMA system has been modified in a manner that enables implementation of the SCI protocols. These protocols are well defined and all details concerning their operation are specified. Therefore, the COMA system was modified to fit SCI instead of modifying SCI to fit COMA. The resulting system uses a hierarchy of rings for communications, thereby enabling the variation of the number of levels in this hierarchy to observe the effect on system performance. Three systems, obtained by varying the number of nodes and levels in the hierarchy, were studied. The three systems had the same memory size and executed the same applications.

Regarding node hit and miss rates, it may be concluded that these rates in a COMA system based on SCI are influenced by the initial data distribution especially when data is referenced by a node for the first time. As far as actual execution times are concerned, all applications execute faster on a 4-node system than they do on an 8-node or 16-node systems. This is due to the extra switching delay encountered in 8-node and 16-node systems. Hence, a multilevel hierarchy of rings can sometimes increase execution times, but a single ring can only connect up to a certain number of nodes after which it becomes more efficient and cost effective to use a hierarchy of rings.

### REFERENCES

Al-Rousan, M. J. K. Archibald and L. R. Beranson 2001. "Evaluating the Impact of Locality on Hierarchical Large-Scale SCI Multiprocessors" *Performance Evaluation Journal*, 46, (Nov), 275-302.

Cho, S., J. Kong, and Gyungho Lee 1999. "Coherence and Replacement Protocol of DICE - A Bus-Based COMA Multiprocessor", *Journal of Parallel and Distributed Computing*, 57, 14-32.

Dahlgren, F., Per Stenstrom, and Mårten Björklund 1997 "Reducing the Read-Miss Penalty for Flat COMA Protocols," *Computer Journal* 40, No.4, 208-219, 1997.

Heinrich, M., R. Soundararajan, J. Hennessy, and A. Gupta. 1999 "A Quantitative Analysis of the Performance and Scalability of Distributed Shared Memory Cache Coherence Protocols". *IEEE Transactions on Computers*, 48, No.2 (Feb), 205-217.

IEEE Standard 1596-1992. 1992, "The Scalable Coherent Interface".

Grujic, A., M. Tomasevic and V. Milutinovic 1996. "A Simulation Study of Hardware Oriented DSM approaches",

IEEE Parallel and Distributed Technology., 4, No. 1 (Spring), 74-83.

Joe, T. and J. L. Hennessy 1994. "Evaluating the Memory Overhead Required for COMA Architectures" In *Proceedings of the 21st International Symposium on Computer Architecture* (Chigaco, IL, April), IEEE Press, CA, 82-93.

Stenstrom, P., T. Joe and A. Gupta 1992, "Comparative Performance Evaluation of Cache Coherent NUMA and COMA Architectures" In *Proceedings of the 19th Annual International Symposium on Computer Architecture* (Queensland, Australia), ACM Press NY, 80-91.

Torrellas and D. Padua 1999. "The illinois aggressive coma multiprocessor project (i-acoma)." In *Proceedings of the 6th Symp. on the Frontiers of Massively Parallel Computing*, October 1996, Vol. 57, 14-32.

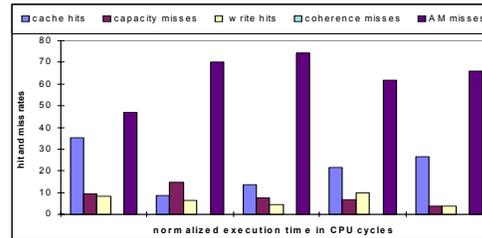


Figure 5: Hit And Miss Rates For An 8-Node System With 4-Way Associative AM

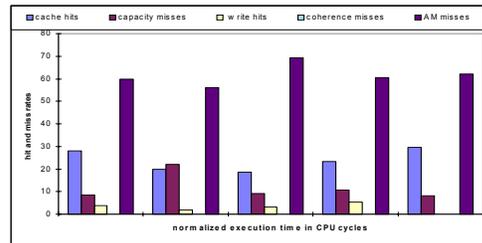


Figure 6: Hit And Miss Rates For An 8-Node System With Fully Associative AM.

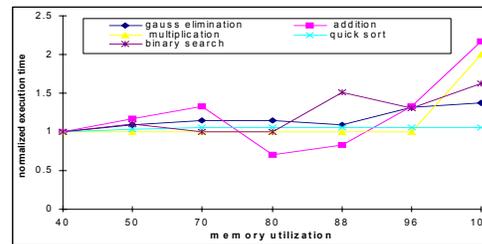


Figure 7: Normalized Execution For a 4-Node System With Fully Associative AM

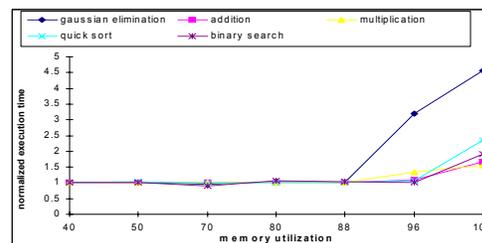


Figure 8: Normalized Execution For An 8-Node System With Fully Associative AM

# MODELLING AND SIMULATION OF COMMUNICATION BETWEEN DISTRIBUTED PROCESSES

Marc Störzel and Ursula Wellen  
Department of Computer Science  
University of Dortmund  
Baroper Str. 301, 44227 Dortmund  
Germany  
E-mail: {mstoerzel,wellen}@ls10.cs.uni-dortmund.de

## KEYWORDS

Simulation, Distributed Processes, Communication, Process Landscaping.

## ABSTRACT

This article presents an approach to modeling and simulating communication between distributed processes. The basic idea is to model processes in a way enabling us to simulate special communication features which have impact on the efficiency of processes taking place at different sites. Therefore, our initial emphasis is on the specification of communication interfaces between process parts distributed amongst different locations. We then use these specifications to simulate the behavior of distributed processes and to implement the software needed for analyzing the resulting simulation data.

The communication analysis is based on discrete simulation of Reference Nets, a special type of Petri Nets. It has been applied by simulating process models showing the distributed processes of a component-based software development project.

## INTRODUCTION

Traditionally, processes are modeled with focus on temporal and causal dependencies of the participating activities in order to understand and support the progress of (business and software) processes (Tully 1995), (Becker et al. 2000). We call this view of processes *logical view*, because it pays most attention to logical dependencies. Within the resulting process models, which we call collectively a process landscape, it is very difficult to recognize and analyze communication between different locations. But because of globalization, companies increase the distribution of their processes and therefore it becomes increasingly important to focus on the communication interfaces and consider them as first class entities while modeling and analyzing processes.

The distribution of processes to different locations requires a different view of a process landscape, which we call the *locational view* in order to differentiate from the terms *local view* and *distributed view* which are also used in the context of distributed processes. A local view of process models focuses on one part of the process landscape taking place at a single location, analogous to local views of distributed systems (Fischer et al. 2000). The term *distributed view* is often used for locally distributed systems, but also for

different views describing different aspects like organizational structures, products, resources or roles (Derniame et al. 1998).

In computer science, the term *communication* is first of all discussed in the context of software and hardware features enabling communication. Krumm for example considers the communication of distributed computer systems with focus on protocols ensuring secure and smooth data exchange between different systems (Krumm 1984). A general and up to date overview to the communication subject can be found e.g. in (Kilat and Lamersdorf 2001). When we consider communication between distributed processes, we mean precise information exchange between two or more activities where information should be available as document files. Simulation of distributed processes supports us to consider those communication features arising from the dynamic of processes' behavior. Therefore, the aforementioned technical communication aspects are not suitable for our research purpose.

In this paper, we apply the Process Landscaping method (Gruhn and Wellen 2000a), (Gruhn and Wellen 2000b) to focus on the communication within a distributed process landscape with respect to

- the amount of communication within a distributed process landscape,
- communication costs based on process dynamic,
- the efficiency of communication paths, and
- the efficiency of information distribution on demand

We do not discuss the Process Landscaping method here in detail, but we use its salient modeling features to develop a process landscape. These are:

- identification of the core activities within a process landscape and their positioning at a top level view,
- explicit modeling of interfaces between processes,
- switching between different levels of refinement, and
- extension by process model details only where needed

As underlying modeling notation we use Reference Nets (Kummer and Wienberg 2000), a special type of Petri Nets. They are based on Colored Petri Nets (Jensen 1997), supporting the analysis of individual, distinguishable information objects by simulation. The reason for preferring discrete instead of continuous simulation (Law and Kelton 1991) is our focus on the analysis of message exchange between distributed processes. Each receiving, processing and sending of a message is modeled as an event, affecting

the dynamic of the simulated Petri Net (Donzelli and Iazeolla 2001). The mechanism of message exchange is modeled with the synchronous channels of Reference Nets (Kummer 1998). The message processing is also modeled with Petri Nets at a more refined level. Thereby, participating activities are identified and modeled as transitions.

Section 2 of this paper gives an introduction to the communication features which are of interest. It discusses their impact on the efficiency of a process landscape and explains how they can be expressed and analyzed within a Petri Net-based simulation model. In section 3 we present first the structure of an example process landscape with respect to the underlying distribution (section 3.1). Afterwards, the main analysis results and some validation steps with a modified model are discussed (section 3.2). Section 4 summarizes our work and gives a prospect to our future research.

## COMMUNICATION FEATURES TO BE ANALYZED

In the introduction, we mentioned four communication features to be analyzed within a distributed process landscape. For this analysis purpose, we have to express them as attributes of process landscape elements and messages to be sent and processed. Before doing so, we define communication channels and communication systems as special landscape elements and discuss the meaning of the term *message* in our context:

- A **communication channel** is defined as abstraction of physical and logic systems enabling two communication partners to send or receive messages. It determines therefore the communication initiating and responding activity as well as the direction of the message interchange.
- A **communication system** is defined as a set of communication channels having equivalent attributes, like capacity or others.
- A **message** is defined as a document to be transported via a communication system for information exchange purposes. Furthermore, it has to state additional information like sender and receiver of the message content, the name of the communication system used, information about message volume, and an ID to identify messages on their ways within a process landscape.

With the terms defined above, we are now able to explain how to measure different communication features of a process landscape.

The varying **amount of communication** of different process parts within a process landscape is measured to evaluate the efficiency of the distribution of tasks. If some process parts are always very idle and others are waiting for them, this is an inefficient process dynamic. It increases communication costs by processes demanding results from idle ones. Answering the demanding processes also costs time and produces additional messages not necessary for the process execution itself. Therefore, the workload can be measured by the amount of demand messages and their answers.

The **communication costs based on process dynamic** can be measured by the extent of time and money for the effort of an activity to communicate with others. These two variables -

time and money - dependent on the distance between communicating activities (the shorter, the less expensive), the duration of each communication, and on the communication system's capacity. Data exchange via WAN (wide area network) with a dial-in connection is for example more expensive than via LAN (local area network) at a single location. Distance, duration and capacity can be used as variables of a charge schedule for the underlying communication system. This approach is equal to those of telecommunication companies.

A **communication path** within a process landscape is defined as the route an information object takes when it is sent from one process to another and is sent back after modification. Its efficiency is seen as minimal, if this way is very short: It let us suppose that the information object has not been modified very much before it has been sent back. This way of proceeding indicates a type of process organization which is only useful for controlling activities but not for activities producing something or handling with something. For the measurement of this communication feature we need information about the message itself, especially a type of identifier to follow the information object's path, and the sending and receiving activities. This information can be expressed as attributes of each message.

The fourth communication feature mentioned in the introduction considers the **information distribution on demand**. It is seen as efficient, if the effort of an activity for getting information is very low. For simulating such situations, we model information exchange as a process, started by the indication of an activity (called initiator) that it wants to communicate with another one. Only if the communication partner (called responder) answers this wish, can the information exchange start. The efficiency of this information distribution on demand can then be measured by the ratio of successful initiator trials to the non-successful ones.

Figure 1 summarizes the discussion of attributes necessary for communication analyses. It indicates which attributes are needed for the consideration of which communication features, and how these attributes are associated to different process landscape elements. By using Colored Petri Nets as underlying notation for a process landscape we express all message attributes as part of the token's signature. For communication channels and communication systems we define their attributes outside the Petri Net in an XML-file. Their values are merged with the simulation results to analyze communication amount and information distribution. Within the Petri Net structure, communication channels are modeled as two transitions connected by one place and two edges. A communication system is then modeled as a set of communication channels presented as a Petri Net on its own.

## EXAMPLE

In this section we discuss an example process landscape, where we analyze the aforementioned communication features. Subsection 3.1 introduces structure and content of the process landscape whereas subsection 3.2 discusses its simulation results.

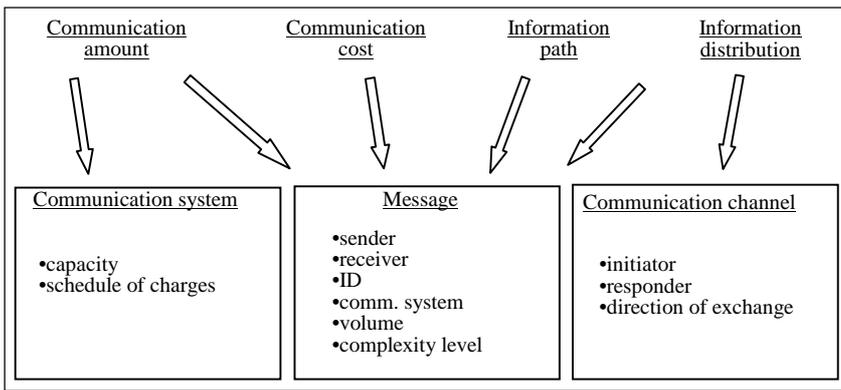


Figure 1: Relation between communication features and attributes

### Example structure

The example process landscape to be analyzed represents distributed processes concerning component-based software development. It includes implementation activities (application engineering and component engineering) as well as quality management, project management and domain engineering activities. The latter contains subactivities like process modeling and administration of component repositories for reuse. They are characteristic activities carried out in component-based software development. A further typical feature of component-based software development projects is the (distributed) coding of a set of software components which have to be integrated in the software application to be developed.

Figure 2 shows a high level locational view of the example process landscape. The gray cycle in figure 2 indicates the headquarters of a software development company, where application engineering *AE-Z*, project management *PM-Z*, domain engineering *DE-Z*, and parts of the quality management (*QM-Z*) as well as parts of component engineering (*CE-Z*) take place. Additionally, the company has one outpost A with component engineering and quality management activities and three outposts B, C and D with only component engineering activities. The bi-directional arrows indicate data exchange between the different locations and between different core activities independent of their location. The distance from the headquarters to outposts B, C and D is longer than to outpost A.

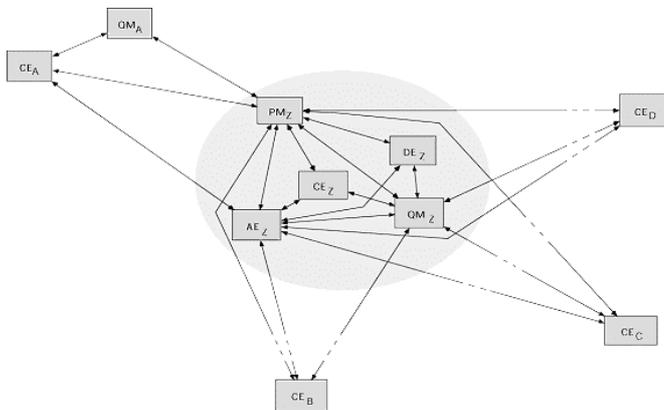


Figure 2: Locational view of the example software process landscape (Störzel 2001)

In figure 2, this is denoted by discontinuous arrows. The process landscape in figure 2 has been modeled with Reference Nets (Kummer and Wienberg 2000) as underlying notation. For simulation purposes the landscape has been split into 27 subnets. Each of which is representing either one core activity (at a single location) or a communication system with a set of similar attributed communication channels connecting two locations with each other. The whole process landscape consists of more than two hundred fine-grained activities and interfaces. It has been developed after interviewing several software

companies and considering domain specific literature (Allen and Frost 1998).

Figure 3 shows four abstract example subnets of the process landscape, representing one of the component engineering locations (gray rectangle) together with its interfaces to

project management, quality management and application engineering (gray rectangles with rounded edges). The interfaces are summarized in three different communication systems. Each circle within a communication system indicates a specific interface via which documents like the development order, guidelines or review reports are interchanged.

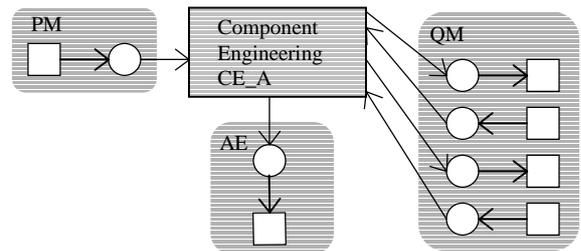


Figure 3: Abstract view of four subnets of the example process landscape

Figure 3 does not explain the detailed underlying Petri Net structure, because this would exceed the limits of this paper. It is just intended to give an idea of how the structure of the subnets is developed. For further modeling details we refer to (Störzel 2001).

### Analysis results

The communication analysis – as part of the Process Landscaping method – is based on discrete simulation of Petri Nets. With the parameterized landscape we started 35 simulation cycles, each consisting of approximately 3500 simulation steps. The parameters themselves range from small software development projects taking about 100 person-days to complex projects taking about 1000 person-days.

For each communication feature introduced in the previous section we now discuss the analysis results. It should be mentioned, that this section can only give a short overview of the most important results for the example process landscape. The algorithms used to come to these results and the

complete set of attribute values and simulation results are not discussed. A more detailed discussion can be found in (Störzel 2001).

Starting with the amount of communication we have to take into account that its costs are always assigned to the sending communication partner. Furthermore, the costs for short distance communication are less expensive than for long distances.

The costs measured for the example process landscape range from

- about 28 € for quality management outpost *QM-A*,
- about 10,000 € for each of the three component engineering outposts *CE-B*, *CE-C* and *CE-D*, and
- about 15,000 € for headquarters quality management *QM-Z*.

The high difference between the two quality management locations indicates possible improvements with regard to the division of responsibility: *QM-Z* is actually responsible for all review and testing activities taking place within *AE-Z* and *CE-B* to *CE-D*. However, *QM-A* is only responsible for review and testing activities concerning the local *CE-A*. A new set of simulation cycles with a modified process landscape, where *QM-A* attends to all four outposts and *QM-Z* to all headquarters processes, shows an improved amount of communication. Additional to the responsibility reorganization, the charge schedule has also been changed for *QM-A* due to its increased amount of communication and is now – for the new set of simulation cycles – less expensive for communication to other locations. The costs are now nearly zero for *QM-Z* because there is only internal communication left. Of course, the costs for *QM-A* increase dramatically, but at about 475 € they are still low-priced compared with the original costs for *QM-Z*.

When analyzing the communication costs based on process dynamic we have to take into account that processes waiting for others are also measured as occupied. Occupation itself is measured on a scale ranging from 0 to 100% where the percentage value shows the ratio of occupied time to simulation time. As expected, application engineering, *AE-Z*, is much occupied within a software development process (76.5%). It communicates with all other locations and core activities shown in figure 2. Additionally, the last activity within a software development process, the software release, is also modeled as part of application engineering. This implicates that the application engineering process is always occupied until the end of each software project. Complying to the first communication feature discussed, *QM-Z* is mostly occupied (85.9%), whereas *QM-A* is only occupied with about 10%. But parallel to the improved amount of communication, this inefficient occupation could also be improved to 64% for *QM-A* and 69% for *QM-Z* with the modified process landscape as basis for a second set of simulation cycles.

Values of path length have to be evaluated in correlation with all possible path lengths with respect to an information object within a process landscape. A precedent static analysis resulted in a range from one to five as possible path length. This scale is used for the evaluation of the simulation results. Low values can then be interpreted as

- either inefficient
- If only a few activities at a considered route treat the

passing information object very little, this indicates an inefficient distribution of tasks

- or the regarded process part consists of controlling activities or of decision making activities which does not change the information objects type but determines further process steps.

Simulation of the example process landscape resulted e.g. in an average path length of 3.5 within core activity *PM-Z* and 2.4 within *QM-A*. The value for the project management is relatively high, but does not indicate a flaw, because within this core activity there are a lot of controlling and decision activities like the ordering of components or estimation of project risks. Summarizing the simulation results for the efficiency of path lengths, this communication features are satisfying the regarded process landscape. Therefore, we do not discuss the further simulation cycles with the modified process landscape in this paper.

The efficiency of information distribution on demand is closely related to communication costs based on process dynamic. An occupied activity cannot answer to an information demanding communication partner at the same time. The more time an activity needs for its execution, the more information requests have to be answered with delay. While searching improvement possibilities within the modeled process landscape, mainly high and low ratios of successful to non-successful communication initiations are of interest. Here it is again *QM-Z* having a high ratio (more than 70%) when it is the initiator of a communication. If *QM-Z* plays the role of the responder, the ratio is very low (17%). This fits again to the high degree of occupation discussed in the context of communication costs based on process dynamic. Because *QM-Z* is responsible for most of the review and testing activities within the origin process landscape, all other activities have to wait for some time for their requested information like guidelines or review reports. Again, the modified process landscape shows improved values for the efficiency of information distribution on demand, especially for *QM-Z* as the responding communication partner.

## Evaluation of our approach

Completing the example discussion, we now consider the benefit of our approach from a more general point of view. The simulation of communication features within a distributed process landscape helps to reach one of the most important of process modeling purposes specified in (Fugetta 2000), namely process simulation in order to “evaluate possible problems, bottlenecks, and opportunities for improvement”. Applied to communication features we have focused on

- the identification of communication delays and
- the optimization of responsibility, task and information distribution

resulting in the reduction of communication costs and a more efficient workload. In the example, we could identify a location where a more low-priced schedule of charges was suitable, underloaded processes and activities with overwork where reorganization of tasks and responsibilities have improved the process landscape,

inefficient communication paths where the reorganization mentioned before could also improve the situation, and inefficient processes because of delayed responding of information demands.

A modified process landscape where a reorganization of tasks and responsibilities was implemented and a more low-priced schedule of charges was introduced for *QM-A*, served as basis for a new set of simulation cycles. Here we could check our improvement ideas, which indeed increased the communication efficiency and decreased the communication costs of the example process landscape.

## CONCLUSION

In this paper, we presented a new view on the communication subject by considering it in the context of a distributed software process landscape with focus on interfaces between processes. For this purpose, we introduced four special types of communication features which are in our opinion very important not only for software processes but for business processes in general (where software processes are seen as one type of business processes). Simulation of these features supports the improvement of communication efficiency within a process landscape and therefore helps to decrease communication costs. We applied our simulation approach to an example process landscape to make clear its benefit especially for locationally distributed processes. With a first set of simulation cycles we recognized the relations between the different communication features and were able to identify improvement possibilities within a process landscape. A second set of simulation cycles validated these improvements. Our future research focuses on simulation improvement. This means

- integration of evaluation support like the batch-mean method (Law and Kelton 1991) to improve the simulation results,
- extending the simulation parameters with respect to resources participating in a process landscape and
- extending tool support for both, simulation and evaluation of communication features in distributed process landscapes.

## REFERENCES

- Allen, P. and S. Frost. 1998. *Component-based Development for Enterprise Systems – Applying the Select Perspective*, Cambridge University Press.
- Becker J.; M. Rosemann; and C.v. Uthmann. 2000. "Guidelines of Business Process Modeling". In *Business Process Management – Models, Techniques and Empirical Studies*, W. v.d. Aalst, J. Desel and A. Oberweis (eds.), 30-49. Springer Verlag, appeared as Lecture Notes in Computer Science No. 1806.
- Derniame J.-C.; B. A. Kaba and D. Wastell. 1998. "The Software Process: Modelling and Technology". In *Software Process: Principles, Methodology and Technology*, J.-C. Derniame, B. A. Kaba, D. Wastell (eds.), 1-12. Springer Verlag, appeared as Lecture Notes in Computer Science No. 1500.
- Donzelli, P. and G. Iazeolla. 2001. "A hybrid software process simulation model". *Software Process – Improvement and Practice*, Vol. 6, No. 2 (Jun), 97-110.

- Fuggetta, A. 2000. "Software Process: A Roadmap". In *The Future of Software Engineering*, A. Finkelstein (ed.), ACM Press, 25-33.
- Fischer, I.; M. Koch and G. Taentzer. 2000. "Local Views on Distributed Systems and their Communications". In *Proceedings of the 6<sup>th</sup> International Workshop on Theory and Application of Graph Transformation (TAGT'98)* H. Ehrig; G. Engels; H.-J. Kreowski and G. Rozenberg (eds.), (Paderborn, Germany, Nov 1998), Springer Verlag, appeared as Lecture Notes in Computer Science No. 1764.
- Gruhn, V. and U. Wellen. 2000a. "Structuring Complex Software Processes by "Process Landscaping"". In *Proceedings of the 7<sup>th</sup> European Workshop on Software Process Technology, EWSPT 2000* (Kaprun, Austria, Feb). Reidar Conradi (ed.), 138-149, Springer Verlag, appeared as Lecture Notes in Computer Science No. 1780.
- Gruhn, V. and U. Wellen, 2000b. "Process Landscaping – Eine Methode zur Geschäftsprozessmodellierung" In *Wirtschaftsinformatik* Vol. 44, No. 4 (Aug), 297-309, in German.
- Jensen, K. 1997. *Coloured Petri Nets – Basic Concepts, Analysis Methods and Practical Use*, Volume 1, second edition, Springer Verlag.
- Kummer, O. 1998. "Simulating synchronous channels and net instances". In *5th Workshop Algorithmen und Werkzeuge für Petri-Netze* J. Desel; P. Kemper and E. Oberweis (eds.). Research Report No. 664, University of Dortmund, Department of Computer Science, Oct. 1998.
- Kummer, O. and F. Wienberg. 2000. "Renew - The Reference Net Workshop". *Petri Net Newsletter* No. 56, 12-16.
- Krumm, H. 1984. "Spezifikation, Implementierung und Verifikation von Kommunikationsdiensten für verteilte DV-Systeme". PhD thesis, University of Karlsruhe, Germany, June 1984, in German.
- Law, A.M. and W.D. Kelton. 1991 *Simulation, Modeling & Analysis*, second edition, McGraw-Hill.
- Killat, U. and W. Lamersdorf (Eds.). 2001. "Kommunikation in Verteilten Systemen (KiVS)", 12. Fachkonferenz der Gesellschaft für Informatik (GI) Fachgruppe "Kommunikation und verteilte Systeme" (KuVS), Feb 2001, Springer Verlag.
- Störzel, M. 2001. "Simulation verteilter Prozesslandschaften", Diplomarbeit at the University of Dortmund, Department of Computer Science, Software Technology, Oct 2001, in German.
- Tully, C. 1995. "The Software Process and the Modelling of Complex Systems" In *Proceedings of the 4<sup>th</sup> European Workshop, EWSPT'95*, (Noordwijkerhout, The Netherlands, Apr 1995). W. Schäfer (ed.), 138-143, Springer Verlag, appeared as Lecture Notes in Computer Science No. 913.

## BIOGRAPHY

**URSULA WELLEN** was born in Dortmund, studied Computer Science at the University of Dortmund and obtained her degrees in 1994. After working as consultant in computer industries, she went back to the University of Dortmund where she is currently writing on her PhDthesis.

**MARC STÖRZEL** was born also in Dortmund, went to the same University and obtained his degrees in Computer Science in 2001. Currently he's working at a Research and Development department for an automotive supplier.

# THE SIMULATION OF COMPRESSION FUNCTION PROCESSING FOR PETRA-256/2 CRYPTOGRAPHIC HASH FUNCTION

Mohannad Najjar and Janusz Stoklosa  
Institute of Control and Information Engineering  
Poznan University of Technology  
60-965 Poznan,  
Poland  
E-mail: Stoklosa@sk-kari.put.poznan.pl

## KEYWORDS

Model analysis, telecommunications, computer networks, cryptology, data integrity, hash function.

## ABSTRACT

In the paper we present a cryptographic hash function called Petra-256/2. When designing Petra-256/2 we formulated some goals, and one of them is especially important: processing an input message, the order of accessing source words should be data driven, i.e., the source words should be defined entirely before processing the input message and should be dependent on this message. We present the structure of the algorithm and results of simulation experiments on it. We have found the resulting structure probably free of collisions.

## INTRODUCTION

The expanding use of information, by sending or saving it, has stimulated a rapid development of effective tools to guarantee that data are not read, changed or damaged by those who do not have authorization to do so (Preneel 1999; Wayne 1996). To ensure that the data have not been tampered with, a special tool called a hash function should be applied. The hash function produces a value called a hash result that is appended to the data. The recipient of data regenerates the hash result and verifies that the two values agree. If so, it is persuaded that the data have not been modified. Such a use of the hash function confirms the integrity of data (Menezes 1997).

The purpose of this paper is to present the new hash function called Petra-256/2 (256 means the bitlength of the hash result and the number 2 – the version of the function) and simulation results in the designing process (Najjar 2002).

## CRYPTOGRAPHIC HASH FUNCTION

A hash function is defined as a computationally efficient algorithm (or function)  $h: M \rightarrow Y$  mapping binary strings of arbitrary length to binary strings of some fixed length, called hash results.

Apart from the property of compression and ease of computation a hash function should have additionally other potential properties (Menezes 1997; Preneel 1999):

- one-way – for almost all  $y \in Y$  it is computationally infeasible to find any input  $m \in M$  such that  $h(m)=y$ ;
- weak collision resistance – for a given  $m$  it is computationally infeasible to find a second input  $m'$  such that  $h(m)=h(m')$ ;
- strong collision resistance – it is computationally infeasible to find any two distinct inputs  $m$  and  $m'$  such that  $h(m)=h(m')$ .

If the hash function is one-way and weak or strong collision resistant then it can be used in real situations, e.g. as an element of digital signature or in construction of digital cash (Najjar and Stoklosa 2001; Wayne 1996).

The purpose of the paper is to present the simulation results of the compression function for a new cryptographic hash function called Petra-256/2.

In designing Petra-256/2 hash function we formulated some goals which must be achieved. These goals are security, simplicity, compactness, and speed.

We also formulate some detailed goals (Najjar 2002):

- processing an input message, the order of accessing source words should be data driven (in, e.g., MD4, MD5, SHA-1 and RIPEMD-160 it is fixed), i.e., the source words should be defined entirely before processing the input message and should be dependent on this message,
- following operations should be used in the structure of compression blocks: selected homogeneous perfect nonlinear functions, bitwise exclusive-or, addition modulo  $2^{32}$ , multiplication modulo  $2^{32}+1$  and rotation of bits.

Thus, our model is the modified version of this used in MD4 and MD5 hash functions (Menezes 1997).

## STRUCTURE OF PETRA-256/2 HASH FUNCTION

A perfect nonlinear  $n$ -argument Boolean function is a function, for which the distance from the set of all affine functions is maximum, i.e., it is equal to  $2^{n-1}-2^{(n/2)-1}$  (Najjar and Stoklosa 1999).

An  $n$ -argument Boolean function is called homogeneous of degree  $k$  if the algebraic normal form of the function is as follows:

$$g(x_1, x_2, x_3, \dots, x_n) = c_0 \oplus c_1 x_1 \oplus c_2 x_2 \oplus \dots \oplus c_n x_n \oplus c_{12} x_1 x_2 \oplus c_{13} x_1 x_3 \oplus \dots \oplus c_{(n-1)n} x_{n-1} x_n \oplus c_{123} x_1 x_2 x_3 \oplus \dots \oplus c_{12\dots n} x_1 x_2 \dots x_n,$$

where  $\oplus$  is the bitwise exclusive-or operation,  $c_i \in \{0, 1\}$ , and  $c_i = 0$  for every conjunction in which the number of variables is different from  $k$  (Najjar 2002; Qu et al. 1999).

Let  $\parallel$  denote the concatenation operation of words and let  $0^r$  be the  $r$ -fold concatenation of 0.

The hash function transformation of the message  $m = m_1 \parallel m_2 \parallel \dots \parallel m_t$  divided into fixed length blocks  $m_1, m_2, \dots, m_t$  can be described as follows:

$$\begin{aligned} H_0 &= IV, \\ H_i &= \varphi(m_i, H_{i-1}) \text{ for } i = 1, 2, \dots, t, \\ h(m) &= H_t, \end{aligned}$$

where  $IV$  is an initial value,  $H_i$  is a chaining variable and  $\varphi$  is a compression function. As a result we obtain a hash result  $h(m)$  of fixed length. The length of processed block  $m_i$  is equal to 512 bits.

The processing diagram of Petra-256/2 hash function is presented in Figure 1. A message  $m$  of the length  $L_m$  is

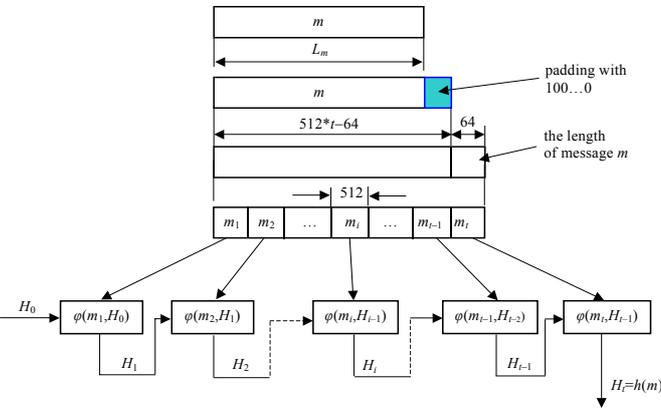


Figure 1: Processing diagram of the hash function

padded with the help of the word  $1 \parallel 0^k$  for an integer  $k \geq 0$  and 64 bits representing the length of the message  $m$ . After padding we have the modified message  $m^* = m_1 \parallel m_2 \parallel \dots \parallel m_t$ , for a positive integer  $t \geq 1$ , where the 16 words of the length equal to 32 bits each, i.e.,  $m_i = s_1 \parallel s_2 \parallel \dots \parallel s_{16}$ . These words are processed in pseudorandom order in 16 steps in each of 3 rounds.

Each word  $H_i$  (for  $i = 0, 1, \dots, t$ ) is considered to be the

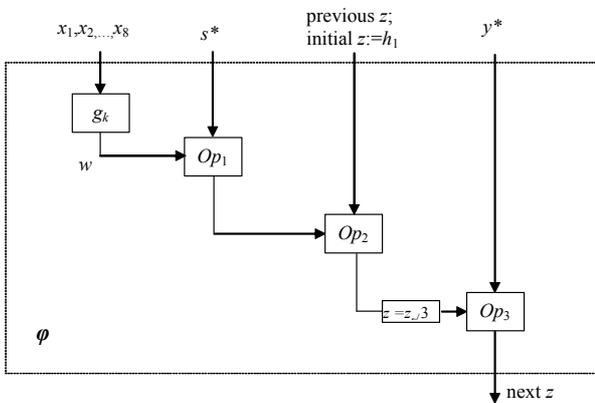


Figure 2: General structure of compression block for Petra-256/2 (one step in a round)

concatenation of 32-bit words  $A, B, C, D, E, F, G, I$ , the chaining variable  $H_i = A \parallel B \parallel C \parallel D \parallel E \parallel F \parallel G \parallel I$ . The words  $A, B, C, D, E, F, G, I$  are arguments  $x_1, x_2, \dots, x_8$ , of a homogeneous Boolean function  $g_k$  which is perfect nonlinear (see Figure 2); as a result we have a 32-bit word  $w = g(x_1, x_2, \dots, x_8)$ . The resulting hash  $H_i$  also depends on a 32-bit constant  $y^* \in \{y_1, y_2, \dots, y_{16}\}$ , on the word  $s^* \in \{s_1, s_2, \dots, s_{16}\}$ , and on the value  $z$ .

In the structure of compression block the following operations are used:

- three perfect nonlinear Boolean functions of degree 2, one per round:
$$\begin{aligned} g_1(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) &= \\ & x_1x_2 \oplus x_1x_3 \oplus x_1x_4 \oplus x_1x_6 \oplus x_1x_7 \oplus x_1x_8 \oplus x_2x_3 \oplus x_2x_4 \oplus \\ & x_2x_5 \oplus x_2x_6 \oplus x_2x_7 \oplus x_2x_8 \oplus x_3x_6 \oplus x_3x_7 \oplus x_3x_8 \oplus x_4x_8 \oplus x_5x_6 \oplus x_5x_7 \oplus \\ & x_6x_8 \oplus x_7x_8, \\ g_2(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) &= \\ & x_1x_2 \oplus x_1x_4 \oplus x_1x_5 \oplus x_1x_6 \oplus x_1x_7 \oplus x_2x_3 \oplus x_2x_5 \oplus x_2x_6 \oplus \\ & x_2x_7 \oplus x_3x_4 \oplus x_3x_7 \oplus x_4x_5 \oplus x_4x_8 \oplus x_5x_8, \\ g_3(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) &= \\ & x_1x_8 \oplus x_2x_3 \oplus x_2x_4 \oplus x_2x_8 \oplus x_3x_5 \oplus x_4x_6 \oplus x_5x_6 \oplus x_5x_7 \oplus x_7x_8. \end{aligned}$$
- bitwise exclusive-or, addition modulo  $2^{32}$  and rotation left through 3 bits (depicted as  $\oplus$ ,  $\boxplus$  and  $\ll 3$ , respectively). Let  $Op_1, Op_2, Op_3 \in \{\boxplus, \oplus\}$ , where  $\oplus$  is used twice.

The main goal in the assigning process is to point out the order of operations in the general structure of the compression block and to choose the permutation of the words  $A, B, C, D, E, F, G, I$  mapped one-to-one onto  $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$ .

## ASSUMPTIONS FOR SIMULATION EXPERIMENTS

To choose the order of operations and the permutation we performed simulation experiments which consisted in changing succeeding single input bit (of all 512 bits which form the word  $m_i$ ) and counting how many of 256 output bits were changed. The following quality criterion of the compression block was used:

$$L = \min (|l_{mi} - s| + |l_{mx} - s|), \text{ for } s \text{ such that } 122 \leq s \leq 134,$$

where  $l_{mi}$  is the minimum number of changed bits,  $l_{mx}$  is the maximum number of changed bits,  $s$  is the average number of changed bits and minimum  $\min$  is taken over all 512 changes of single input bits. All possibilities were tested:  $3!$  orders of operations and  $8!$  permutations ( $3! \cdot 8! = 241920$  possibilities).

All permutations of the  $n$ -element set are generated in antilexicographical order. We use it to find all orderings of operations and also all permutations of the set  $\{1, 2, 3, 4, 5, 6, 7, 8\}$  modeled the words  $A, B, C, D, E, F, G, I$  mapped onto variables  $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$ .

Let the indices 0, 1, 2 represent three orderings of operations ( $Op_1, Op_2, Op_3$ ) as follows:

- 0: ( $\boxplus, \oplus, \oplus$ )
- 1: ( $\oplus, \boxplus, \oplus$ )
- 2: ( $\oplus, \oplus, \boxplus$ )

Let the indices 0, 1, ..., 40319 represent the following permutations:

0: (1, 2, 3, 4, 5, 6, 7, 8)  
1: (2, 1, 3, 4, 5, 6, 7, 8)  
2: (1, 3, 2, 4, 5, 6, 7, 8)  
3: (3, 1, 2, 4, 5, 6, 7, 8)  
4: (2, 3, 1, 4, 5, 6, 7, 8)  
.....  
40319: (8, 7, 6, 5, 4, 3, 2, 1)

In the sequel, by *NoOper* we denote a number from the set of indices {0, 1, 2} representing three orderings of operations ( $Op_1, Op_2, Op_3$ ) and by *NoPerm* – we denote a number from the set of indices {0, 1, ..., 40319} representing one of the above permutations.

### GENERAL ALGORITHM FOR PETRA-256/2

Let *Round*=3 denote the number of rounds, *Step*=16 be the number of steps per round. *TabPerm* is the permutation table, and *TabOper* – the table of operations.

#### Algorithm 1 – Petra-256/2

INPUT:  $y, m, h_1, h_2, \dots, h_8, g_k, Round, Step, NoPerm, NoOper$ .

OUTPUT:  $TabA[0]||TabA[1]||\dots||TabA[7]$ .

METHOD:

1. {Initial processing}
  - $h_1 := DAA4361D_H,$
  - $h_2 := B16EC431_H,$
  - $h_3 := 615BE562_H,$
  - $h_4 := 98256F2E_H,$
  - $h_5 := E565B322_H,$
  - $h_6 := 98D8EB18_H,$
  - $h_7 := 53A0EF98_H,$
  - $h_8 := 6EA9D519_H.$
  - $TabA[0]:=h_1, TabA[1]:=h_2, TabA[2]:=h_3, TabA[3]:=h_4,$
  - $TabA[4]:=h_5, TabA[5]:=h_6, TabA[6]:=h_7, TabA[7]:=h_8,$
  - $A^* := TabA[0], B^* := TabA[1], C^* := TabA[2],$
  - $D^* := TabA[3], E^* := TabA[4], F^* := TabA[5],$
  - $G^* := TabA[6], I^* := TabA[7],$
  - $y_1 := 4CC59886_H, y_9 := CA0ECF98_H,$
  - $y_2 := 33166219_H, y_{10} := 29F1A748_H,$
  - $y_3 := F31941D9_H, y_{11} := EEB268AA_H,$
  - $y_4 := E9053E34_H, y_{12} := 4822D4C0_H,$
  - $y_5 := 155DD64D_H, y_{13} := 4766EA27_H,$
  - $y_6 := 9809045A_H, y_{14} := 864CC598_H,$
  - $y_7 := 44E8ECDD_H, y_{15} := 767CC650_H,$
  - $y_8 := B310C998_H, y_{16} := 8D3A414F_H.$
  - $x := TabA[0]$
2. for  $l = 1$  to  $r$  do
  - for  $k = 1$  to *Round* do
    - a. for  $m_l := s_1||s_2||\dots||s_{16}$  do  $m_l^* := s_1^*||s_2^*||\dots||s_{16}^*$ , where  $s_j^*$  is chosen pseudorandomly from the set  $\{s_1, s_2, \dots, s_{16}\}, s_j^* := S_{q_k[j]}$
    - b.  $y_j^* := y_j'$ , where  $y_j'$  is chosen pseudorandomly from the set  $\{y_1, y_2, \dots, y_{16}\}, y_j^* := y_{q_k[j]}$
    - c. {Compression and permutation}
      - for  $j = 1$  to *Step* do

$z := \varphi(NoOper, g_k, TabA[0], \dots, TabA[7], x, s_j^*, y_j^*),$   
 $A := TabA[TabPerm[NoPerm,0]-1],$   
 $B := TabA[TabPerm[NoPerm,1]-1],$   
 $C := TabA[TabPerm[NoPerm,2]-1],$   
 $D := TabA[TabPerm[NoPerm,3]-1],$   
 $E := TabA[TabPerm[NoPerm,4]-1],$   
 $F := TabA[TabPerm[NoPerm,5]-1],$   
 $G := TabA[TabPerm[NoPerm,6]-1],$   
 $I := TabA[TabPerm[NoPerm,7]-1],$   
 $x := z,$   
 $TabA[0]:=A, TabA[1]:=B, TabA[2] :=C,$   
 $TabA[3]:=D, TabA[4]:=E, TabA[5]:=F,$   
 $TabA[6]:=G, TabA[7]:=I,$

- d.  $TabA[0]:=A \oplus A^*, TabA[1]:=B \oplus B^*,$   
 $TabA[2]:=C \oplus C^*, TabA[3]:=D \oplus D^*,$   
 $TabA[4]:=E \oplus E^*, TabA[5]:=F \oplus F^*,$   
 $TabA[6]:=G \oplus G^*, TabA[7]:=I \oplus I^*.$

The way of choosing the best order of operations  $Op_i$  and permutations for Petra-256/2 is presented in the next algorithm. As previously we assume that the number of rounds is equal to three, *Round*=3. We also assume that  $w \gg n$  denotes the word  $w$  with the  $n$ -th bit from the left negated, and  $w_1 \neg w_2$  denotes the number of bits in  $w_1$  different from  $w_2$  ( $w_1$  and  $w_2$  must have the same length).

#### Algorithm 2 – Generation of the structure of compression block for Petra-256/2

INPUT:  $m^*$ .

OUTPUT:  $TabR$ .

METHOD:

1.  $i := 0, j := 0,$
2. Pad the message  $m^*$  with the word 100...0 and the word representing its length;  $x := m^*$ ,
3. Generate all possible orders of operations and permutations,
4. Let  $hh^*$  be the output of Algorithm 1 for  $y, x, h_1, h_2, \dots, h_8, g_k, Round:=3, Step:=16, i, j, ii:=0,$
5.  $ii := ii+1, x := m^* \gg ii,$
6. Let  $hh$  be the output of Algorithm 1 for  $y, x, h_1, h_2, \dots, h_8, g_k, Round:=3, Step:=16, i, j,$
7.  $p[i, ii] := hh^* \neg hh,$
8. if  $ii < 513$  then go to 5,
9.  $l_{mi} := \min(p), l_{mx} := \max(p), s := \text{average}(p),$
10. if not  $122 \leq s \leq 134$  then go to 13,
11.  $L = \min(|l_{mi}-s| + |l_{mx}-s|),$
12.  $TabR[i, j] := L,$
13. if  $j < 40319$  then  $\{j := j+1 \text{ go to } 4\},$
14. if  $i < 5$  then  $\{i := i+1, j := 0, \text{ go to } 4\}.$

### RESULTS OF EXPERIMENTS

The best ordering of operations and permutations for Petra-256/2 was done by means of Algorithm 2.

For the next steps in the designing process we chose ordering with *NoOper* = 0 and the permutation with *NoPerm* = 5943 indicating the order of operations and permutations in the compression block (see the grayed row in Table 1, and also Figure 3).

Table 1: The best orderings of operations and permutations for Petra-256/2

Ordering	Permutation	$L$
0	5943	41
1	14123	44
0	5921	45
0	5926	45
1	14061	46
2	19173	46
1	14075	47
2	19153	47
2	19174	47
0	5914	48
0	5932	48
1	14058	48
1	14083	48
0	5931	49
1	14080	49
1	14084	49
2	19157	57

During the tests it was observed that permutation plays more important role than the order of operations.

Operations	Permutation	$l_{mi}$	$l_{mx}$	$s$	$L$
$Op_1 = \boxplus$ $Op_2 = \oplus$ $Op_3 = \oplus$	(8, 1, 4, 2, 3, 5, 6, 7)	106	147	127	41

Figure 3: The best structure of the compression block and its parameters

For the chosen result additional tests were performed with about  $2^{20}$  random input texts of the length changing from 0 to 20. For each text we computed the average  $s$ . The minimum average we received equals 127 and the maximum average is equal to 128. This confirmed the pertinence of our selection.

Three kinds of tests were performed only for Petra-256/2 for the first round with the purpose to find collisions:

- The first test was conducted for all words of the length from 0 to 20; they were stored and all pairs were checked. No collision was found.
- The second test relied on random generation of 1000 inputs and processing each input  $in$   $2^{20}$  times in such a way that  $in := h(in)$  trying to find collisions. The results were negative.
- In the third test of memoryless Yuval's birthday attack was implemented (Menezes 1997). We used Algorithm 3. No collision was found.

#### Algorithm 3 – Memoryless Yuval's birthday attack

INPUT: Petra-256/2.

OUTPUT: messages  $m_1$  and  $m_2$  such that  $h(m_1) = h(m_2)$ .

METHOD:

1. Generate randomly two  $n$ -bit messages  $m_1$  and  $m_2$ .
2.  $H := h(m_1)$ .

3. If the  $H$ 's least significant bit equals 1 then  $i:=1$ , if it equals 0 then  $i:=2$ .
4. Modify  $m_i$  in the following way:
  - a.  $H := H || 0^{n-256}$ ,
  - b.  $m_i := m_i \oplus H$ ,
5. Compute  $h(m_1)$  and  $h(m_2)$ .
6. Is  $h(m_1) = h(m_2)$ ? If so, increase the collision counter, if not  $H := h(m_i)$  and go to 3.

In order to estimate the speed of Petra-256/2 its processing rate was measured. On average the processing rate is 15.39 Mbits/sec. It is satisfactory in comparison with the other hash functions tested in the same environment (Najjar 2002).

## REFERENCES

- Menezes A. J., van Oorschot P.C., Vanstone S. A. 1997. "Handbook of Applied Cryptography". CRC Press, Boca Raton, FL.
- Najjar M. 2002. "Petra Family of Cryptographic Hash Functions". Ph.D. dissertation. Department of Electrical Engineering, Poznan University of Technology, Poznan, Poland.
- Najjar M., Stoklosa J. 1999. "The nonlinearity of homogeneous Boolean functions and the design of strong cryptographic algorithms". In *Proceedings of the Regional Conference on Military Communication and Information Systems. CIS Solutions for an Enlarged NATO*. Burakowski, A. Wiczorek (Eds.), Zegrze, Poland, vol. 2, 71–76.
- Najjar M., Stoklosa J. 2001. "Cryptographic hash functions in digital economy". In *Information Systems Architecture and Technology ISAT 2001. Proceedings of the 23rd International Scientific School 'Digital Economy Concepts, Tools and Applications'*. Grzech A., Wilimowska Z. (Eds.), Wroclaw University of Technology Press, Wroclaw, Poland, 140–146.
- Preneel B. 1999. "The state of the cryptographic hash functions". In *Lectures on Data Security. Modern Cryptology in Theory and Practice*. Damgård I. (Ed.), LNCS 1561, Springer, Berlin, 158–182.
- Qu C., Seberry J., Pieprzyk J. 1999. "On the symmetric properties of homogeneous bent functions". In *Information Security and Privacy*. Pieprzyk J., Safavi-Naini R., Seberry J. (Eds.). LNCS 1587, Springer, Berlin, 26–35.
- Wayner P. 1996. "Digital Cash". AP Professional, Boston.

## AUTHORS' BIOGRAPHIES

**MOHANNAD NAJJAR** was born in Amman, Jordan. After finishing high school he won a scholarship for the Ministry of Education of Jordan to learn in Poland. In 1998 he defended his MSc. in computer networks, and in 2002 a Ph. D. thesis in cryptology, both at Poznan University of Technology.

**JANUSZ STOKLOSA** is a professor in computing sciences at Poznan University of Technology, Poznan, Poland. He is the author of numerous books, "Algebraic and Structural Automata Theory" (North-Holland, 1991, co-author), "Cryptographic Algorithms" (in Polish, 1994), "Security of Data in Information Systems" (in Polish, 2001, co-author) being most important of them.

# SIMULATION OF THE RATE-DISTORTION BEHAVIOUR OF A MEMORYLESS LAPLACIAN SOURCE

*Nasir M. Rajpoot*

Department of Computer Science  
University of Warwick  
Coventry, CV4 7AL  
United Kingdom

email: [nasir@dcs.warwick.ac.uk](mailto:nasir@dcs.warwick.ac.uk)

## ABSTRACT

In the context of transform coding of still images, the distribution of high frequency transform coefficients has been approximated by a Laplacian distribution. In this short paper, an efficient way of simulating the rate-distortion behaviour of a memoryless Laplacian source is presented. Simulation results show that the estimated behaviour very closely matches the empirical one.

## 1. INTRODUCTION

In transform coding, it has been established that the statistics of individual high-frequency subbands (wavelet or wavelet packet) and non-DC discrete cosine transform (DCT) coefficients can best be approximated by a Laplacian distribution, for most natural images. Earlier research [5, 2] on the DCT transform coding schemes employed the intuitive assumption that the DC and the non-dc DCT coefficients for images are more likely to follow Gaussian and Laplacian distributions respectively. Later on, these assumptions were experimentally verified to be valid upto a reasonable extent by the goodness-of-fit tests using  $l_\infty$ -norm as a distance measure [3] between both the distributions: the actual one and the one under consideration. Some researchers have also proposed generalized Gaussian distribution (GGD) for both DCT and wavelet coefficients in image coding applications [4]. However, it was shown recently, in [1], that the GGD assumption performs better than the Laplacian one by no more than 0.08 bits/pixel using a mean square error (MSE) distortion measure.

In this short paper, analytical expressions are developed for rate (entropy) and distortion (MSE) resulting from the uniform scalar quantization of data generated by a memoryless Laplacian source. This is helpful in simulating the

rate-distortion behaviour of a memoryless Laplacian source without actually doing the quantization.

## 2. THE RATE-DISTORTION RELATIONSHIP

In this section, we will derive a relationship between rate and distortion, as defined earlier in the previous section, for a memoryless Laplacian source (with zero mean) which may provide the governing distribution for non-DC DCT or high-frequency wavelet transform coefficients. Let  $p(x)$  denote the probability of occurrence, also known as the probability density function (pdf), of the source value  $x$  given by

$$p(x) = \frac{\lambda}{2} e^{-\lambda|x|} \quad (1)$$

where  $\lambda$  is the distribution parameter related directly to the source variance  $\sigma^2$  by  $\lambda = \sqrt{2/\sigma^2}$ . In order to investigate the distortion-rate function for this distribution, let us suppose that the uniform step size  $\Delta$  produces  $2n + 1$  equally spaced points  $(x_{-n}, \dots, x_{-1}, x_0, x_1, \dots, x_n)$ ;  $x_i = i\Delta$ , yielding the distortion  $D_\delta$  and requiring  $R_\delta$  bits/sample, at the least, for encoding. The quantities  $R_\delta$  and  $D_\delta$  can thus be expressed as follows,

$$R_\delta = - \sum_i p'(x_i) \log_2 p'(x_i) \quad (2)$$

and

$$D_\delta = \sum_i d_i, \quad (3)$$

where

$$p'(x_i) = \int_{a_i}^{b_i} p(x) dx \quad (4)$$

is a new pdf for the quantized source data, and

$$d_i = \int_{a_i}^{b_i} p(x)(x - x_i)^2 dx \quad (5)$$

---

Part of this work was completed while the author was visiting Department of Mathematics, Yale University, USA.

is the distortion due to mapping of the source data values  $x \in (a_i, b_i)$  to  $x_i$ , with  $a_i = x_i - \frac{\Delta}{2}$  and  $b_i = x_i + \frac{\Delta}{2}$ . The new pdf  $p'(x_i)$  for quantized coefficients given in equation (4) can be simplified as follows,

$$p'(x_i) = \frac{2 \sinh \frac{\lambda \Delta}{2}}{\lambda} \cdot p(x_i) \quad (6)$$

The above equation can be derived by integrating (4). It shows that  $p'(x_i)$  is directly proportional to  $p(x_i)$ , with the value of constant increasing monotonically with the step size  $\Delta$ . The obvious fact that  $p(x_i) = p'(x_i)$  for small values of  $\Delta$  is also clearly decipherable from this equation. Now putting the value of  $p(x)$  from (1) in (5), we obtain

$$d_i = \frac{\lambda}{2} \int_{a_i}^{b_i} (x - x_i)^2 e^{-\lambda|x|} dx; \quad \forall i, i \neq 0$$

and

$$d_0 = \lambda \int_0^{\frac{\Delta}{2}} x^2 e^{-\lambda x} dx.$$

Integration of the right hand sides of above equations yields the following expression for  $d_i$  and  $d_0$

$$d_i = \left\{ \frac{\Delta^2}{4} + \frac{2}{\lambda^2} - \frac{\Delta}{\lambda} \coth \frac{\lambda \Delta}{2} \right\} p'(x_i); \quad \forall i, i \neq 0 \quad (7)$$

and

$$d_0 = \frac{2}{\lambda^2} - \left( \frac{\Delta^2}{4} + \frac{2}{\lambda^2} + \frac{\Delta}{\lambda} \right) e^{-\frac{\lambda \Delta}{2}}. \quad (8)$$

We also know that  $p'(x_0) = p'(0) = 2 \int_0^{\Delta/2} p(x) dx = 1 - e^{-\frac{\lambda \Delta}{2}}$  and so

$$\sum_{i; i \neq 0} p'(x_i) = 1 - p'(0) = e^{-\frac{\lambda \Delta}{2}}.$$

Replacing  $d_i$  in (3) with the expressions in (7) and (8), we obtain the following expression for total distortion

$$D_\delta = \frac{2}{\lambda^2} - \frac{\Delta}{\lambda} \left( 1 + \coth \frac{\lambda \Delta}{2} \right) e^{-\frac{\lambda \Delta}{2}} \quad (9)$$

The above equation confirms the fact that the distortion  $D_\delta$  approaches towards the source variance  $\sigma^2 = 2/\lambda^2$  as the step  $\Delta$  increases. As for the rate  $R_\delta$ , (2) can be re-written as

$$R_\delta = -p'(0) \log_2 p'(0) - \sum_{i; i \neq 0} \log_2 \left( \sinh \frac{\lambda \Delta}{2} e^{-\lambda|x_i|} \right) p'(x_i)$$

or

$$R_\delta = -p'(0) \log_2 p'(0) - e^{-\frac{\lambda \Delta}{2}} \log_2 \sinh \frac{\lambda \Delta}{2} + \frac{\lambda}{\log 2} S \quad (10)$$

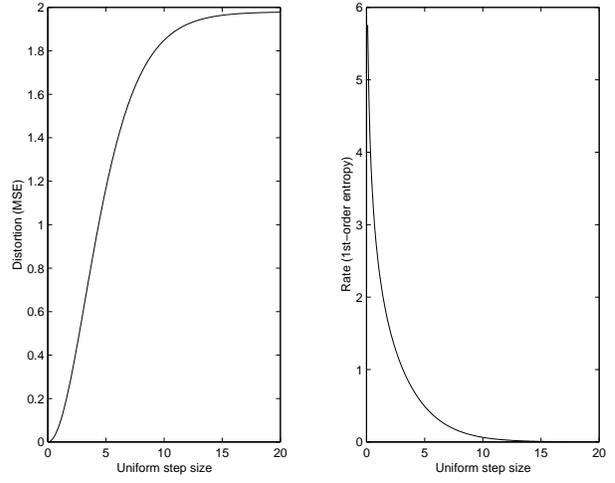


Figure 1: Typical plots of  $D_\delta$  and  $R_\delta$  versus  $\Delta$

where  $p'(0) = 1 - e^{-\lambda \Delta/2}$  and  $S$  is given by

$$S = \sum_i |x_i| \cdot p'(x_i) = 2\Delta \cdot \sinh \frac{\lambda \Delta}{2} \sum_{i=0}^n i e^{-i\lambda \Delta} \quad (11)$$

From (10), it is clear that as the step size  $\Delta$  increases, both  $e^{-\lambda \Delta/2}$  and  $S$  approach towards zero and thus  $R_\delta$  has more of the contribution to its value from the source values  $x \in [-\frac{\Delta}{2}, \frac{\Delta}{2}]$ , which are quantized to zero, than for smaller values of  $\Delta$ . The typical curves for  $D_\delta$  and  $R_\delta$  when plotted against  $\Delta$  (for  $\lambda = 1$ ) are shown in Figure 1. As is clear from their corresponding expressions,  $R_\delta \rightarrow 0$  and  $D_\delta \rightarrow \sigma^2$  as the step size  $\Delta$  becomes very large.

### 3. SIMULATION RESULTS

Equations (9) and (10) provide us a fast way of approximating the  $(R_\delta, D_\delta)$  points of the rate-distortion curve for a given data which is assumed to follow a Laplacian distribution. The only parameter that we need to know about this data is its variance  $\sigma^2$  (with  $\lambda$  given by  $\lambda = \sqrt{2/\sigma^2}$ ).

We generated data sets of various sizes coming from a memoryless Laplacian source with  $\lambda = 1$ . The operational rate-distortion curve was obtained by simply quantizing the data uniformly with different step size  $\Delta$  values and then computing the corresponding values for distortion (MSE) and rate (order-1 entropy). To obtain a fast approximation to the operational rate-distortion curve, (9) and (10) were used to compute the  $(R_\delta, D_\delta)$  points. Both the empirical and the approximated operational curves for rate-distortion for the number of data points  $N = 4K$  and  $N = 16K$  are shown in Figures 2 and 3 respectively.

## 4. CONCLUSIONS

A fast method of approximating the rate-distortion curve of a memoryless Laplacian source is presented. The simulation results verify the validity of expressions (9) and (10). An important problem in data compression where this may be helpful is the following: Given fixed rate and/or distortion value, find a fixed step size ( $\Delta$ ) to quantize the transform coefficients such that the given rate and/or distortion value are achieved. However, finding an analytical solution for this problem does not appear to be straightforward and needs some attention.

## 5. REFERENCES

- [1] K. A. Birney and T. R. Fischer, *On the modeling of DCT and subband image data for compression*, IEEE Trans. on Image Processing **4**, 2 (1995), 186–193.
- [2] H. Murakami, Y. Hatori, and H. Yamamoto, *Comparison between DPCM and Hadamard transform coding in the composite coding of the NTSC color TV signal*, IEEE Trans. on Communications **COM-30** (1982), 469–479.
- [3] R. C. Reininger and J. D. Gibson, *Distributions of the two-dimensional DCT coefficients for images*, IEEE Trans. on Communications **COM-31**, 6 (1983), 835–839.
- [4] N. Tanabe and N. Farvardin, *Subband image coding using entropy-coded quantization over noisy channels*, (1995), *Computer Science Technical Report Series*, University of Maryland, College Park, MD.
- [5] A. G. Tescher, *Advances in Electronics and Electronic Physics* (New York), Academic, 1979, pp. 113–115.

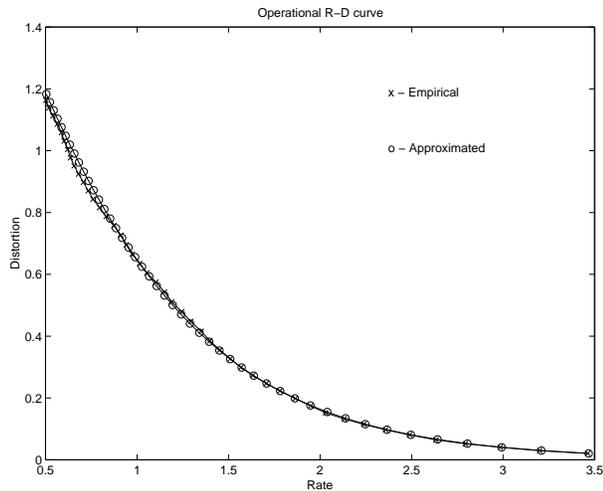


Figure 2: Empirical and approximated operational rate-distortion curves for a Laplacian data with  $\lambda = 1$  and  $N = 4K$

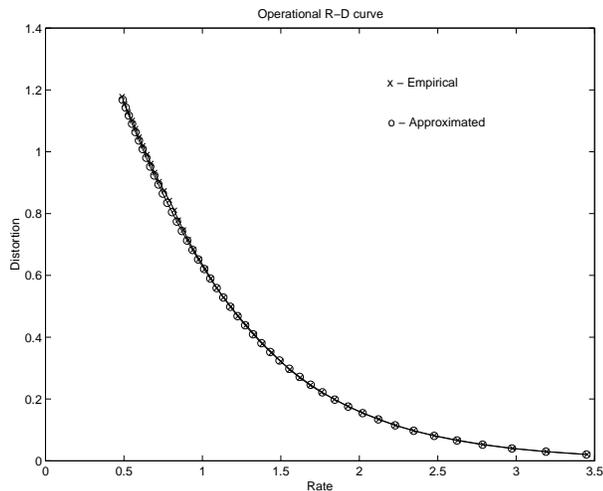


Figure 3: Empirical and approximated operational rate-distortion curves for a Laplacian data with  $\lambda = 1$  and  $N = 16K$

# Performance Evaluation of MQAM in Fading Channels with Diversity

O. Alani, M. Al-Akaidi, D Alnsour

School of Engineering and Technology  
De Montfort University, Leicester  
LE1 9BH, UK  
Email: mma@dmu.ac.uk

## Abstract

The radio spectrum is a finite resource and it is important that all users exploit it efficiently. Consequently modulation scheme used for mobile environment should utilise the RF channel bandwidth and the transmitted power as efficiently as possible.

The introduction of adaptive MQAM (M-ary Quadrature Amplitude Modulation) has enhanced the performance of the communication system as shown in [4]. The impact of using diversity techniques in such systems is investigated. The gain, obtained from utilizing this combination of adaptive MQAM and diversity techniques in the system is shown.

## 1 Introduction

In real world wireless communication systems, the link channels are not simple AWGN (additive white Gaussian noise), in fact they are characterised by different kinds of fading due to environmental factors and interference from other users that adversely affect signals propagate in these channels. Therefore we cannot rely on the formula set by Shannon that relates the capacity of channel (hence spectral efficiency) to SNR (signal to noise ratio) since this relation is meant for AWGN channels. LEE [1] derived a relation which gives the average channel capacity of flat fading channel and it's obtained by averaging the capacity of AWGN channel,  $C = W \log_2(1 + \gamma)$ , over the distribution of the received SIR (signal to interference ratio),  $\gamma$ .

To enhance the spectral efficiency, adaptive transmission schemes are used. The main idea about these schemes is real time balancing of the link budget through adaptive variation of the transmitted power level, symbol rate, constellation size, coding rate/scheme, or any combination of these parameters [2].

Adaptive transmission schemes provides much higher average spectral efficiency by taking advantage of the "time varying" nature of the propagation channel without sacrificing BER. Other fading compensation techniques (such as increased link budget margin or interleaving with channel coding) are designed for the worse case channel conditions which result in a poor utilisation of the full channel capacity for majority of the time (under negligible or shallow fading conditions) [2].

Shannon's channel capacity theorem states that, for an arbitrary small probability of error, the maximum possible bandwidth efficiency is limited by the noise in the channel, and is given by:  $C/W = \log_2(1 + \gamma)$ . This theorem applied to an AWGN time-invariant channel with constant  $\gamma$ , in this case the noise of the channel rather than interference is present; where C and W are the capacity and bandwidth of the channel respectively. In cellular system where interference is a major issue, the performance of a modulation scheme in terms of spectral efficiency is interference limited rather than noise limited. For fading channels, LEE [1] has derived the capacity of the channel to consider the variation of  $\gamma$  due to different propagation impairment in a Rayleigh channel. The authors in [3] have derived an expression for the capacity of fading channels for different adaptation schemes when information about the channel conditions was provided for both transmitter and receiver or one of them. Goldsmith [2] used the theory that developed in [3] to obtain a closed-form expression for the capacity of Rayleigh fading channels under different adaptive transmission and diversity combining techniques.

However, if the channel fade level is known to the transmitter, then Shannon capacity is achieved by adapting the transmit parameters (power, data rate, and coding schemes) relative to this fade level.

In [4], the authors' proposed adaptive variable rate-variable power uncoded MQAM in which they derived the spectral efficiency for their proposed system. They have shown that a power loss of  $K$  depending on BER and independent of the fading distribution has been introduced when using MQAM relative to the optimal transmission scheme. In this paper, the same system is considered with adaptive rate and fixed power transmission combined with diversity at the receiver to observe the improvement that a diversity-combining scheme would introduce to the spectral efficiency.

## 2 Channel model

We assume a slow varying channel at a rate much slower than the data rate and a Rayleigh fading channel with exponentially distributed PDF of carrier to noise ratio, CNR

$$p(\gamma) = \frac{e^{-\gamma/\bar{\gamma}}}{\bar{\gamma}} \quad (1)$$

Where  $\bar{\gamma}$  is the average received CNR.

In this work, the adaptive rate fixed-power MQAM system is combined with a well-known diversity combining techniques. In particular we use the maximal ratio combining (MRC) and selection combining (SC) of the received signal. The former requires the  $M$  signals to be weighted proportionately to their CNR and then summed coherently. Perfect knowledge of the branch amplitudes and phase is assumed. The disadvantage of MRC is that it requires knowledge of the branch parameters and independent processing of each branch. The PDF of the received CNR at the output of a perfect  $M$ -branch MRC is derived in [5] to be:

$$P^{mrc}(\gamma) = \gamma^{M-1} e^{-\gamma/\bar{\gamma}} / (M-1)\bar{\gamma}^{-M} \quad (2)$$

In the SC technique only one of the  $M$  receivers having the highest baseband CNR is connected to

the output. Unlike the MRC it does not require coherent reception.

The PDF of the received CNR at the output of  $M$ -branch is again derived in [5] and it is given by:

$$P^{sc}(\gamma) = \frac{M}{\gamma} (1 - e^{-\gamma/\bar{\gamma}})^{M-1} e^{-\gamma/\bar{\gamma}} \quad (3)$$

## 3 Adaptive MQAM

The spectral efficiency of a communication link can be increased by using a multilevel modulation scheme, such as MQAM, which tends to send a multiple bits per symbol.

The radio channels in a wireless mobile communication system are affected by different types of fading (multipath, shadowing, etc), therefore, they will have negative effect on signals carried on these channels.

To compensate for these channel impairments imposed by fading, adaptive modulation scheme is used. In adaptive MQAM, information about the channel conditions at the receiver is fed back to the transmitter so that it will adjust its' transmitted modulation level (constellation size) accordingly. This channel information is usually acquired by using a pilot signal or inserting a training sequence into the stream of MQAM data symbol to extract the channel induced attenuation and phase shift [6].

The BER of a coherent MQAM over an AWGN channel assuming a perfect clock and carrier recovery can be well approximated by [4]

$$BER(M, \gamma) \approx 0.2 e^{-\frac{3\gamma}{2(M-1)}} \quad (4)$$

This expression for the BER approximation is invertible [7] as it provides a closed form for the spectral efficiency of MQAM as a function of CNR and BER.

For given CNR and assuming ideal Nyquist pulses the spectral efficiency of a continuous rate MQAM can be approximated by inverting Eq.4 giving;

$$\frac{R}{W} = \log_2(M) = \log_2\left(1 + \frac{3\gamma}{2K}\right) \quad (5)$$

Where  $R$  is channel bit rate, and  $K = -\ln(5\text{BER})$ .

In practice the CNR is not fixed but fluctuates due to channel impairments, therefore the spectral efficiency is calculated by integrating the RH side of Eq. 5 over the distribution of CNR i.e. PDF of the received CNR as shown below:

$$\frac{R}{W} = \int \log_2 \left( 1 + \frac{3 \cdot \gamma}{k} \right) \cdot P(\gamma) d\gamma$$

Hence for the MRC diversity case, we integrate over MRC distribution function as given in Eq. 2:

$$\frac{R}{W} = \int \log_2 \left( 1 + \frac{3\gamma}{2K} \right) \frac{\gamma^{M-1} e^{-\gamma}}{(M-1)! \gamma^M} d\gamma \quad (6)$$

To yield the following

$$\begin{aligned} \frac{C}{W}^{mrc} &\approx [\log_2(e)] \times P_M \left( \frac{-1}{\bar{\gamma}_o} \right) \left( -E + \ln \bar{\gamma}_o + \frac{1}{\bar{\gamma}_o} \right) \\ &+ \sum_{k=1}^{M-1} \frac{P_k \left( \frac{-1}{\bar{\gamma}_o} \right) - P_{M-k} \left( \frac{-1}{\bar{\gamma}_o} \right)}{k} \end{aligned} \quad (7)$$

Where  $\bar{\gamma}_o = \frac{3}{2K} \bar{\gamma}$ , and  $\bar{\gamma}$  is the average received CNR and  $k$  is an integer.

The achievable spectral efficiency is compared with that achieved by an  $M$ -array independent AWGN channels, optimal combining (MRC), and to that of optimal rate and constant power without diversity i.e. a normal Rayleigh channel, which are given by

$$C_{\text{AWGN}} = W \log_2(1 + M \cdot \gamma) \quad (8)$$

$$C_{\text{opt-rate}} = \int_0^{\infty} W \log(1 + \gamma) \rho(\gamma) d\gamma, \quad (9)$$

where  $\rho(\gamma)$  is given in Eq.1. In the case of SC diversity we integrate Eq.5 over Eq.3 to yield the following approximation [2].

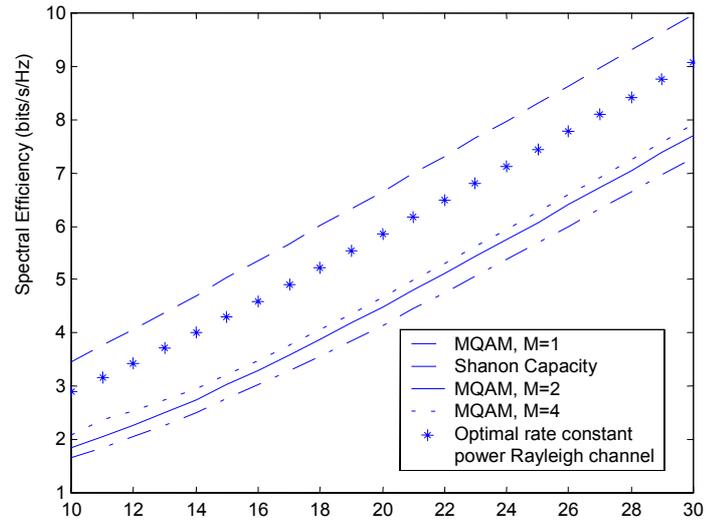
$$\begin{aligned} \frac{C}{W}^{Sc} &\approx M \log_2(e) \sum_{k=0}^{M-1} \frac{(-1)^k}{1+k} \binom{M-1}{k} e^{-\frac{(1+k)}{\gamma}} \\ &\left[ E + \ln \left( \frac{1+k}{\gamma} \right) - \left( \frac{1+k}{\gamma} \right) \right] \end{aligned} \quad (10)$$

Where  $E$  is Euler constant ( $E = 0.57721566$ ).

However our results will be restricted to those of MRC diversity scheme.

## 4 Results & Discussions

In Figure 1, the spectral efficiency for different diversity levels of MQAM system in the case of MRC is shown. The spectral efficiency of Rayleigh channel with optimal rate and constant power, and the Shannon capacity are shown as well.



**Figure 1: Capacity of MQAM with MRC diversity system for average combined received SIR value.**

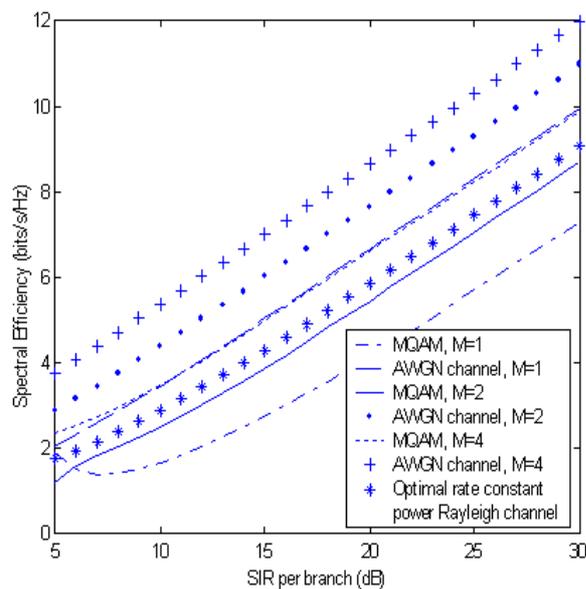
We notice that when the value of CNR was taken as the average combined received CNR an increase of about 1/4 bits/sec/Hz as the diversity goes from  $M=1$  to  $M=2$ , and another 1/4 bit/sec/Hz as it increases to  $M=4$ .

The results are compared to those of optimal rate and constant power Rayleigh fading channel and to the Shannon AWGN channel capacity  $C/W = \log_2(1 + \gamma)$ .

It could be realised that when the value of CNR was taken as the average combined received CNR, the obtained results from MRC diversity were identical to those obtained in Nakagami fading channels with the level of diversity corresponding to the value of the Nakagami fading parameter ( $m$ ). This is an indication that less fading is being exposed to the channel (increasing  $m$ ) when the diversity increased. It could be noticed that as the diversity increased the spectral efficiency

approaches that of the Shannon capacity  $C/W = \log_2(1 + \gamma)$ .

When the value of CNR was taken as the per branch power, then we can see that as the diversity level increases the spectral efficiency approaches that of the Shannon capacity of an M-array of independent AWGN channels  $C_{AWGN} = W \log_2(1 + M.\gamma)$  as shown in Figure 2. More significant improvement can be seen when the calculations are based on the per branch value of CNR. An increase of 1.3 bits/sec/Hz achieved when the diversity increased from M=1 to M=2, and 1.2 bits/sec/ Hz achieved as the diversity increased from M=2 to M=4.



**Figure 2: Capacity of MQAM system with MRC diversity for per branch SIR value**

It could be noticed that when M=4, the spectral efficiency of the MQAM system has exceeded that of the optimal rate constant power Rayleigh channel and it coincides with the M=2 of an M array AWGN channel (Eq.8).

The above explanation could be applied for the results obtained when SC diversity technique was used with the results showing less performance in this case.

## 5 Conclusion

As the work in [4] shows that there is a constant gap between the channel capacity and the maximum efficiency of adaptive MQAM which is a simple function of BER. We have combined diversity schemes in order to see the effects on the spectral efficiency of adaptive MQAM

communication system. The results show slight improvement in the spectral efficiency as the diversity level increases. These results indicate that the complexity introduced to the system by the diversity technique is unjustified as they have very little significance. However, if the per branch CNR were taken in consideration the results have improved but still cannot justify the complexity of using diversity technique.

## References

- 1- William. C. Lee, "Estimate of channel capacity in Rayleigh fading environment" IEEE transaction. Veh. Technology, vol. VT-39, pp187-190, August 1990.
- 2- Mohammed Alouini, A Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity combining techniques" IEEE Transaction on Vehicular Technology, vol. 48, NO. 4, pp. 1165-1181, July 1999.
- 3- A. Goldsmith, P. Varaiya, "Capacity of fading channels with channel side information," IEEE trans. Inform. Theory, Vol. IT-43, pp. 187-190, November 1997.
- 4- A. Goldsmith, and S. Chua, "Variable rate variable power M-QAM for fading channels", IEEE trans. On communication, Vol. COM\_45, pp.1218-1230, October 1997.
- 5- William Jakes, *Microwave mobile communications*, John-Wiley & sons, 1974.
- 6- S. Sampei and Sunga, "Rayleigh fading compensation for QAM in Land mobile Radio communications." IEEE trans.Veh. Technology, Vol. VT-43, pp. 294-302, May 1993.
- 7- Mohammed Alouini, A Goldsmith, "Adaptive Modulation over Nakagami Fading channels", Kluwer Journal on Wireless Communications, Vol. 13, No. 1-2, pp.119-143, May 2000.

# Modelling of Saturation Intensity in Principal Channel for Dye Solutions with Coincident Absorption and Emission Bands

Jihad S. M. Addasi  
Department of Basic Sciences  
Tafila Applied University College  
Al-Balqa' Applied University  
P.O. Box 179, Tafila-66110, Jordan  
Tel.: +962-3-2250659  
E-mail: Jihad\_add@yahoo.com

## Abstract

A phase response of dye solution can be taken into consideration to develop many nonlinear phenomena, such as four-wave mixing, holography, self-diffraction, amplification [1, 2, 5, 6 and 7]. Nonlinear processes strongly depend on intensity and frequency of radiations involved in these processes.

A nonlinear medium can be modeled by a typical three-level configuration ( $S_0$ - $S_1$ - $S_2$ ) [2], for which the transitions of molecules in principal channel ( $S_0$ - $S_1$ ) are realized by the light fields of intensity  $I_{12}$  at frequency  $\omega_0$ . At the same time, radiations with intensity  $I_{23}$  at frequency  $\omega$  interact with excited molecules to realize their transitions in excited channel ( $S_1$ - $S_2$ ). For this model, the conditions of phase response saturation are determined in principal channel.

The saturation of phase response of dye solution in principal channel is realized at saturation intensity  $I_{12}^{sat}$ , decreasing with increasing of radiation intensity in excited channel  $I_{23}$ . The saturation intensity  $I_{12}^{sat}$  has its optimum (minimum) values when the frequency of light fields in principal channel is tuned into the centre of principal absorption band. In addition, the saturation intensity  $I_{12}^{sat}$  has its optimum when the radiations in excited channel have enough big intensity  $I_{23}$  and a frequency tuning into the centre of absorption excited band.

## Introduction

A cubic nonlinearity of dye solution represents a basic information to study nonlinear processes, four-wave mixing amplification holography [1, 2, 3 and 5]. The

lifetimes of vibration energy levels of dye solutions are significantly lower than the lifetimes of electronic energy levels [7]. In this case, the electronic states can be taken as a homogeneously broadened levels, which gives ability to use two-, three-level models and others with averaged Einstein coefficients for many nonlinear medium [2]. Two-level model can be used to study nonlinear processes involved by radiations of intensity  $I_{12}$  at frequency  $\omega_0$ , but in this model the radiations can't have two different independent frequencies, see figure 1 (a).

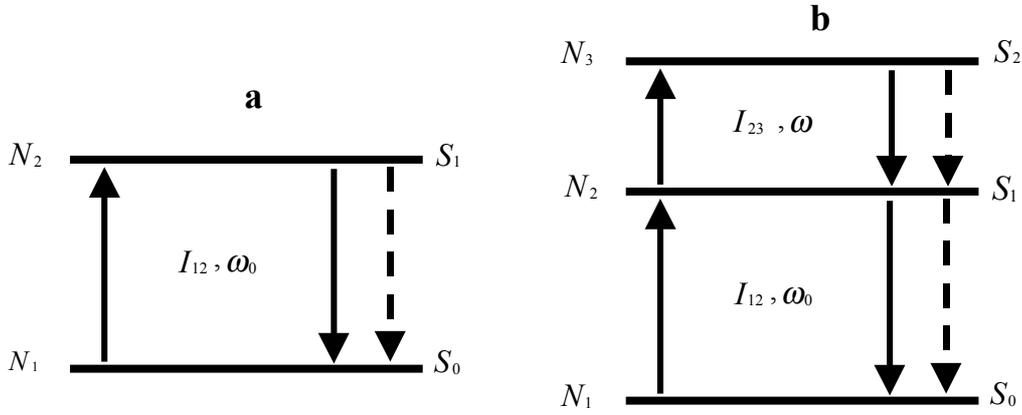
In three-level model, illustrated in figure 1 (b), the dye solution can be excited by light fields with two different independent frequencies, one group of light fields acts in principal ( $S_0$ - $S_1$ ) channel (with intensity  $I_{12}$  at frequency  $\omega_0$ ) and other group acts in excited ( $S_1$ - $S_2$ ) channel (with intensity  $I_{23}$  at frequency  $\omega$ ). Light fields in one channel can involve nonlinear processes, and other light fields act in second channel to control these processes [3]. The balance equations under a double frequencies excitation of dye solution modeled by three-level model can be written as follows:

$$\begin{aligned} N_1 B_{12}(\omega_0) I_{12} &= N_2 (B_{21}(\omega_0) I_{12} + \nu p_{21}) ; \\ N_2 B_{23}(\omega) I_{23} &= N_3 (B_{32}(\omega) I_{23} + \nu p_{32}) ; \quad (1) \\ N &= N_1 + N_2 + N_3 ; \end{aligned}$$

where

$N_i$  - is the population of  $i$ - energy level;  $N$  - is the number of molecules in the unit volume of nonlinear medium;  $P_{ij}$  - is the total probability of spontaneous and radiationless transitions in the  $i$ - $j$  channel;  $\nu = c/n$  - is the light velocity in the nonlinear medium. The Einstein coefficients  $B_{12}(\omega_0)$ ;  $B_{21}(\omega_0)$  - are determined at the frequency of radiations  $\omega_0$  in principal

( $S_0$ - $S_1$ ) channel. At the same time  $B_{23}(\omega)$ ;  $B_{32}(\omega)$  - are determined at frequency of radiations  $\omega$  in excited channel. The refractive



**Figure 1** Diagram of (a) two- (b) three-level models. The solid lines denote the radiation-induced transitions of molecules, and the dashed lines denote spontaneous and radiationless transitions. Where:  $S_0$  - is the ground state;  $S_1$  ( $S_2$ ) - are the first ( second ) excited states;  $N_i$  - is the population of  $i$ - energy level;  $I_{12}, \omega_0$  ( $I_{23}, \omega$ ) - are the intensity, frequency of radiation in principal (excited) channel.

## Theory

The saturation intensity  $I_{12}^{sat}$  in principal channel is defined as the value of radiation intensity, acting in principal channel, for which the absorption is decreasing in half of its initial value ( $K_{12}(I_{12}^{sat}) = (1/2)K_{12}(I_{12} = 0)$ ). The extinction coefficient in principal channel at frequency  $\omega_0$  can be found by the following expression:

$$\chi_{12}(\omega_0) = \frac{c}{2\nu} K_{12}(\omega_0); \quad (2)$$

where

$$K_{12}(\omega_0) = \frac{\hbar\omega_0}{\nu} (N_1 B_{12}(\omega_0) - N_2 B_{21}(\omega_0)) - \text{is}$$

the absorption coefficient in principal channel. From equations (1) and (2) the extinction coefficient  $\chi_{12}(\omega_0)$  will be:

$$\chi_{12}(\omega_0) = \chi_0 (1 + aI_{23}) / K; \quad (3)$$

where

$$K = 1 + JI_{12} + aI_{23} + bI_{12}I_{23};$$

$$a = B_{32} / \nu p_{32}; \quad J = (B_{12} + B_{21}) / \nu p_{21};$$

$$b = B_{12}B_{23} / \nu^2 p_{21}p_{32} + aJ;$$

$\chi_0 = N\hbar c B_{12}(\omega_0) / 2\nu$ ;  $\chi_0(\omega_0)$  - is the linear extinction coefficient. The extinction coefficient, equation (3), has a monotonic proportionality with intensity of radiations in each channel ( $I_{12}$  and  $I_{23}$ ) and has its maximum value ( $\chi_{12} = \chi_0$ ) at intensities

index determined by balance equations can be used to study many nonlinear processes such as bleaching processes.

$I_{12} = I_{23} = 0$ . The extinction coefficient has the half of its maximum value at saturation intensity in principal channel ( $I_{12}^{sat}$ ) with value:

$$I_{12}^{sat} = \frac{1 + aI_{23}}{J + bI_{23}}; \quad (4)$$

From equation (4) the saturation intensity ( $I_{12}^{sat}$ ) has a monotonic dependence on radiation intensity in excited channel ( $I_{23}$ ). To study the dependence of saturation intensity  $I_{12}^{sat}$  on frequency tuning of radiations in principal and excited channels, some parameters of nonlinear medium must be determined, especially this dependence has an optimum values as a function of frequency tuning.

## Analysis and Discussion

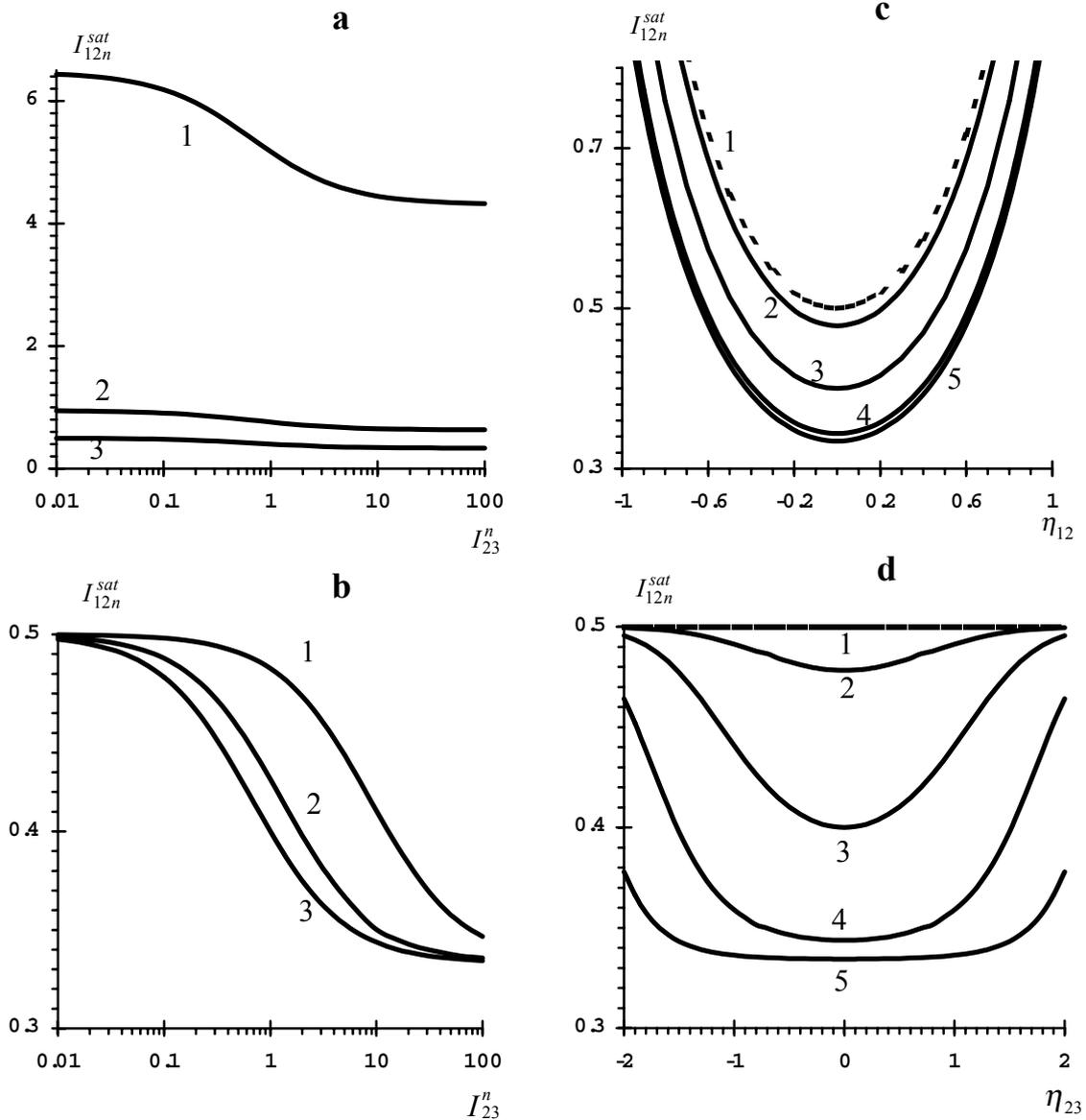
Taking into consideration a nonlinear medium with a Gaussian form of coincident mirror-symmetric absorption and emission bands ( $\omega_{ij} = \omega_{ji} \Rightarrow \delta_{ij} = (\omega_{ij} - \omega_{ji}) / \Delta_{ij} = 0$ ), where  $\Delta_{ij}, \omega_{ij}$  - are the centre; the profile halfwidth of  $i$ - $j$  band. For this matter the frequency tuning of radiations in principal ( $\eta_{12} = (\omega_0 - \omega_{12}) / \Delta_{12}$ ) and excited ( $\eta_{23} = (\omega_0 - \omega_{23}) / \Delta_{23}$ ) channels

are used to find Einstein coefficients  $B_{ij}$ . From equation (4) the saturation intensity ( $I_{12}^{sat}$ ) in principal channel, for this form of absorption and emission bands, has optimum values (minimum) at radiations frequency in principal channel tuned into the centre of principal absorption band ( $\eta_{12}^0 = 0$ ). At the same time, the optimum of  $I_{12}^{sat}$  is realized at radiations frequency in excited channel tuned into the centre of excited absorption band ( $\eta_{23}^0 = 0$ ), see figure 2 (c) and (d). All equations (1), (2), (3) and (4) will be true for

two-level model, when there is no radiations in excited channel ( $I_{23} = 0 \Rightarrow N_3 = 0$ ). In this case let us rewrite equation (4) for two-level model:

$$I_{12}^{sat} = (1/J); \quad (4^*)$$

The saturation intensity  $I_{12}^{sat}$  for two-level model, equation (4\*), has its optimum values at radiations frequency tuned into the center of absorption band ( $\eta_{12}^0 = 0$ ) for nonlinear medium with coincident absorption and emission bands, see figure 2 (c) curve 1.



**Figure 2** - Dependence of saturation intensity  $I_{12n}^{sat}$  on: radiation intensity  $I_{23}^n$  (a); (b); frequency tuning of radiations in principal  $\eta_{12}$  (c) and in excited  $\eta_{23}$  (d) channels. Curves: 1, 2 and 3 at: (a)  $\eta_{12} : \pm 1.6; \pm 0.8$  and 0; (b)  $\eta_{23} : \pm 1.6; \pm 0.8$  and 0. Curves: 1; 2; 3; 4 and 5 at (c) and (d)  $I_{23}^n : 0$  (two-level model); 0.1; 1; 10 and 100 respectively. Where for: (b), (d)  $\eta_{12} = 0$ ; (a), (c)  $\eta_{23} = 0$ .

In figures 2, the following statements are considered: the maximum values of Einstein coefficients are the same for all bands  $B_{12}^{\max} = B_{21}^{\max} = B_{23}^{\max} = B_{32}^{\max}$ , the radiations intensities are normalized to the value  $\nu p_{21} / B_{12}^{\max}$  and  $\nu p_{32} / B_{32}^{\max}$  in the principal and in the excited channels, respectively ( $I_{12n}^{(sat)} = I_{12}^{(sat)} B_{12}^{\max} / \nu p_{21}$ ;  $I_{23}^n = I_{23} B_{32}^{\max} / \nu p_{32}$ ). Figure 2 (a) and (b) show the monotonic dependence of saturation intensity  $I_{12}^{sat}$  on radiation intensity in excited channel  $I_{23}$ . The saturation in principal channel is realized for enough small intensity in principal channel  $I_{12}^{sat}$  when an effective excitation of molecules happens in excited channel, radiations in excited channel must have enough big intensity and frequency tuned into the centre of excited absorption band ( $I_{23}^n \gg 1$ ;  $\eta_{23} \approx 0$ ), see figure 2.

Equation (4\*), for two-level model, is illustrated in figure 2 (c) and (d) by curve 1, where the optimum value of saturation intensity  $I_{12n}^{sat} = 0.5$  is realized at frequency tuning  $\eta_{12}^0 = 0$ . Figure 2 (c) and (d) show that, the saturation intensity  $I_{12}^{sat}$  in principal channel for three-level model (curves: 2, 3, 4 and 5) always lower than for two-level model at any frequency tuning or intensity of radiations in excited channel ( $I_{23n}$ ;  $\eta_{23}$ ) and at any frequency tuning in principal channel ( $\eta_{12}$ ).

It's easy to realized and control the saturation and other nonlinear processes in principal channel by using external radiations in excited channel [2, 3, 6]. Figure 2 shows that a little frequency shift of radiations from centre of principal absorption band ( $\eta_{12} \neq 0$ ) increases the saturation intensity ( $I_{12}^{sat}$ ) significantly, but the frequency shift in excited channel ( $\eta_{23} \neq 0$ ) makes a little change in saturation intensity ( $I_{12}^{sat}$ ) of principal channel.

## Conclusion

For double frequencies excitation of nonlinear medium, the saturation intensity in principal channel  $I_{12}^{sat}$  is decreasing with increasing

radiation intensity in excited channel. The optimization of saturation intensity is realized with the following statements:

1. An effective excitation of molecules in excited channel ( $I_{23} \gg \nu p_{32} / B_{32}$ ;  $\eta_{23} \approx 0$ ).
2. The radiations in the principal channel must be tuned into the centre of principal absorption band, frequency shift from centre of absorption band not more than (1/3) of its halfwidth.
3. The radiations in excited channel must be tuned into the centre of excited absorption band, frequency shift from centre of absorption band not more than its halfwidth.

## References

1. Ivakin E V, Petrovich I N and Rubanov A S 1973 Russian J. of Applied Spectroscopy **18** N. 6 1003.
2. Agishev I N, Ivanova N A and Tolstik A L 1998 Optics Communications **156** 199-209.
3. Addasi J S, Tolstik A L and Chaley A V 1993 Trans. of Byelorussian University I N. 1 9-12.
4. Vasileva M A, Vishcakas J and Gulbinas V 1987 Quantum Electronics **14** 1691.
5. Gancherenok I I, Romanov O G, Tolstik A L, Fleck B and Wenke L 2001 J. of Optics B **3** 220-224.
6. Rubanov A S, Tolstik A L, Karpuk S M and Ormachea O 2000 Optics Communications **181** 183-190.
7. Tichonov E A and Shpak M T 1979 Nonlinear Optical Effects in Organic Compounds Kiev Naukova Dumka 90-100.

## Biography

Jihad Addasi was born in Jordan in 1964. He received the Master of Science degree in Physics and Mathematics from Faculty of Physics of Byelorussian State University with honours in 1989. His Doctor of Philosophy degree is in Physics and Mathematics with Specialty Laser Physics (1993). His main research interests are in Laser Physics and Nonlinear Optics. He is a lecturer in Department of Basic Sciences at Tafila Applied University College in Al-Balqa Applied University, Jordan.

# COMPUTATIONAL ALGORITHM FOR VOLTAGE ASYMMETRY MEASUREMENT IN ELECTRICAL NETWORKS

Dr. Mahmoud Salama Awad  
Al-Balqa' Applied University, Engineering Technology College  
P.O. Box 15008, Amman, Jordan.  
Tel.: +962.6.4894291, Fax: +962.6.4894292  
E-mail: [Dr\\_Awad\\_M@yahoo.com](mailto:Dr_Awad_M@yahoo.com)

Increasing of efficiency and reliability of electrical networks and systems operation depends on the quality of electrical energy. In three-phase networks voltage asymmetry is one of the most factors, which affects the electrical energy quality. This asymmetry is usually defined by the approximate method. In this paper the error of the approximate method is evaluated to define the restriction of its use and other methods of voltage asymmetry measurement, which are based on data acquisition, are suggested.

For a general case, assume that  $\dot{V}_A$ ,  $\dot{V}_B$  and  $\dot{V}_C$  compose asymmetrical vectors system. Then they can be resolved into symmetrical components:

$$\left. \begin{aligned} \dot{V}_1 &= \frac{1}{3}(\dot{V}_A + a\dot{V}_B + a^2\dot{V}_C) \\ \dot{V}_2 &= \frac{1}{3}(\dot{V}_A + a^2\dot{V}_B + a\dot{V}_C) \\ \dot{V}_0 &= \frac{1}{3}(\dot{V}_A + \dot{V}_B + \dot{V}_C) \end{aligned} \right\} \quad (1)$$

where  $a = e^{j\frac{2\pi}{3}}$  - rotation operator. If the system is symmetrical, then:

$$\begin{aligned} \dot{V}_B^0 &= a^2\dot{V}_A^0 \\ \dot{V}_C^0 &= a\dot{V}_A^0 \end{aligned}$$

Introducing two vectors, which are describe the deviation of the vectors system from the symmetry:

$$\dot{g}_B = \frac{\dot{V}_B}{\dot{V}_B^0},$$

$$\dot{g}_C = \frac{\dot{V}_C}{\dot{V}_C^0},$$

then:

$$\left. \begin{aligned} \dot{V}_B &= \dot{g}_B \dot{V}_B^0 = a^2 \dot{V}_A^0 \dot{g}_B \\ \dot{V}_C &= \dot{g}_C \dot{V}_C^0 = a \dot{V}_A^0 \dot{g}_C \end{aligned} \right\} \quad (2)$$

Substituting (2) into (1), the voltage negative and zero components can be written as follows:

$$\left. \begin{aligned} \dot{V}_2 &= \frac{1}{3} \left( 1 + a\dot{g}_B + a^2\dot{g}_C \right) \dot{V}_A^0 \\ \dot{V}_0 &= \frac{1}{3} \left( 1 + a^2\dot{g}_B + a\dot{g}_C \right) \dot{V}_A^0 \end{aligned} \right\} \quad (3)$$

Introducing vectors:

$$\begin{aligned} \Delta\dot{g}_B &= a\dot{g}_B - 1; \\ \Delta\dot{g}_C &= a\dot{g}_C - 1. \end{aligned}$$

Then the voltage negative component can be written as follows:

$$\dot{V}_2 = \frac{1}{3} \dot{V}_A^0 \left( a\Delta\dot{g}_B + a^2\Delta\dot{g}_C \right).$$

Here the following is taken into account:

$$1 + a + a^2 = 0.$$

To find out the voltage negative component:

$$V_2 = |\dot{V}_2|,$$

and using:

$$|a^2| = aa^*,$$

where  $a^*$  - the complex conjugate of the rotation operator, we have:

$$V_2^2 = |\dot{V}_2|^2 = \frac{1}{9} V_A^2 \left( a \Delta \dot{g}_B + a^2 \Delta \dot{g}_C \right) \left( a^2 \Delta g_B^* + a \Delta g_C^* \right).$$

After some calculations we have:

$$V_2^2 = \frac{1}{9} V_A^2 \left\{ |\Delta \dot{g}_B|^2 + |\Delta \dot{g}_C|^2 + 2 \operatorname{Re} \left( a \Delta \dot{g}_C \Delta g_B^* \right) \right\}. \quad (4)$$

Using the following formulas:

$$\dot{g}_B = g_B e^{j\Delta\varphi_B};$$

$$|\dot{g}_B| = g_B;$$

$$\dot{g}_C = g_C e^{j\Delta\varphi_C},$$

$$|\dot{g}_C| = g_C;$$

we find:

$$|\Delta \dot{g}_B|^2 = |g_B e^{j\Delta\varphi_B} - 1|^2 = g_B^2 + 1 - 2 g_B \cos \Delta \varphi_B; \quad (5)$$

$$|\Delta \dot{g}_C|^2 = |g_C e^{j\Delta\varphi_C} - 1|^2 = g_C^2 + 1 - 2 g_C \cos \Delta \varphi_C; \quad (6)$$

$$\Delta \dot{g}_C \Delta g_B^* = g_B g_C e^{j(\Delta\varphi_C - \Delta\varphi_B)} - g_B e^{-j\Delta\varphi_B} - g_C e^{j\Delta\varphi_C}; \quad (7)$$

$$2 \operatorname{Re} \left( a \dot{g}_C \Delta g_B^* \right) = -g_B g_C \cos(\Delta \varphi_C - \Delta \varphi_B) + g_B \cos \Delta \varphi_B - g_C \cos \Delta \varphi_C - 1 + \sqrt{3} \left[ g_B g_C \sin(\Delta \varphi_B - \Delta \varphi_C) - g_B \sin \Delta \varphi_B + g_C \sin \Delta \varphi_C \right]. \quad (8)$$

Substituting (5), (6) and (8) into (4) we get:

$$V_2^2 = \frac{1}{9} V_A^2 \left\{ g_B^2 + g_C^2 + 1 - g_B - g_C + [ g_B (1 - \cos \Delta \varphi_B) + g_C (1 - \cos \Delta \varphi_C) + g_B g_C (1 - \cos(\Delta \varphi_C - \Delta \varphi_B)) + \sqrt{3} g_B g_C \sin(\Delta \varphi_B - \Delta \varphi_C) + \sqrt{3} g_B \sin \Delta \varphi_B + g_C \sin \Delta \varphi_C ] \right\}$$

Introducing the following notations:

$$A = \frac{1}{2} \left[ (g_B - 1)^2 + (g_C - 1)^2 + (g_B - g_C)^2 \right], \quad (9)$$

$$A = g_B^2 + g_C^2 + 1 - g_B - g_C - g_B g_C;$$

$$B = [ g_B (1 - \cos \Delta \varphi_B) + \dots + ].$$

Value  $A$ , which can be written as follows:

depends only on the vectors values ( $V_A$ ,  $V_B$  and  $V_C$ ) and does not depend on the phase of these vectors. Value  $B$  depends on the phase  $\Delta \varphi_B$ ,  $\Delta \varphi_C$ . While calculating the voltage negative component, in order to neglect the phases, the value  $B$  must be small; that means

phase deviation  $\Delta\varphi_B$ ,  $\Delta\varphi_C$  from symmetrical system must be small. If the value  $B$  is negligible in comparison with value  $A$ , then:

$$g_B = \frac{V_B}{V_A};$$

$$g_C = \frac{V_C}{V_A},$$

$$V_2 = \frac{1}{3} \sqrt{(g_B - 1)^2 + (g_C - 1)^2 - (g_B - 1)(g_C - 1)}.$$

we have:

As:

$$V_2 = \frac{1}{3} \sqrt{(V_B + V_A)^2 + (V_C - V_A)^2 - (V_B - V_C)(V_C - V_A)}. \quad (10)$$

Taking (9) in consideration, formula (10) can be written as follows:

$$V_2 = \frac{1}{3\sqrt{2}} \sqrt{(V_B - V_A)^2 + (V_C - V_A)^2 + (V_B - V_C)^2}. \quad (10)$$

The error in finding the voltage negative component can be evaluated using formulas (10) and (11). Exact formula can be written as follows:

Then relative error will be:

$$\delta_2 = \frac{B_2}{2A}. \quad (14)$$

$$V_2 = \frac{1}{3} V_A \sqrt{A - B_2}. \quad (12)$$

If  $B_2 \ll A$ , then resolving (12) into series we get:

$$V_2 \cong \frac{1}{3} V_A \left( \sqrt{A} + \frac{B_2}{2\sqrt{A}} \right). \quad (13)$$

For small  $B_2$ , which means small  $\Delta\varphi_A$ ,  $\Delta\varphi_B$  and considering formulas (10), (11), (13) and (14) we get:

$$B_2 \cong \frac{1}{2} \left[ g_B \Delta\varphi_B^2 + g_C \Delta\varphi_C^2 + g_B g_C (\Delta\varphi_B - \Delta\varphi_C)^2 \right] + \sqrt{3} \left[ g_B (g_C - 1) \Delta\varphi_B - g_C (g_B - 1) \Delta\varphi_C \right]. \quad (15)$$

The first component in formula (15) is second order infinitesimal, while the second component is first order infinitesimal. Thus, when the second component goes to zero, it is bigger than the first. Thus, we have:

$$\delta_2 = \frac{\sqrt{3}}{2} \frac{g_B (g_C - 1) \Delta\varphi_B - g_C (g_B - 1) \Delta\varphi_C}{(g_B - 1)^2 + (g_C - 1)^2 - (g_C - 1)(g_B - 1)}. \quad (16)$$

$$B_2 \cong \sqrt{3} \left[ g_B (g_C - 1) \Delta\varphi_B - g_C (g_B - 1) \Delta\varphi_C \right],$$

Substituting the following:

$$g_B = \frac{V_B}{V_A};$$

$$g_C = \frac{V_C}{V_A},$$

and the relative error:

we get:

$$\delta_3 = \sqrt{3} \frac{V_B (V_C - V_A) \Delta \varphi_B + V_C (V_A - V_B) \Delta \varphi_C}{(V_A - V_B)^2 + (V_C - V_A)^2 - (V_C - V_A)^2}. \quad (17)$$

It is obvious from (2) that in order to calculate the zero sequence component, the transformation from the negative sequence component to the zero sequence component can be obtained by switching the indexes  $B \leftrightarrow C$ . The value  $A$  here is kept constant, thus the

approximate voltage negative and zero components are the same:

$$V_0 \cong V_2. \quad (18)$$

Formula for error calculation follows from the  $B \leftrightarrow C$  switching.

$$\delta_0 = \sqrt{3} \frac{V_C (V_B - V_A) \Delta \varphi_C + V_B (V_A - V_C) \Delta \varphi_B}{(V_A - V_B)^2 + (V_B - V_C)^2 + (V_C - V_A)^2}. \quad (19)$$

If  $\delta_2$  and  $\delta_0$  are not small, then it is not allowed to use the above mentioned approximation, and formula (12) must be for accurate calculation of the voltage negative and zero sequence components, but in this case we have to know the phase shift between the vectors (in such a case it is easy to use other methods to calculate the voltage negative and zero sequence components).

In order to measure voltage asymmetry by three-phase voltage system transformation into two-phase system, then measuring voltage symmetrical components, the following should be done: voltage positive and negative components should be written in the following form:

$$\begin{aligned} \dot{V}_1 &= \frac{1}{3} (\dot{V} + j \dot{W}) ; \\ \dot{V}_2 &= \frac{1}{3} (\dot{V} - j \dot{W}) , \end{aligned} \quad (20)$$

where

$$\left. \begin{aligned} \dot{V} &= \dot{V}_A - 0,5(\dot{V}_B - \dot{V}_C) \\ \dot{W} &= \frac{\sqrt{3}}{2} (\dot{V}_B - \dot{V}_C) \end{aligned} \right\} \quad (21)$$

Then the effective values of the voltage positive and negative sequence components:

$$V_1^2 = \frac{1}{9} [V^2 + W^2 + 2VW \sin(\varphi_V - \varphi_W)] ;$$

$$V_2^2 = \frac{1}{9} [V^2 + W^2 - 2VW \sin(\varphi_V - \varphi_W)] , \quad (22)$$

where

$\varphi_V, \varphi_W$  - initial phases of  $V$  and  $W$  respectively.

The component  $VW \sin(\varphi_V - \varphi_W)$  can be presented as follows:

$$VW \sin(\varphi_V - \varphi_W) = V_Y W_Y - V_X W_X ,$$

where

$$V_X = V \cos \varphi_V ;$$

$$V_Y = V \sin \varphi_V ;$$

$$W_X = W \cos \varphi_W ;$$

$$W_Y = W \sin \varphi_W ,$$

are square components of the effective voltage values  $V$  and  $W$ . Taking into account (22) we have:

$$\left. \begin{aligned} V_1^2 &= \frac{1}{9} (V_X^2 + V_Y^2 + W_X^2 + W_Y^2 + 2V_Y W_X - 2V_X W_Y) \\ V_2^2 &= \frac{1}{9} (V_X^2 + V_Y^2 + W_X^2 + W_Y^2 - 2V_Y W_X + 2V_X W_Y) \end{aligned} \right\} (23)$$

Square components can be found by discrete Fourier transform:

$$\begin{aligned} V_X &= \frac{2}{n} \sum_{i=0}^{n-1} V(t_i) \sin \omega t_i; \\ V_Y &= \frac{2}{n} \sum_{i=0}^{n-1} V(t_i) \cos \omega t_i; \\ W_X &= \frac{2}{n} \sum_{i=0}^{n-1} W(t_i) \sin \omega t_i; \\ W_Y &= \frac{2}{n} \sum_{i=0}^{n-1} W(t_i) \cos \omega t_i, \end{aligned} \quad (24)$$

where the sampling instants for  $V$  and  $W$  signals:

$$t_i = \frac{2\pi}{n\omega} i;$$

$n$  – the number of the sampling points;  $\sin \omega t_i$  and  $\cos \omega t_i$  – numbered harmonics instantaneous values at the sampling points;

$V(t_i)$  and  $W(t_i)$  – instantaneous voltage values at the sampling points.

The block diagram of the suggested method for voltage asymmetry measurement is shown in the figure. Phase voltages are applied to the analog to digital converter and to the control block, which gives the instants of phase voltages sampling, which are applied to the analog to digital converter. Instantaneous phase voltages codes at the output of the analog to digital converter are applied to the phase converter, where they are converted into binary system. From the output of the phase converter the voltages codes are applied to two codes multiplier input terminals, while to the third input the signal of harmonic numbering code is applied by the control block, which form harmonic codes set-point device. At the output of the codes multiplier square components codes of the phase voltages are composed. These codes are applied to the arithmetical device, where the symmetrical voltage components are calculated. The output information can be obtained by using the registration block.

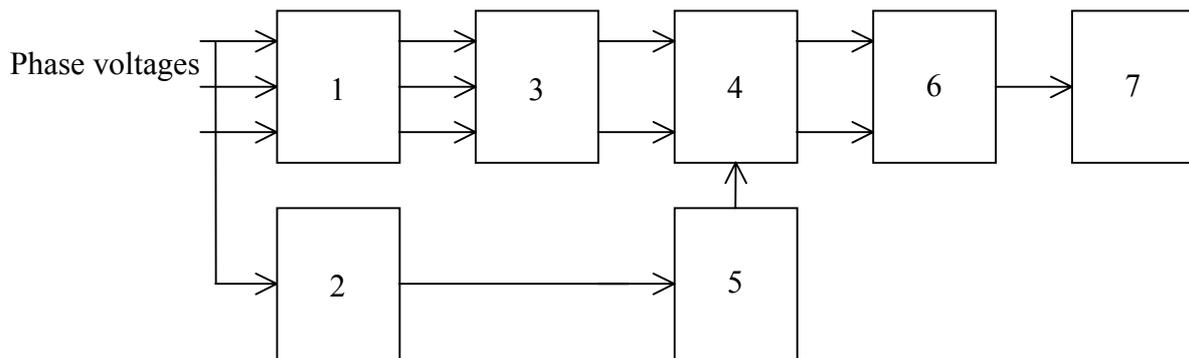


Figure 1: - Block diagram for voltage asymmetry measurement: 1- analog to digital converter, 2- control block, 3- phase converter, 4- multiplier, 5- set-point device, 6- arithmetical device, 7- registration block.

## References

1. Smith, R.L., and S. Herman. 1999. *Electrical Wiring Industrial*. Delmar Publishers, USA.
2. Jukov, V. 1994. *Short-Circuit in Electric System Nodes with Complex Load*. MEI, Moscow, Russia.

# **SCHEDULING AND FLOW CONTROL**



# SUMO (Simulation of Urban MObility)

## An open-source traffic simulation

Daniel Krajzewicz, Georg Hertkorn and Peter Wagner  
German Aerospace Centre,  
Institute for Transportation Research  
Rutherfordstr. 2, 12489 Berlin, Germany  
E-mail: Daniel.Krajzewicz@dlr.de,  
Georg.Hertkorn@dlr.de, Peter.Wagner@dlr.de

Christian Rössel  
Centre for Applied Informatics Cologne  
Weyertal 80, 50931 Köln, Germany  
E-mail: roessel@zpr.uni-koeln.de

### KEYWORDS

traffic simulation, microscopic, continuous, multimodal, open source, car-driver model, traffic research, road traffic

### ABSTRACT

As no exact model of traffic flow exists due to its high complexity and chaotic organisation, researchers mainly try to predict traffic using simulations. Within this field, many simulation packages exist and differ in their software architecture paradigm as well as in the models that describe traffic itself. We will introduce yet another system which, in contrast to most of the other simulation software packages, is available as an open-source program and may therefore be extended in order to fit a researcher's own needs and also be used as a reference testbed for new traffic models.

### INTRODUCTION

When trying to improve traffic, a valid model to work with is needed. Although some people may assume traffic can be described by departure times and routes with certain durations, traffic is highly conditioned by an individual's private wish for mobility – making up around 60% of traffic – and due to this, neither departure times nor fixed and earlier known routes are available. This is a great problem for modelling traffic itself. Especially the private transit leads to an impossibility of describing traffic by the use of mathematical formulas. Both, a modern human being's wishes to leave and arrive at certain places and at certain times on the one hand, and the movement of the vehicle on the street on the other, influence traffic and one another: the street network's work load depends on the drivers' departure times and determines the speed of movement. Vice versa the load affects the departure times of the drivers' as they wish to move fast and arrive at a certain time. Beside this, traffic is conditioned by values like the weather, the infrastructure within the region or other incidents affecting the system. This complexity yields in varied behaviour of the whole system where system means the generation of traffic and traffic itself, and as no valid mathematical models that take into account all these influences are

available, simulation is the only way to show weak points of the street network or predict its traffic. For this purpose, many simulation software packages were developed. Some of them were tested within the SMARTTEST-project (SMARTTEST 1999). Such traffic simulation software packages differ in their portfolios of modelled artifacts as well as in their usage paradigm: Some are conceived as Windows-using applications while other which are rather tools for traffic researchers, are simple command-line tools or programs. SMARTTEST also holds a list of features that such packages have. SMARTTEST was mostly dealing with microscopic simulations. Simulations of this type regard single vehicles as atomic parts, not the whole traffic flow itself (macroscopic) or single parts of the vehicles or drivers (sub-microscopic). Such vehicle models may be discrete in time and space using cellular automata or only discrete in time (Nagel and Schreckenberg 1992, Brockfeld et al 2001) or may even be fully continuous models. Figure 1 shows the difference between space-discrete and space-continuous simulations.

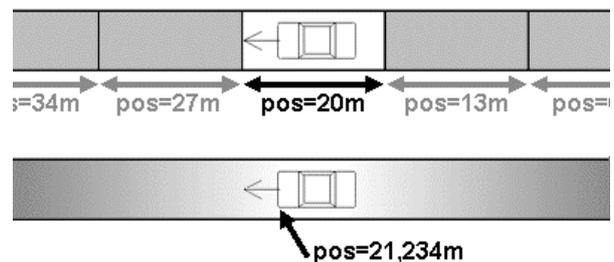


Figure 1: Space-discrete vs. Space-continuous simulation

From a researcher's point of view, when using the available simulation software packages, several problems may arise due to their availability as a ready-to-use software black-box, especially when commercial products are regarded. At first, one can not examine the underlying model of a simulation. Also, due to different software architectures, a comparison of different models' features (like simulation speed, its ability to describe the reality, etc.) is difficult if not impossible. Furthermore, such simulation tools can not be spontaneously extended by introducing ones own ideas such as new types of sensors, measurements or models.

To introduce a tool which accomplishes these not yet supported tasks, our institute – together with the Centre for Applied Informatics (Cologne, Germany) – is working on a traffic simulation software called SUMO („Simulation of Urban MObility“). In fact, this software is a continuous, microscopic and multi-modal traffic simulation and is – in spite of it’s name - also capable of modelling traffic on networks larger than single cities, e.g. highway networks, without any changes.

## THE SIMULATION

### Basic Paradigms

SUMO is conceived to simulate a traffic road network of the size of a city.

As the simulation is multi-modal, which means that not only are car movements within the city modelled, but also public transport systems on the street network, including alternative train networks, the atomic part of the simulation is a single human being. This human being is described by a departure time and the route he/she takes which again is made up of subroutes that describe a single traffic modality.

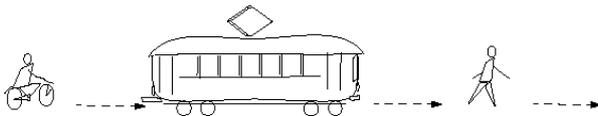


Figure 2: Multimodality

Thus, a simulated person may take his/her car to the nearest public transportation system station and continue his travel by other means of transport. Apart from movements using motorised vehicles, a person may also walk. Walking is not simulated at all but is modelled estimating the time the person needs to reach the destination. Figure 2 displays a such a compound route.

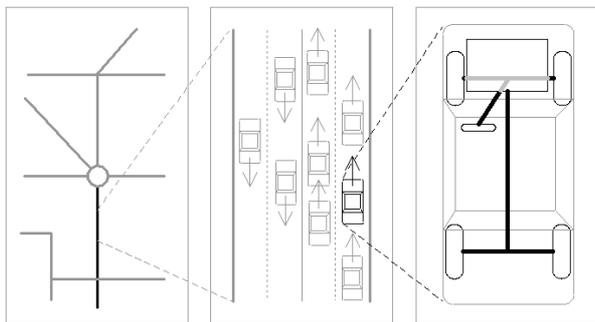


Figure 3: Different simulation classes (from left to right: macroscopic, microscopic, sub-microscopic)

The traffic flow is simulated microscopically. This means, that every vehicle that moves within the simulated network is modelled individually and has a certain place and speed. In every time step which has a duration of 1s, these values are updated in dependence to the vehicle ahead and the street network the vehicle

is moving on. The simulation of street vehicles is time-discrete and space-continuous. As our car-driver model is continuous - as the majority of car-driver models are - we decided to use this approach.

When simulating traffic, the street attributes, such as maximum velocity and right of way rules, are regarded.

### Features

In the current version – 0.7 – SUMO contains the following features:

- collision free vehicle movement
- different vehicle types
- multi-lane streets with lane changing
- junction-based right-of-way rules (junctions with streets having equal / different priorities, e.g. right-before-left)
- lane-to-lane connections
- a XML-raw-output containing information about the state of the net for every time step
- detectors with independent GnuPlot or CSV (comma separated value) – output
- input from XML-files which may be spread over several files for a better handling

### Car-Driver Model

The model used currently within SUMO is the Gipps-model extension (invented and described in: Krauß 1998, Janz 1998). It is capable of displaying main features of traffic like free and congested flow.

In each time step the vehicle’s speed is adapted to the speed of the leading vehicle in a way that yields to a collision-free system behaviour within the following simulation step(s).

This velocity is called the safe velocity  $v_{safe}$ , and is computed using the following equation:

$$v_{safe}(t) = v_l(t) + \frac{g(t) - v_l(t)\tau}{\frac{\bar{v}}{b(\bar{v})} + \tau}$$

$v_l(t)$ : speed of the leading vehicle in time  $t$

$g(t)$ : gap to the leading vehicle in time  $t$

$\tau$ : the driver’s reaction time (usually 1s)

$b$ : the deceleration function

To bind the acceleration to the vehicle’s physical abilities, the resulting “wished” or “desired” speed is computed as the minimum of the vehicle’s possible maximum velocity, the vehicle’s speed plus the maximum acceleration with the safe velocity computed as shown above, therefore a vehicle will not “drive” or “accelerate” faster than is possible for it:

$$v_{des}(t) = \min[v_{safe}(t), v(t) + a, v_{max}]$$

Further, the driver is simulated by assuming he is making errors and so fails to perfectly adapt to the

desired velocity. This is done by subtracting a random “human error” from the desired speed:

$$v(t) = \max[0, \text{rand}[v_{des}(t) - \epsilon a, v_{des}(t)]]$$

As the vehicle must not drive backwards, once again – after the previous computations – the maximum of the computed speed and zero must be taken and will be the vehicle’s final speed for the current time step.

This model is collision-free to allow simulations without any artifacts that arise from the imperfection of the underlying model (An impressive in-depth description of the model and the underlying assumptions and rules may be found in Krauß 1998 and Janz 1998).

## Traffic Lights

Traffic lights play an important role within the traffic management as they improve traffic flow. Apart from simple right-of-way rules, each simulated junction may also be a junction with traffic lights. As some junctions in Germany allow to ignore the red stoplight when turning right, an extension to the right-of-way rules regarding this is being implemented.

## Simulation Output

By now, two different outputs are available. The first is a so-called „raw“ output which contains all edges (streets) and all lanes along with the vehicles driving on them for every time step, where vehicles are described by their name, position and speed. This output is complete and may be used as input to post-processing tools for evaluation. However, a large simulation produces a nearly unmanageable amount of data, so other outputs have been invented.

The first processed output that may be generated by the simulation, is a log-file made by simulated detectors which may be positioned on a certain position of a certain lane. These detectors are a simulation of induct loops and are able to compute the flow, the average velocity on the lane, and other values. The results of this computation are written into a file using the CSV or the GnuPlot-format. Every detector has it’s own file.

## Software Development

Being a tool for traffic researchers, SUMO is designed to be fast and exact instead of trying to be a software that is pleasant to look at. So, although the implementation of a GUI will be one of our next tasks, the main program is meant to be started from a command line and produce an output which must be post-processed when one wants visual results. This prevents data arising from the GUI from slowing down the system, giving more memory and system time to the simulation itself.

SUMO is implemented in C++. During development, we try to use only standardised parts of this language. One of them is the STL which – when not coming directly with the compiler – may be additionally

downloaded as free implementations exist (STLPort 2001). The code is well documented and formatted and we will follow Ellementel-guidelines (Ellementel 1990-1992) that assure portability compatibility with most systems.

Due to this, our software is compilable using most platforms and we validated this for the following environments:

- Windows using MSVC++
- Solaris using SUN-C++-compiler and STL-Port
- Linux using gcc

## Simulation Benchmarks

The simulation is capable of simulating large cities like Berlin, Munich or Cologne on a normal desktop computer. Most of our tests were done on Intel PC’s running at 933MHz and having 256MB of memory.

The simulation is capable of simulating around 1Million vehicle movements per second depending on the network’s complexity. Further optimisations will follow.

## Extensibility

Hoping for the participation of other interested persons, we try to supply potential developers with all information needed to extend and modify the program. Beside things like intelligent traffic lights, models for cars equipped with ACC systems, etc., we hope to supply interested researchers with a common testbed for their microscopic models.

The documentation (SUMO 2002) shows how to extend the simulation – and other tools.

## ONGOING PROJECTS

Within our institute some projects use SUMO to validate their assumptions about new technologies. The following projects have already started using SUMO or will use it within the next months:

- A project from California, investigates whether the detector loops spread on different highway lanes may be used to predict jams.
- We try to validate our traffic flow predictions based on floating car data retrieved from Taxis in Berlin, Vienna and Würzburg .
- An internal project is trying to predict the advantages of new sensor technologies, on route prediction and the concluding traffic light phases.
- A project from the US tries to improve traffic within a highway off-ramp area.

## ADDITIONAL TOOLS

SUMO consists of more than a single application. Some other modules exist that allow to build assigned data needed for simulations and research.

The following modules are now being developed:

## Sumo-netconvert

Due to its high complexity, the SUMO's network description is not meant to be generated by a human user. Instead, we use this tool to convert common data like lists of edges and optional nodes into a complete SUMO-network.

During this process, SUMO-NETCONVERT reads in the available data, computes the needed input for SUMO and writes the results into a XML-file.

By now, four different input formats may be converted into SUMO-networks:

- simple XML-data containing edge types, nodes and edges
- CSV-data containing edge types, nodes and edges
- Cell-input files (Cell is a queue-simulation developed by the ZAIK)
- VISUM-networks

As SUMO is used within the INVENT-project, some further import functions will be implemented within this year: ArcView, VISSIM and possibly GDF where an indirect import of GDF is already possible by converting it into XML using a script and then using the generated XML-descriptions as input to SUMO-NETCONVERT.

The next figure shows which data may be computed from a simple list of nodes and edges. The flow is aligned from left to right and from top to bottom. The first step is to determine the priorities on the junction, the second to compute the relationships between the lanes and edges that may be reached, and within the third step, the destination edges are split among the incoming lanes. The computation is flexible and depends on the number of incoming and outgoing edges and their sizes as well as the priorities they have within the network and on the resulting type of junction.

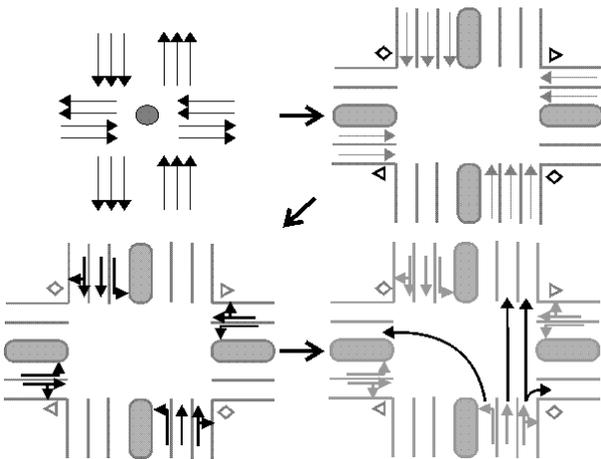


Figure 4: Conversion of simple, plain network data into a complete description

The next picture shows the need of such a net-building tool when a whole city is wished to be modelled. Neither the number of the streets nor the complexity of their relationships allows the processing of the network description by a human user.

SUMO-NETCONVERT is also responsible for the building of traffic light phases where either the priorities of the streets that build up a certain junction or their traffic flows are used. As in the real world – at least in Germany – phases of traffic are adapted individually for each junction and the certain phase data may not be available for all junctions, we use heuristics to generate a realistic output.



Figure 5: The city of Berlin build from high-quality digital map data

## Sumo-router

Beside the static part – the network – the simulation consists of moving vehicles. With the increase of the quality of simulations, the need to model a population's mobility has increased as well. In such cases, vehicles are not spread statistically over the network, instead a single person's daily plan consisting of routes with certain departure times is used. While data needed to describe the departure times and a route's origin and destination are given, the routes themselves must be computed. To avoid online-computation of these during the simulation, this computation is done using a separate module, the SUMO-ROUTER. This module reads the departure times, origins and destinations for a set of virtual humans that will be simulated, then computes the routes through the network itself using the well-known Dijkstra routing algorithm (Dijkstra 1959).

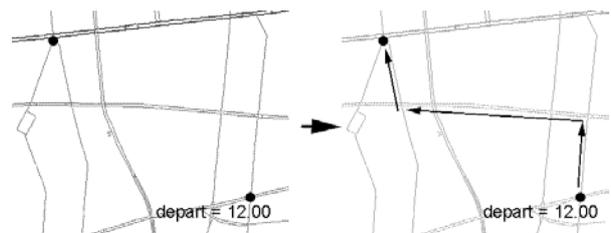


Figure 6: Routing in a small network

As the speed on the streets changes with the traffic amount and therefore the computation of routes using a network where the traffic is not yet known does not

regard the real-world situation, the routing will be done using the Dynamic User Equilibrium approach developed by Christian Gawron (Gawron 1998) where routing and simulation are repeated several times to achieve a real-world behaviour of drivers.

Furthermore, the router supports dynamic network load: the fact that the load on a depends on the time of day is also regarded.

## Planned

Other tools are planned due to the growing number of different projects investigating various topics and methods.

Some planned modules are:

- traffic light optimisation
- post-processing tools for raw-output evaluation

## CONCLUSIONS

We try to build and establish a common platform for traffic research by providing a simulation tool that may be applied by non-programming users in a simple way supporting them with methods and tools mostly needed when working on traffic simulations. Due to its high portability, the tool may be used on different operating systems.

Besides this, the platform is also extendable by others in order to allow them to improve the software and include ideas we have not taken into account.

Additionally, models originally implemented may be replaced by own methods allowing their comparison to the existing models in respect to the simulation quality and speed.

## FUTURE WORK

### Ongoing Work

Further work will be done as this simulation will be used by our institute for different purposes. Projects concerned about the verification of floating car data, the prediction of improvements on traffic foreseeing and traffic light optimisation through new sensors that observe traffic will increase the need to extend the simulation software by artifacts reflecting these devices and so will increase the simulation's complexity.

To trace and validate the simulation, a GUI will be implemented where the loaded map, together with vehicle movements and generalised information, will be displayed. To stay portable, the GUI will be implemented using the Qt-windowing library (TrollTech 2002) which is free to use and available for most systems including Windows, UNIX/Linux and Macintosh. Two different types of GUI support will be implemented. While the first application is the simulation itself extended by a windowing system, the second will be an interface that solely displays data coming from simulation module(s).

Some running projects have shown the need to integrate the software into other software packages or to build interfaces between the software and other

programming languages. This, too, will be investigated.

Also, as the amount of traffic to be simulated is growing and some research experiments need several simulations to compute a single value (for instance in case of traffic light optimisation where several timing schemes must be tested), the system will be extended to allow the usage of computer clusters.

## Participation

We also hope to gain help from other persons or institutes that are interested in traffic simulation and want to participate on SUMO's development by extending, improving or simply using and criticizing it. We highly invite you to visit the SUMO-pages at <http://sumo.sourceforge.net>.

## REFERENCES

- Brockfeld, E. et al. 2001. „Optimizing Traffic Lights in a Cellular Automaton Model for City Traffic“. In: Physical Review E 64, 056132
- E. W. Dijkstra. 1959. „A note on two problems in connexion with graphs“. In: Numerische Mathematik, 1:269-271.
- Directorate General VII - Transport of the European Commission. „SMARTTEST home page“. <http://www.its.leeds.ac.uk/projects/smartest/>
- Ellemtel. 1990-1992. Coding Standards Page. <http://membres.lycos.fr/pierret/cpp2.htm>
- Christian Gawron. 1998. „Simulation-Based Traffic Assignment“. Inaugural Dissertation.
- Stefan Janz. 1998. „Mikroskopische Minimalmodelle des Straßenverkehrs“. Diploma Thesis.
- Stefan Krauß. 1998. „Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics“. Hauptabteilung Mobilität und Systemtechnik des DLR Köln. ISSN 1434-8454
- K. Nagel, M. Schreckenberg. 1992. Journal of Physics I 2, 2221
- STLPort. 2001. Company Homepage. <http://www.stlport.org/index.html>
- SUMO: G. Hertkorn, D. Krajzewicz, C. Rössel. 2002. SUMO Homepage. <http://sumo.sourceforge.net>
- TrollTech. 2002. Troll Tech Homepage. <http://www.trolltech.com>

## BIOGRAPHY

Born in Bydgoszcz, Poland, 1972, **DANIEL KRAJZEWICZ** has finished his study of computer science at the Technical University in Berlin by the middle of the year 2000 with artificial intelligence and computer graphics as main topics. After work on text classification he changed to the Institute for Transportation Research of the German Aerospace Centre where he now works on a cognitive driver model and an open-source urban traffic simulation.

# PERFORMANCE IMPLICATIONS OF BUFFER OVERFLOW AS A KEY (DISTURBING) ELEMENT IN THE FLOW CONTROL OF A JOB/FLOW SHOP FACILITY

Erland Hejn Nielsen  
Department of Management Science and Logistics,  
The Aarhus School of Business, Denmark  
Fuglesangs Allé 4, DK-8210 Aarhus V  
Denmark  
E-mail: [ehn@asb.dk](mailto:ehn@asb.dk)

## KEYWORDS

Flow/Job shop Scheduling; Buffer limitations; Re-entrant systems; CONWIP; Numerical simulation.

## ABSTRACT

From time to time most real life job/flow shops will experience more or less severe buffer problems in relation to the operation at its various stations in the shop. Due to temporary non-predictable variations in the job flow intensity the assigned buffer space in front of a given critical station will now and then turn out simply to be quite insufficient. Unless they exercise either a consequent JIT-control system or a CONWIP set-up of some sort where the latter strictly speaking is a system with the overflow concentrated at the entrance of the production facility, this consequently leads to temporary buffer overflow situations. This paper will investigate the impact of buffer overflow on several key performance measures of relevance to the overall operation of a job/flow shop facility. Several simple job/flow shop structures will be considered in this paper including re-entrant system set-ups (a re-entrant system is a system where some of the job-routes revisit a given station more than once).

## INTRODUCTION

Many roads of events can lead to a situation of frequent buffer overflows. Business is good and order books are "full" and increasing or business can be depressed, and costs have to be cut which may imply that some space has to be freed to serve other purposes. Both cases may lead to a situation where space consequently gets limited or put in a different perspective "gets very well utilized", the job/flow shop will either get partly blocked occasionally or there will be occasional buffer overflows. A blocked station is as harmful to productivity as an idle station given that operations at the various stations are subject to significant time variations that cannot be foreseen.

On the other hand permitted overflow of buffers tends to introduce some chaos into the operation of the job/flow shop in that when overflow occurs, items are often placed at random in places where internal transportation etc. also

has to take place, however, typically in places where they simply constitute a practical nuisance and often a security hazard. So the basic hypothesis will, of course, be that either way the result of not enough buffer space available in a production facility can easily result in a severe loss of productivity as the controlled flow is disturbed by the randomly occurring overflow phenomenon. Tight physical constraints on production calls for serious production planning and control and for instance the subject of choice of scheduling policy should be thought to be important. However, it is my impression that many shop managers do not seem to pay very much attention to the actual scheduling policy employed as if it did not matter that much.

## THE EXPERIMENTAL SET-UP

The structures considered will be simulated in order to obtain results on the performance measures "Average Unit System Time" and "Units Processed per Time Unit" as functions of given regular buffer sizes allocated to the various stations. Small regular buffer sizes will obviously imply frequent situations of buffer overflow, whereas large regular buffer sizes obviously will not, so the chaotic impact on operations will be most noteworthy for small regular buffer levels. The interesting question is clearly what will be a necessary and thereby relevant regular buffer size given an overall traffic-intensity in the job/flow shop where the possible chaotic disturbance element, due to buffer overflow, becomes almost insignificant? The allocated regular buffers will be operated on a set of priority schemes that is assumed to be possible choices by a shop management, whereas overflow "buffers" are operated on a strictly random basis. Further, if present, the removal of overflow "buffers" will always take priority over the regular ones.

The actual set of simulation experiments that will be conducted in this work will be based on

- a simple M/M/1/n queuing system controlling the regular buffer by the SPT (Shortest Processing Time) operating discipline.
- a simple 2-station/5-class re-entrant system (see figure 1) as used by Dai in his very illustrative



priority of queue 2 over 4 on machine 2. And vice versa with respect to a seemingly well functioning scheduling principle like LBFS, which by the way can be proven in general to be globally long-run stable if not the most optimal in all situations.

The measure UPTU does not show much of a variation in table 2. The explanation is that despite the fact that severe congestion takes place as the 100K units pass the re-entrant system, the 100K units are far from delivering steady-state results, where UPTU should be expected to be limited by the virtual bottleneck and converge towards  $1/1.2 = 0.83$ ! Because this is not the case the results for AUST are also not truly steady-state results, however, the structural tendency seems clear.

A slightly different experimental set-up where the arriving units do not arrive individually according to a poisson-process, but arrive simultaneously in a bunch at time zero, turns out to make the UPTU measure much more active. It should be expected to lie between a worst performance of  $1/1.2 = 0.83$  and a best performance of  $1/0.8 = 1.25$ .

Table 3: Units Processed per Time Unit (UPTU) - for a 2-station/5-class re-entrant queuing system with limited buffer capacity (BC) and permission to overflow temporarily and the whole batch arrives at time zero. The number of replications is 20, traffic intensity is 0.8 on each of the 2 machines with machine 1 processing times with respect to class(1,3,5) = (ex(0.1),ex(0.1),ex(0.6)) and with machine 2 processing times with respect to class(2,4) = (ex(0.6),ex(0.2)). The class priority on machine 1 is fixed to (5,3,1) in all cases.

BC	Batch=5K units		Batch=15K units	
	Class Priority M2 "VB"	(4,2) "LBFS"	(2,4) "VB"	(4,2) "LBFS"
1	1.09 (±0.01)	1.09 (±0.01)	1.09 (±0.01)	1.09 (±0.01)
50	1.09 (±0.01)	1.09 (±0.01)	1.09 (±0.01)	1.09 (±0.00)
Inf.	0.83 (±0.01)	1.24 (±0.01)	0.83 (±0.01)	1.25 (±0.00)

So the "VB" case produces its worst at BC = "Infinity" and, as should be expected, "LBFS" produces its best at BC = "Infinity". The RAN scheduling principle is ruling whenever BC is close to 0 and produces an UPTU of 1.09.

Getting back to the observation that many shop managers do not seem to pay much attention to the actual scheduling policy employed as if it did not matter that much, one simple explanation could easily be that they have tried different scheduling principles and found that sometimes UPTU is greater than one and sometime it is lower, but always in the vicinity of one! For lumped batch-arrivals combined with  $BC < 50$  and embedded re-entrant systems

in the overall job/flow shop structure (which is not that unusual), the overflow scheduling principle will dominate.

Finally, I turn to the well-known CONWIP system set-up. Not to study the steady state case, but, in many situations, the equally realistic one of lumped batch-arrivals of a moderate size. The system is an 11-stations flow line with the first station as a gate to the CONWIP part of the shop. The individual queue priority discipline is SPT.

Table 4: Average Unit System Time (AUST) - for an 11-station-TOTAL/10-station-CONWIP system with limited buffer capacity (BC) equal for all stations and permission to overflow temporarily and the whole batch = 5K arrives at time zero. The number of replications is 20 and the traffic intensity on each of the 11 machines is 0.9

BC	Max. units in shop (CONWIP)			
	1	10	50	Inf.
1	9.90 (±0.02)	16.78 (±0.05)	43.38 (±0.015)	1348.67 (±11.30)
10	9.90 (±0.02)	15.66 (±0.05)	37.61 (±0.17)	1258.11 (±15.09)
50	9.90 (±0.02)	15.66 (±0.05)	37.18 (±0.17)	1206.68 (±16.78)
Inf.	9.90 (±0.02)	15.66 (±0.05)	37.18 (±0.17)	1074.08 (±12.64)

Table 5: Units Processed per Time Unit (UPTU) - for an 11-station-TOTAL/10-station-CONWIP system with limited buffer capacity (BC) equal for all stations and permission to overflow temporarily and the whole batch = 5K arrives at time zero. The number of replications is 20 and the traffic intensity on each of the 11 machines is 0.9

BC	Max. units in shop (CONWIP)			
	1	10	50	Inf.
1	0.10 (±0.00)	0.53 (±0.00)	0.74 (±0.00)	1.04 (±0.01)
10	0.10 (±0.00)	0.56 (±0.00)	0.80 (±0.01)	1.06 (±0.00)
50	0.10 (±0.00)	0.56 (±0.00)	0.80 (±0.00)	1.07 (±0.00)
Inf.	0.10 (±0.00)	0.56 (±0.00)	0.80 (±0.00)	1.08 (±0.01)

The findings can be interpreted as follows. Looking carefully at table 4 and 5 where SPT is dominating below the diagonal we see that for any given CONWIP value, an increasing BC, which in this case means increasing dominance of SPT, both the AUSTS measure as well as the UPTU measure improve. I admit that one might need some theoretical background knowledge to see this clearly from the results especially in table 5. But then again the results in table 4 and 5 seem to indicate that for lumped batches of size 5000 units the actual buffer size at the individual stations given  $CONWIP < 50$  is not of much importance!

## CONCLUDING REMARKS

Contrary to what common intuition tells us, that unorganised buffer overflow should be a problem, this analysis seems to put it somewhat differently and allow me to formulate the message somewhat provocatively

- buffer overflow is not always a problem, on the contrary it can occasionally be the cure!
- when buffer overflow has a negative impact on performance, my findings seems to indicate that it most likely has a practical significance for relative tight buffer area set-ups ( $BC < 10$ ).
- increasing structural complexity or poor choice of part of the overall scheduling policy, seems to overrule the importance and question of an exact determination of the size of buffer limitation, if present at all

The set of experiments, that have lead to the results in this paper and especially to the last of my three concluding remarks above, has also been conducted with the SPT queue discipline replaced by the FIFO discipline as the local queue discipline. As size of buffer limitation loses importance, so does the actual choice of local queue discipline as well.

The possible fading importance of SPT or more generally any local queue discipline set-up, whenever overall structure gets increasingly more complex, stands in somewhat contrast to the general accepted understanding of the SPT loading rule as being one of the best if not the best all-round choice. The SPT rule has to a great extent gained its reputation of superiority by generalising from the simpler situations where its virtue can be proven exactly to the more complex ones where some simulation results support that view according to Anderson (Anderson 1994) and Yao (Yao 1994).

However, emphasis on the element of job re-entrance and its possible consequence on the choice of loading rule have not really been present until the early nineties, when a series of interesting results on pure re-entrant flow shop structures emerges. It now seems that alternative loading rules as for instance the LBFS (Last Buffer First Served) due to its strong and general long run stability property attracts quite some interest. The LBFS scheduling policy can also be seen as the natural choice if one subscribes to lean thinking (Womack and Jones 1996).

Some may finally argue that pure re-entrant systems are not very common in practice, but from my discussions with colleagues that on a daily basis visit many different job/flow shop production facilities I have learned that though pure re-entrant systems may not be seen very often, embedded re-entrant systems are quite common. So even if my results in this paper are based, amongst others, on a

simple and pure re-entrant system structure, I think they support a view that complexity have many dimensions attached and is not limited to simple scaling of a simple M/M/1 system as for instance in unidirectional serial and/or parallel set-ups without constraints of any kind either on buffers or on structure.

## BIBLIOGRAPHY

- Anderson, E.J. 1994. *The Management of Manufacturing: Models and Analysis*. Addison-Wesley.
- Dai, J. G. 1995. "Stability of open multiclass queuing networks via fluid models". In *Proceedings of the IMA workshop on stochastic networks*. Frank Kelly and Ruth Williams (editors), Springer-Verlag, New York, p.71-90
- Nielsen, E.H. and D. Simons. 2000. "Lean thinking in systems with non-negligible process variability". In *Proceedings of the Logistics Research Network 5<sup>th</sup> Annual Conference*, The Cardiff Business School, Lean Enterprise Research Centre, p.438-446
- Kumar, S. and P.R. Kumar. 1995. "The last buffer first served priority policy is stable for stochastic re-entrant lines", Research Report.
- Womack, J. and D.T. Jones. 1996. *Lean thinking*. Simon and Schuster.
- Yao, D. D. 1994. "Stochastic Modeling and Analysis of Manufacturing Systems". Springer-Verlag, p.325-360

## AUTHOR BIOGRAPHY

**ERLAND HEJN NIELSEN**, associate professor, was born in Padborg, Denmark, and studied mathematics and economics at the University of Aarhus, Denmark, from where he obtained his MSc(Econ/Math) degree in 1980. After a period of three years as EU financed contract researcher at the Danish National Energy Laboratories RISØ he started his present employment at the Aarhus School of Business, Department of Management Science and Logistics.

# SCHEDULING JOBS ON PARALLEL BATCH PROCESSING MACHINES USING DISPATCHING RULES AND MACHINE LEARNING TECHNIQUES

Lars Mönch  
Technische Universität Ilmenau  
Institut für Wirtschaftsinformatik  
PF 100565  
98684 Ilmenau, Germany  
E-mail: Lars.Moench@tu-ilmenau.de

Peter Otto  
Technische Universität Ilmenau  
Institut für Automatisierungs- und Systemtechnik  
PF 100565  
98684 Ilmenau, Germany  
E-mail: Peter.Otto@tu-ilmenau.de

## KEYWORDS

ATC Dispatching Rule, Scheduling, Neural Networks, Inductive Decision Trees, Semiconductor Manufacturing

## ABSTRACT

In this paper we consider the solution of batch scheduling problems with incompatible job families on parallel machines. A simple solution to the problem of minimizing the performance measure total weighted tardiness could be based on the apparent tardiness cost dispatching rule and its extension to batching problems in combination with a deterministic forward simulation. The setting of a look ahead parameter is crucial for the use of this dispatching rule. Therefore, we suggest two different machine learning techniques, i.e., neural networks and inductive decision trees in order to overcome this problem. The performances of the two different approaches are compared and the results of simulation experiments are presented.

## INTRODUCTION

Recently, the electronics industry has become the largest industry in the world. The most important area in this industry is the manufacturing of integrated circuits (IC). In the past, sources of reducing costs were decreasing the size of the chips, increasing the wafer sizes and improving the yield, simultaneously with efforts to improve operational processes inside the semiconductor wafer fabrication facilities (wafer fab). Currently, it seems that the improvement of operational processes creates the best opportunity to realize the necessary cost reductions. Therefore, the development of efficient planning and control strategies is highly desirable in the semiconductor manufacturing domain. In order to create, implement and test the required strategies, it is necessary to take the new opportunities of information technology, especially soft computing techniques, into account.

In this paper we present a method to support the solution of batch scheduling problems with incompatible job families on parallel machines (a batch is a set of jobs that

are processed simultaneously on the same machine). We assume, that jobs of different families cannot be processed together. Problems of this type arise, for example, in the diffusion process in semiconductor manufacturing and have great practical relevance (see Atherton and Atherton 1995 for details of the wafer fabrication process).

A simple solution approach could be based on the use of the ATC (apparent tardiness cost) dispatching rule and its extension to batch processing problems in combination with a deterministic forward simulation. Even in its simplest form the ATC rule requires the choice of a look ahead parameter  $k$ . The proper choice of this look ahead parameter is crucial for the quality of the schedule derived using the dispatching heuristic. Simple heuristics for the parameter setting (cf. Lee and Pinedo 1997 and Park et al. 2000) fail in certain situations or result in sub optimal solutions.

We suggest two different machine learning techniques, namely neural networks and inductive decision trees to derive proper values for the look ahead parameter  $k$ . We extend the work of Park et al. (2000) to the more general batch scheduling problem.

However the application of decision trees is new in this situation. We compare the performance of the two different approaches and present results of computational experiments.

The paper is organized as follows. The next section describes the problem under consideration. Then we suggest two different machine learning approaches for choosing the look ahead parameter for the BATC dispatching rule. We finish the paper by presenting results of simulation experiments with the neural network and the decision tree approach.

## STATEMENT OF THE PROBLEM

We will use the following notations throughout the rest of the paper.

1. There exist  $f$  different families of products to be processed.
2. There are  $m$  homogenous machines in parallel.

3. There are  $n$  jobs waiting to be scheduled in front of the machines.
4. Job  $i$  of family  $j$  is denoted by  $ij$ .
5. The due date of job  $ij$  is represented by  $d_{ij}$ .
6. The notation  $c_{ij}$  is used for the completion time of job  $ij$ .
7. The weight of job  $ij$  is denoted by  $w_{ij}$ .
8. The size of a single batch, i.e., the capacity of one of the homogenous machines, is  $B$ .
9. The processing time for jobs of family  $j$  is denoted by  $p_j$ .
10. The weighted tardiness of job  $ij$  is calculated by

$$w_{ij}T_{ij} := w_{ij} \max(c_{ij} - d_{ij}, 0) = w_{ij} (c_{ij} - d_{ij})^+.$$

Scheduling problems are usually represented in the form  $\alpha/\beta/\gamma$  (cf. Pinedo 2002). Here, the  $\alpha$  field describes the machine environment, the  $\beta$  field is used as a notation for process specifics and the  $\gamma$  field indicates the used performance measure. Using this notation the problem under consideration can be represented in the form

$$P_m | \text{batch, incompatible} | \sum w_j T_j.$$

The problem is NP-hard, hence heuristics have to be used to obtain an efficient solution.

The ATC heuristic is a composite dispatching rule that tends to schedule jobs with higher priority indices first, given by the combination of two other rules: MS, Minimum slack (time until due date) and WSPT, Weighted Shortest Processing Time (cf. Pinedo 2002). The ATC index for job  $ij$  at time  $t$  is given by:

$$I_{ij}(t) := \frac{w_{ij}}{p_j} \exp\left(\frac{-(d_{ij} - p_j - t)^+}{k\bar{p}}\right), \quad (1)$$

where  $\bar{p}$  denotes the average processing time of all remaining jobs and  $k$  is the look ahead parameter. The common ATC heuristic can be extended to the BATC (Batched Apparent Tardiness Cost) heuristic (cf. Devapura et al. 2000 and Perez et al. 2002) in the following way:

$$BATC_{ij} := \sum_{i \in B_{lj}} I_{ij}(t), \quad (2)$$

where  $BATC_{lj}$  is the BATC index for batch  $l$  of family  $j$ . By  $B_{lj}$  we denote the batch  $l$  of family  $j$ . The batch with the highest BATC index (among the remaining jobs) is formed in each step of the deterministic forward simulation. The parameter  $k$  is a free parameter in (1). It ranges between 0.5 and 5.0.

The proper choice of  $k$  is important in order to ensure schedules of high quality with respect to the performance measure  $G := \sum w_j T_j$ . We indicate the dependency of  $G$  on the parameter  $k$  by using the notation  $G(k)$ .

## CHOICE OF THE LOOK AHEAD PARAMETER

In (Lee and Pinedo 1997) the authors derived several formulas in order to determine the look ahead parameter  $k$ . A parametric adjustment method for the ATC rule is described by Chao and Pinedo (1992). A neural network approach is presented for scheduling jobs with sequence dependent setup-times by Park et al. (2000). We extend this approach to the BATC rule.

The  $k$  value depends on the following quantities:  $R, T, \mu$ . These quantities are defined as follows. The factor  $R$  represents the range of the due dates. For its calculation we use the following expression

$$R := \frac{1}{\bar{d}} \sqrt{\frac{12}{n-1} \sum (d_{ij} - \bar{d})^2}. \quad (3)$$

The second factor  $T$  is used to measure the tightness of the schedule and can be expressed as

$$T := 1 - \frac{\bar{d}}{\mu\bar{p}}, \quad (4)$$

where we set  $\bar{d} := \frac{1}{n} \sum d_{ij}$ . Note that  $\mu\bar{p}$  is a (rough) estimator for the makespan. The third factor  $\mu$  is called batch-machine factor and is given by

$$\mu := \frac{n}{mB}. \quad (5)$$

It describes the average number of batches per machine. Note, that the factors  $R, T, \mu$  in our study are modifications of the factors suggested by Perez et al. (2002).

## Neural Network Approach

We use a multi layer perceptron neural network with a sigmoid function for activation purposes. The hidden layer consists of seven nodes. The values of the network weights are learnt through repeated examination of training cases. We use the following input nodes:  $R, T, \mu$ . The  $k$  values are given by the output layer of the neural network. The used neural network is given in Figure 1

We use the NeuralWorks Professional II/Plus tool from NeuralWare for carrying out the neural network experiments.

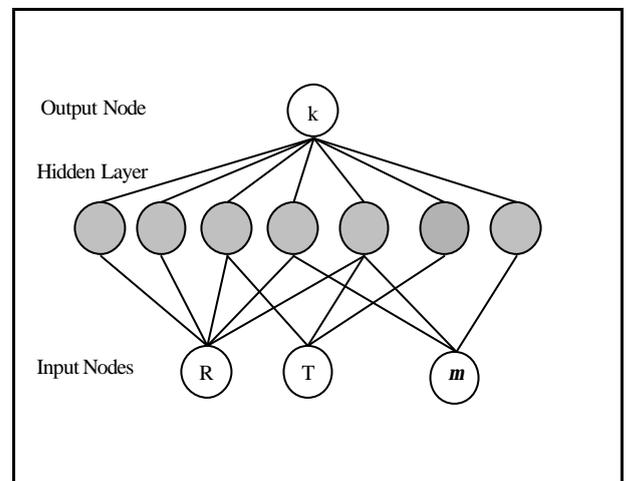


Figure 1: Configuration of the Used Neural Network

### Inductive Decision Tree Approach

The second approach for choosing proper value for  $k$  is an inductive decision tree (cf. Otto 1995 and Quinlan 1993 for more detailed information on decision trees). The basic idea of the algorithm consists in the generation of an initial decision tree from a set of training cases. For that purpose we divide the range of  $k$ , i.e., the interval  $[0.5, 5.0]$ , in equidistant classes.

The decision tree is a data structure containing leaves, indicating a class, and decision nodes specifying tests that have to be carried out on the attributes (i.e., the input factors) of the test case in order to branch in other sub trees. The structure of a learned decision tree is a collection of disjoint hyper cubes that are created by partitioning the domains of attributes into intervals. We used the ID3 algorithm (cf. Quinlan 1993) in order to derive the tests. A rule is basically a traversing of the decision tree. We obtained 948 different rules. We used our own developed software for constructing the trees.

The decision tree algorithm is sensitive to the number of classes of  $k$ . Considering too many classes has the drawback that too much noise is incorporated into the decision tree. Therefore, we get poorer results because of the described data over fitting.

### Complexity Aspects of the Two Approaches

The training phase of the neural network takes several minutes. However, the usage of the learned neural network requires only seconds. A successful training phase of the inductive decision tree is within seconds possible. The application of the (non optimized) decision tree requires the same computational effort as in the case of the neural network. However, it takes some time and experience to perform the optimization of the Fuzzy membership functions that correspond to the attribute intervals of the decision tree.

### EXPERIMENTAL DESIGN AND RESULTS

We started from the test data described by Perez et al. (2002) and generate test cases according to the following factor setting:

$$R = (0.50, 0.75, 1.00, \dots, 2.50)$$

$$T = (0.30, 0.35, 0.40, \dots, 0.90)$$

$$\mu = (10, 15, 20, \dots, 60).$$

We consider all combinations of these factors. We use three incompatible families. The processing time of the families are uniformly distributed. Batch sizes of one, two four and eighth are considered in our studies. The due dates of job  $ij$  are derived from a uniform distribution over the interval  $[\bar{d}(1 - R/2), \bar{d}(1 + R/2)]$ . The weights of the jobs are generated from a uniform distribution over  $[0, 1]$ . For each factor combination we compute the optimal value

for  $k$ . We take the average of the results of ten independent simulation runs in order to get statistically significant results for  $k$ . In Table 1 we find the results for the neural network (NN) approach. We measure in each case the mean square error between the  $k$  values given by the machine learning approach and the optimal value for  $k$ . In Table 2 and Table 3 we present the corresponding results for the inductive decision tree (IDT) approach. Here, we also present how the error depends on the number of fixed classes for the  $k$  parameter. In the case of 20 classes the corresponding decision tree contains 948 different rules. In Figure 2 we present results for the IDT approach using training data. The corresponding results for the test data case are shown in Figure 3. In Figure 4 we show results for test data for the developed neural network.

We compare the values of the objective function  $G$  obtained by using the optimal value  $k_{opt}$  and the value  $k_{IDT}$  obtained by the IDT approach. We determine the (normalized) square error between  $G(k_{opt})$  and  $G(k_{IDT})$ . The corresponding results are presented in Table 4. In order to obtain the results of the first row we used  $B = 1$ , for the results in the second row we used  $B = 8$ . The error is small in both causes; the error in the second row of Table 4 is slightly larger because of the batch decision.

Table 1: Mean Square Error for Neural Network Approach

Nodes of the Hidden Layer	Error Training Data	Error Test Data
4	0.0680	0.0829
7	0.0535	0.0717

Table 2: Mean Square Error for Inductive Decision Tree Approach

Number of Considered Classes	Error Training Data	Error Test Data
5	0.0641	0.1020
10	0.0450	0.0940
15	0.0415	0.0941
20	0.0396	0.0797
25	0.0399	0.0914

Table 3: Mean Square Error for Inductive Decision Tree Approach (Optimized Tree)

Number of Considered Classes	Error Training Data	Error Test Data
5	0.0645	0.1020
10	0.0422	0.0807
15	0.0383	0.0784
20	0.0354	0.0679
25	0.0360	0.0752

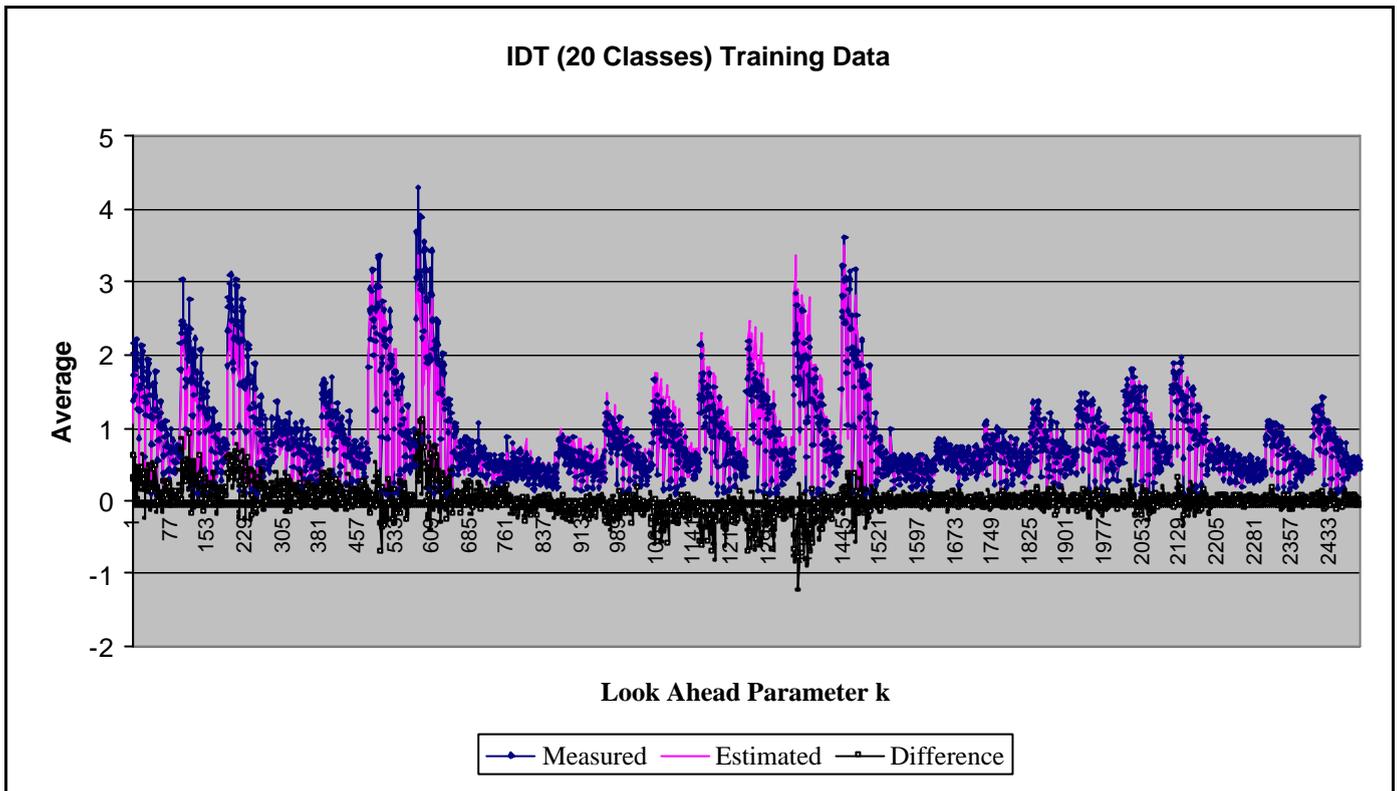


Figure 2: Mean Square Error for an Inductive Decision Tree: Training Data

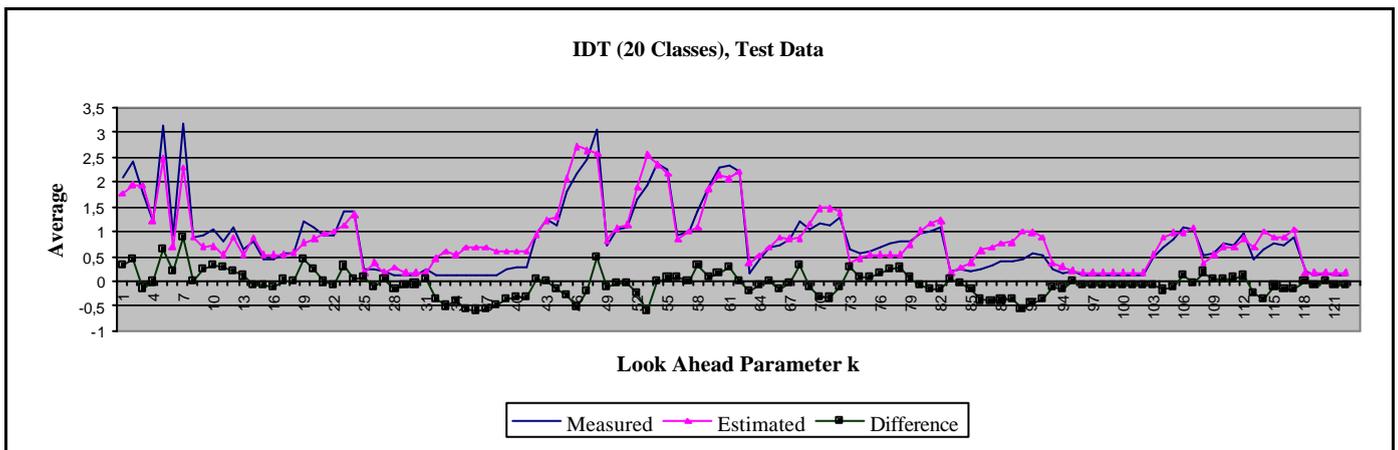


Figure 3: Mean Square Error for an Inductive Decision Tree: Test Data

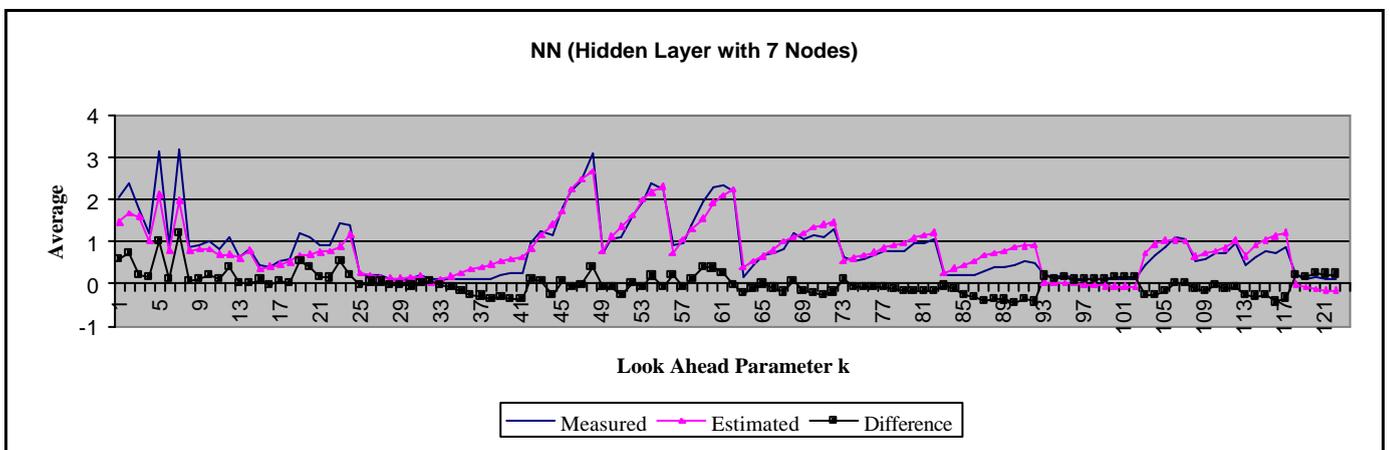


Figure 4: Mean Square Error for a Neural Network: Test Data

Table 4: Mean Square Error between  $G(k_{opt})$  and  $G(k_{IDT})$

T	R	$\mu$	Error
0.3271	0.5431	37.5000	0.0013
0.7777	0.6666	11.2500	0.0016

We conclude from our experiments that the optimized version of the inductive decision tree performs better than the neural network. Especially, it requires less effort to derive the tree from training data as in the case of the neural network. Therefore, we conclude that the usage of inductive decision trees in the given situation has some advantage over the usage of neural networks.

## CONCLUSIONS AND OUTLOOK TO FUTURE RESEARCH

In this paper we suggested two different machine learning approaches, i.e., a neural network approach and a decision tree approach, in order to choose the look ahead in the ATC dispatching rule applied to batch scheduling problems. We described the used algorithms in detail. We presented the results of computational experiments and compared the performance of the two different algorithms. A possible direction of future research is to be focus on the incorporation of ready times into the algorithm. Here, instead of using the pure ATC rule the X-dispatch ATC rule (cf. Morton and Ramnath 1995) has to be considered.

**ACKNOWLEDGEMENT:** This research was partially supported by a research grant of the Deutsche Forschungsgemeinschaft (DFG). The first author gratefully acknowledges the support of the Semiconductor Research Corporation (2001-NJ-880).

## REFERENCES

- Atherton, L.F. and R. W. Atherton. 1995. *Wafer Fabrication: Factory Performance and Analysis*. Kluwer Academic Publishers, Boston, Dordrecht, London, 1995.
- Chao, X. and M. Pinedo. 1992. *A Parametric Adjustment Method for Dispatching*. Technical Report, Department of Industrial Engineering, Columbia University, New York.
- Devpura, A., J. W. Fowler, M. W. Carlyle, and I. Perez. 2000. Minimizing Total Weighted Tardiness on Single Batch Process Machines with Incompatible Job Families. In *Proceedings Symposium on Operations Research*, 366-371.
- Lee, Y.-H. and M. Pinedo. 1997. Scheduling Jobs on Parallel Machines with Sequence-Dependent Setup Times. *European Journal of Operational Research*, 100, 446-474.
- Morton, T. E. and P. Ramnath. 1995. Guided Forward Search in Tardiness Scheduling of Large One Machine Problems. In *Intelligent Scheduling Systems*, eds.: D. E. Brown, W. T. Scherer, Kluwer Academic Publishers, Hingham, MA.
- Otto, P. 1995. Fuzzy Modelling of Nonlinear Dynamic Systems by Inductive Learned Rules. In *Proceedings 3rd European Congress on Intelligent Techniques and Soft Computing (EUFIT' 95)*, 858-864.

- Park, Y., S. Kim, and Y.-H. Lee. 2000. Scheduling Jobs on Parallel Machines Applying Neural Network and Heuristic Rules. *Computers & Industrial Engineering* 38, 189-202.
- Perez, I., J. W. Fowler, and M. W. Carlyle. 2002. Minimizing Total Weighted Tardiness on a Single Batch Process Machine with Incompatible Job Families. Submitted to *Computer and Operations Research*.
- Pinedo, M. 2002. *Scheduling Theory, Algorithms, and Systems*. Prentice Hall, Second Edition.
- Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, Inc. 1993.

## AUTHOR BIOGRAPHIES

**LARS MÖNCH** is an Assistant Professor in the Department of Information Systems at the Technical University of Ilmenau, Germany. He received a master's degree in applied mathematics and a Ph.D. in the same subject from the University of Göttingen, Germany. After his Ph.D. he worked at Softlab GmbH in Munich in the area of object-oriented software development for two years. His current research interests are in simulation-based production control of semiconductor wafer fabrication facilities, applied optimization and artificial intelligence applications in manufacturing. He is a member of GI (German Chapter of the ACM), GOR (German Operations Research Society) and SCS. His email address is <Lars.Moench@tu-ilmenau.de>.

**PETER OTTO** is an Associate Professor in the Department of System Analysis at the Technical University of Ilmenau, Germany. He received a master's degree in computer science from the Moscow State University and a Ph.D. in the same subject from the Technical University of Ilmenau, Germany. His current research interests are in machine learning techniques applied to control problems. His email address is <Peter.Otto@tu-ilmenau.de>.

# THE CIRCULATING MULTISEQUENCER SCHEDULING ALGORITHM IN A REAL-TIME DISTRIBUTED TRANSACTIONAL SYSTEM: MATHEMATICAL MODEL AND ITS VALIDATION BY SIMULATION

Leïla Azouz Saïdane  
Ecole Nationale des Sciences de l'Informatique (E.N.S.I.)  
Campus universitaire de la Manouba - 2010 – Ariana - Tunisia  
*email:leila.saidane@ensi.rnu.tn*

## KEYWORDS

Queuing systems, real time systems, distributed systems, scheduling, EDF, transaction.

## ABSTRACT

In a real-time distributed transactional system, customers generate transactions, which should be scheduled to be executed on different servers. To schedule these transactions the circulating multisequencer algorithm has been considered to obtain a global view of the system. In this paper, a mathematical model is developed to obtain the average stay time of a transaction within the system. This model introduces a bulk arrival M/G/1 station with K classes of customers where bulks are considered according to FIFO discipline, customers (actions) are scheduled according to EDF within a group and the algorithm is processed according to the HOL discipline. To validate this mathematical model, mathematical and simulation results are confronted. This study shows that the mathematical model is realistic in most situations.

## 1. INTRODUCTION

A real-time distributed transactional system is defined as being a system composed of customers and servers interconnected by a communication network (customers/servers network) (Lelann 1994). The system is distributed, there is no global memory and the servers can communicate only by exchanging messages over a network of servers, which is generally a local area network. The transactions are generated by the customers and must be executed by the servers. The system is real-time, that is, the transactions have a temporal constraint. Their executions must be finished before their deadlines. A transaction is composed of actions; an action is considered as a program, which must be run completely by a single server.

One of the points of interest elaborated by the project team REFLECS of INRIA France, is the development of algorithms, which satisfy the properties of serialisability and punctuality of the transactions. The principle which has been retained is to apply scheduling algorithms recognised as being optimal in the centralised case (local), like NP-EDF (Non-Pre-emptive- Earliest- Deadline First) (Lelann 1994, Lelann 1990, Georges et al 1995). In order to do this, a global view of the transactions to be executed by the different servers will have to be obtained.

To that end, algorithms are being developed: the circulating multisequencer algorithm, the consensus algorithm and the multi-token algorithm (Anceaume et al. 1995; Mettali 1998).

These algorithms enable one or many servers to have the global view of the system, and so to realise the scheduling of the transactions. They diffuse afterwards the result of this scheduling to the different servers.

Each server maintains two waiting queues:

-An unscheduled queue, where are stocked the transactions received from customers, and which will be called customer queue

-And a scheduled waiting queue, where are stocked the scheduled actions waiting to be run. This queue will be called execution queue.

The traffic is predefined. The generation of transactions is such that, during each period D, the customers generate N transactions ( T1..... TN ). K types of transactions are defined. Each transaction (i) has a deadline  $E_i$  and is activated  $n_i$  times such as  $\sum_{i=1}^K n_i = N$ . The traffic is

characterised by the matrix (a<sub>ij</sub>) with

$$a_{ij} = \begin{cases} 1 & \text{if server } j \text{ is concerned by the transaction } i. \\ 0 & \text{else} \end{cases}$$

Each action inherits from the deadline of its transaction.

The contents of execution queues are the result of the algorithm NP-EDF, which has been realised after having obtained a global view of the system, thanks to one of the algorithms mentioned above.

**Assumption 1:** We suppose that there exists no possibility of errors at the level of the system. The customers/servers and servers/servers networks are reliable. The messages are then always transmitted successfully. Similarly, the processors cannot be on any account, failing.

**Assumption 2:** We suppose that the networks used are diffusion networks with bounded delay and that the propagation time is uniformly distributed within the intervals [delaymin, delaymax] for the customers/servers network and [Jmin, Jmax] for the servers network.

A mathematical model has been developed to evaluate the average stay time of a transaction within a system using the circulating multisequencer algorithm (Azouz and Kamoun 1999a, Azouz and Kamoun 1999b, Azouz and Kamoun 2001; Azouz 2002). In this paper, we present the outlines of

this model and we focus on its validation by comparing its results to simulation results developed in (Mettali 1998).

## 2. PRINCIPLE OF THE CIRCULATING MULTISEQUENCER ALGORITHM

This algorithm supposes that a virtual ring is formed between the servers. On this ring circulates a token consisting of a field called announcement table, and in which the servers empty their customer queues at each passage of the token. This token also contains a multisequencer (one sequencer per server).

Two types of servers are defined:

- Non-diffuser servers which, at each passing of the token empty their customer queues in the announcement table and make the token pass to the next server.
- Diffuser servers, which, in addition to the previous operations, are responsible for the scheduling and the labelling of all actions, which exist in the announcement table, according to NP-EDF. They diffuse, afterwards, the result of this scheduling, to the different servers before passing the token to the next server. The servers receive then the lists of the actions to be executed (in their execution queues). The order is reflected by the labelling of the actions, carried out with the help of the multisequencer, which circulates in the token. We find a diffuser server at all  $d$  servers.  $d$  is called diffusion step.

## 3. MATHEMATICAL MODEL

Each time the token goes through a diffuser server, we have the arrival of a scheduled bulk of actions to the different execution queues of different servers. Let  $T$  the random variable representing the time, which separates the passing of the token through two consecutive diffusers. This duration  $T$  is the sum of  $d$  random variables.

$$T = \underbrace{X_1 + X_1 + \dots + X_1}_{(d-1)\text{times}} + X_d$$

where  $(d-1) X_1$  corresponds to the processing time of the token at the non-diffuser servers and  $X_d$  corresponds to the time of its processing at a diffuser server.

The average arrival rate of scheduled actions groups to different servers corresponds then to:  $\lambda_{dif} = \frac{1}{T}$

The arrival at the execution queues is a bulk arrival of actions scheduled according to EDF. The groups of actions are served according to FIFO discipline.

**Assumption 3:** The random variable  $T$  is exponentially distributed with the parameter  $\lambda_{dif}$ .

A server  $j$  is then modelled by a bulk arrival M/G/1 station with  $K$  classes of customers. Bulks are considered according to FIFO discipline. Customers (actions) are scheduled according to EDF within a group. The token is processed, according to the head of line (HOL) discipline without pre-emption of the server at the arrival of the token.

This station can be used to model each execution queue.

We will denote by  $\overline{g_{ij}^k}$  the  $k^{\text{th}}$  moment of the random variable  $r_{ij}$ , which corresponds to the number of actions of type  $i$ , present within a group arriving at sever  $j$ .

## 4. AVERAGE RESPONSE TIME OF THE SYSTEM

The life cycle of a transaction comprises four stages, which will be noted down by Stage $k$  ( $k=1\dots4$ ).

Stage 1 corresponds to the propagation delay of the transaction actions between the customer and the different servers. According to assumption 2,

$$\overline{Stage1} = \frac{\text{delaymin} + \text{delay max}}{2} \quad (1)$$

Stage 2 consists in the wait for the token acquisition by any concerned server with the generated transaction. On the basis of the matrix describing the transaction traffic

$$\overline{Stage2} = \frac{1}{2} \sum_{i=1}^K \frac{TRJ}{L} \frac{ni}{\sum_{i=1}^K a_{ij} \bar{X}} T_{transition} \quad (2)$$

Where  $TRJ$  is the total rotation time of the token at the level of the virtual ring,  $X$  is the random variable corresponding to the token service time at a server,  $L$  is the number of servers and  $T_{transition}$  is the average duration of the token transition between two servers.  $N$ ,  $n_i$  and  $a_{ij}$  have been defined while describing the traffic.

Stage 3 corresponds to the time needed to reach a diffuser server and to diffuse the scheduling result once the transaction has been announced. By analysing the traffic matrix

$$\overline{Stage3} = \sum_{i=1}^d \left[ (d-i) \bar{X}_1 \sum_{j=i \bmod(d)}^{j!} \frac{\sum_{i=1}^N a_{ij}}{N L} \right] + \bar{X}_{diffusion} \quad (3)$$

Where  $d$  is the diffusion step (distance between two diffusers) and  $X_{diffusion}$  is the random variable corresponding to the duration, which separates the arrival time of the token at a diffuser server, from the arrival time of the scheduled actions groups to the servers.

Stage 4 corresponds to the transaction execution time, which corresponds to the time needed to execute all its actions. Execution queues are equivalent to  $L$  queuing systems working in parallel.

$$\overline{Stage4} = \sum_{i=1}^K \frac{ni}{N} \max_j a_{ij} T_{sij} \quad (4)$$

$T_{sij}$  being the average response time of the server  $j$  execution queue for the action  $i$ , given by

$$T_{sij} = \frac{\frac{\lambda d i f \overline{Xg_j^2}}{2} + \rho_{jtoken} \frac{\overline{X^2}}{2X}}{1 - \sum_{i=1}^K \rho_{ij} - \rho_{jtoken}} + \frac{\overline{W_{sij}} - \overline{X_{ij}}}{1 - \rho_{jtoken}} + \overline{X_{ij}} \quad (5)$$

Where  $X_{ij}$  is the execution time of the transaction action  $i$  at the server  $j$ ,  $Xg_j$  is the group service time,  $X_{g_j}^n$  is its  $n^{\text{th}}$  moment,  $\rho_{jtoken}$  is the utilisation factor of the server  $j$  by the token processing,  $\rho_{ij}$  is the utilisation factor of the server  $j$  by the execution of actions of type  $i$  and  $X^n$  is the  $n^{\text{th}}$  moment of the random variable  $X$ .  $W_{sij}$  corresponds to the necessary time for an action  $A_{ij}$  of type  $i$  to be served, once its group has been considered by the server  $j$  and is given by

$$W_{sij} = \overline{X_{ij}} + \sum_{\substack{k=1 \\ k \neq i}}^K \left( \frac{\overline{X_{kj}} P[\text{preced } A_{ij}(\text{type } k)]}{\overline{g_j} \mid_{r_i > 0}} \right) + \overline{X_{ij}} \left( \frac{P[\text{preced } A_{ij}(\text{type } i)] \left[ \overline{g_{ij}^2} - \overline{g_{ij}}^2 + \overline{g_{ij}} \overline{g_j} \right]}{(1 - P[r_i = 0]) \overline{g_j} \mid_{r_i > 0}} - P[\text{preced } A_{ij}(\text{type } i)] \right) \quad (6)$$

Where  $P[\text{preced } A_{ij}(\text{type } k)]$  is the probability that a transaction of type  $k$  precedes the action  $A_{ij}$  of type  $i$ , within the same group arriving to server  $j$ . The average stay time of a transaction within the system is therefore:

$$T_{stay} = \overline{\text{Stage 1}} + \overline{\text{Stage 2}} + \overline{\text{Stage 3}} + \overline{\text{Stage 4}} \quad (7)$$

## 5. Results Presentation

A simulation has been realized using SAMSON (Toutain 1991) and its results are presented in (Mettali 1998). For the validation of the mathematical models, we present a confrontation of mathematical and simulation results (Azouz 2002). Mathematical and simulation models are based on common assumptions except assumption 3, which supposes that the interarrival of groups to execution files is exponentially distributed. Two types of server's networks

are considered:  $R1$  and  $R2$ .  $R1$  is characterized by  $[J_{min}=0.2ms, J_{max}=0.5ms]$  and  $R2$  is characterized by  $[J_{min}=8ms, J_{max}=10ms]$ .

Figures 1 and 2 present a confrontation of mathematical and simulation results concerning the average stay time of a transaction within the system versus the utilization factor of the busiest server ( $\rho_{jmax}$ ). These figures correspond respectively to networks  $R1$  and  $R2$  and consider a diffusion step  $d$  equal to 1, 2, 4 and 8 and a poisson generation process for transactions.

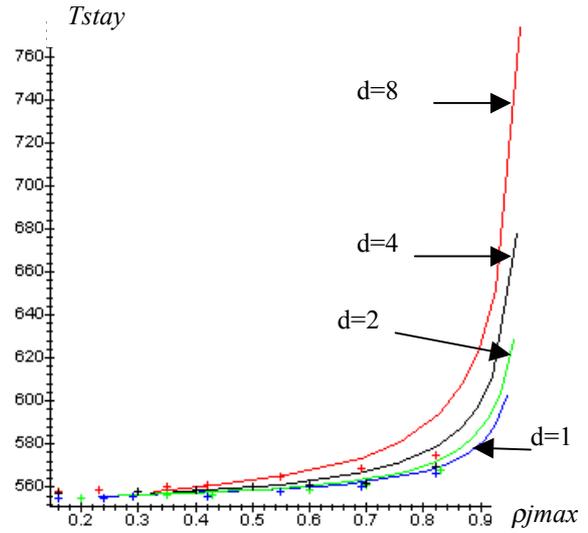


Figure 1: Average stay time versus  $\rho_{jmax}$   
 $K=4 - R1 : [0.2ms, 0.5ms]$   
 Mathematical model (continued lines) – Simulation (+).

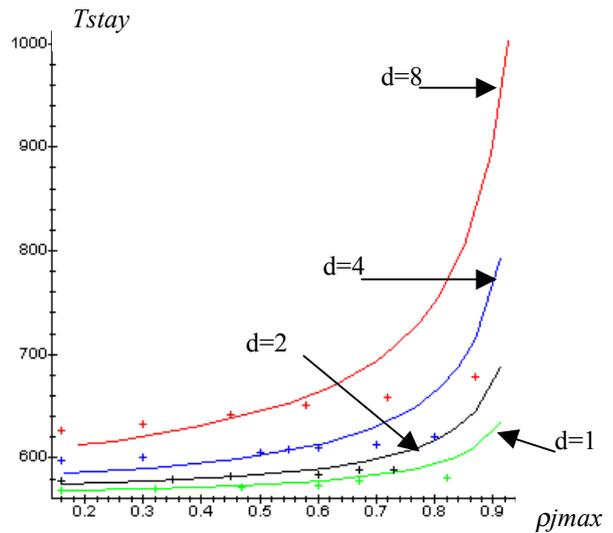


Figure 2 : Average stay time versus  $\rho_{jmax}$   
 $K=4 - R2 : [8ms, 10ms]$   
 Mathematical model (continued lines) – Simulation (+).

This comparison shows that for network of type R1, mathematical and simulation results are practically astounded, except for  $d=8$ , for which we begin to observe a difference for a load greater than 80%. However, for network of type R2, theoretical and simulation results converge until a utilization factor of 70%, after which we begin to observe a divergence, for  $d=4$  and 8. This divergence increases, when the charge or the diffusion step and consequently, the bulk size increases. Mathematical results are in this case greater than simulation ones. We note that simulation results are stable and that no saturation is observed! In fact we note that we begin to observe a difference between the two models, each time we begin to observe the saturation with the mathematical model. To try to explain this difference, we present a mathematical result analysis in tables 1, 2, 3 and 4, which respectively correspond to the situations (R1 and  $d=2$ ), (R1 and  $d=8$ ), (R2 and  $d=2$ ) and (R2 and  $d=8$ ).

These tables present the variations of  $\rho_{jmax}$  (the utilization factor of the busiest server),  $W0_j$  (the average residual service time of a bulk),  $W1_j$  (the average waiting time of a bulk),  $\overline{Stage4}$  (the average execution queues response time) and  $Tstay$  (the average stay time of a transaction within the system), versus the number of generated transactions.

Table 1 : Mathematical results analysis - R1 –  $d=2$

$N$ (trans /sec)	$\rho_{jmax}$	$W0_j$ (ms)	$W1_j$ (ms)	$\overline{Stage4}$ (ms)	$Tstay$ (ms)
100	0.39	0.12	0.26	3.05	556.28
200	0.67	0.63	1.99	4.72	558.54
300	0.96	1.78	39.12	42.62	597.34

Table 2 : Mathematical results analysis - R1 –  $d=8$

$N$ (trans /sec)	$\rho_{jmax}$	$W0_j$ (ms)	$W1_j$ (ms)	$\overline{Stage4}$ (ms)	$Tstay$ (ms)
100	0.43	0.33	0.67	3.64	558.76
200	0.69	1.98	6.39	11.62	568.98
300	0.97	5.92	165.56	174.18	733.78

Table 3 : Mathematical results analysis - R2 –  $d=2$

$N$ (trans /sec)	$\rho_{jmax}$	$W0_j$ (ms)	$W1_j$ (ms)	$\overline{stage4}$ (ms)	$Tstay$ (ms)
100	0.31	0.89	1.31	5.81	575.98
200	0.61	3.74	9.66	17.25	588.52
300	0.91	8.78	99.25	110.61	682.57

Table 4 : Mathematical results analysis - R2 –  $d=8$

$N$ (trans /sec)	$\rho_{jmax}$	$W0_j$ (ms)	$W1_j$ (ms)	$\overline{Stage4}$ (ms)	$Tstay$ (ms)
100	0.31	3.44	5.03	18.15	616.47
200	0.61	14.43	37.25	63.09	663.67
300	0.91	33.94	384.86	424.5	1027.32

This analysis shows that the difference between the two models is essentially due to  $W1_j$  (the average waiting time of a bulk), which depends essentially of  $W0_j$  (the average residual service time of a bulk). However  $W0_j$  and  $W1_j$  are the direct results of M/G/1 station study and particularly of assumption 3 which supposes that bulks interarrival is exponentially distributed. This implies that the difference observed between mathematical and simulation results, when bulk size increases and mathematical model begins to saturate, is due to this simplifying assumption.

We can then conclude that the mathematical model developed here is very near from the reality, while the bulk size isn't sufficiently important to begin to observe saturation with the mathematical model. This corresponds to these situations:

- Servers networks with low transmission delays,
- Servers networks with important transmission delays,
  - with a small value for diffusion step and a load not exceeding 80%,
  - with a high value for diffusion step and a load not exceeding 70%.

Mathematical model is then valid for all these situations and the assumption 3 which supposes that bulks interarrival is exponentially distributed becomes penalizing only when the mathematical model begins to display a saturation. Mathematical model can be judged satisfying, since

1. Actual technologies make local area networks offering low transmission delays.
2. Mathematical model and simulation have shown that the use of low diffusion step gives better and convergent results.
3. Out of these situations leading to valid model, mathematical model gives greater results than those obtained by simulation. In fact, this isn't very constringing, since this model will be used to obtain the maximum stay time of a transaction within the system, to determine the feasibility conditions (Azouz and Kamoun 1999b, Azouz and Kamoun 2002).

## 6. Conclusion

In this paper, We have presented a mathematical model developed for a transactional distributed system using the circulating multisequencer algorithm. This model gives the average stay time of a transaction within the system. A bulk arrival M/G/1 station with K classes of customers where bulks are considered according to FIFO discipline, actions are scheduled according to EDF within a group and the processing of the token is executed according to the head of

line (HOL) discipline, has been introduced to model servers execution queues.

To validate the mathematical model, mathematical and simulation results are confronted. This study shows that the mathematical model developed here is satisfying in most situations.

## REFERENCES:

- Lelann, G. 1994. « Systèmes transactionnels répartis temps réel ». Journées Internationales des Sciences Informatiques - Tunis -Tunisia
- Anceaume, E. Georges, L. Lelann, G. 1995. « Timeliness proofs for real time distributed systems in the presence of high loads » research report INRIA.
- Lelann, G. 1990. « Critical issues for the development of Distributed real time computing systems », research report INRIA n° 1274.
- Georges, L; Muhlethaler, P; and Rivière, N. 1995. « Optimality and non-preemptive real-time scheduling revisited ». Research report INRIA no 2516.
- Azouz Saidane, L and Kamoun, F. 1999. « Modelling and Performance Evaluation of the Circulating Multisequencer Scheduling Algorithm in a Real-time Distributed Transactional System », CNDS'99, Computer Networks and Distributed Systems modeling and analysis, San Francisco, USA.
- Azouz Saidane, L and Kamoun, F. 1999. « Modelling and Performance Evaluation of the Multi-tokens and the Circulating Multisequencer Scheduling Algorithms in a Real-time Distributed Transactional System », Performance evaluation International Journal. (n° 36-37(1999) 429-456), NH Elsevier.
- Azouz Saidane, L and Kamoun, F. 2001. « Performance Evaluation of the Circulating Multisequencer, the Multi-tokens and the Consensus Algorithms in a Real-time Distributed Transactional System », The International Journal of Foundations of Computer Science (IJFCS), Vol.12 No. 6 (2001) 719-749, World Scientific Publishing Company.
- Mettali Gammar, S. 1998. "Evaluation des performances d'algorithmes d'ordonnancement pour des systèmes transactionnels répartis temps réel", Thèse de doctorat, Université de Tunis II.
- Azouz Saidane, L. 2002. « Modélisation et Etude Analytique des Performances d'Algorithmes d'Ordonnancement dans un Système Transactionnel Réparti Temps Réel », Thèse de doctorat d'état, Université d'El Manar.
- Toutain, L. 1991. "SAMSON Simulation Analysis and Monitoring of Systems and Open Networks", Thèse de doctorat, Université du Havre.

## AUTHOR BIOGRAPHY

**LEILA AZOUZ SAIDANE** was born in Carthage (Tunis), Tunisia and went to the university of Tunis where she obtained the computer engineering degree in 1982, the doctoral degree (doctorat 3<sup>ième</sup> cycle) in 1985 and the second doctoral degree ( doctorat d'état) in 2002. She was assistant professor at the faculty of science of Tunis (1982-1985) and is teaching as associate professor at ENSI which is a computer engineering school at the university of Mannouba, since 1985. Her research domain is performance evaluation and she is teaching probability, queuing theory, performance evaluation and networking and data communication.

# A TECHNIQUE FOR SIMULATION OPTIMIZATION

Mahmoud H. Alrefaei

Department of Mathematics and Statistics  
Jordan University of Science & Technology  
P.O. Box 3030, Irbid 22110, JORDAN  
E-mail: [alrefaei@just.edu.jo](mailto:alrefaei@just.edu.jo)  
URL: <http://www.just.edu.jo/~alrefaei>

## KEYWORDS

Global, Optimization, Queuing, Simulation, Stochastic.

## ABSTRACT

We investigate the simulation optimization problem in which the objective function values are not known exactly, but have to be estimated using simulation. The objective function is also assumed to be a multimodal. In this paper, we present an approach that can be used for solving this problem. This approach uses the modified simulated annealing algorithm together with the standard clock simulation technique. The hill climbing feature of the simulated annealing is used to escape the local minimality trap and the standard clock simulation technique is used for improving the computational efficiency.

## 1 INTRODUCTION

In this paper, we are interested in solving discrete optimization problems in both transient and steady-state simulation in which the objective function is the expected performance measure of some complex stochastic system. In the transient case, we are interested in solving the following optimization problem:

$$\min_{x \in \mathcal{S}} f(x) = E[h(x, Y_0(x), \dots, Y_{\tau_x}(x))], \quad (1)$$

where  $\mathcal{S}$  is discrete set of feasible solutions,  $h$  is a deterministic function, for all  $x \in \mathcal{S}$ ,  $\{Y_i(x)\}$  is a stochastic process whose distribution depends on the parameter  $x$ , and  $\tau_x$  is a stopping time with respect to  $\{Y_i(x)\}$ . We assume that it is possible to draw independent samples  $Y_0^j(x), \dots, Y_{\tau_x}^j(x)$ ,  $j = 1, 2, \dots$ , from  $Y_0(x), \dots, Y_{\tau_x}(x)$  for all  $x \in \mathcal{S}$ .

In the steady-state simulation case, we are interested in solving the following optimization problem:

$$\min_{x \in \mathcal{S}} f(x) = E[h(x, Y(x))], \quad (2)$$

where  $Y(x)$  is the steady-state limit of the stochastic process  $\{Y_i(x)\}$  whose distribution depends on the parameter  $x$  for all  $x \in \mathcal{S}$  (i.e.,  $Y_i(x) \Rightarrow Y(x)$  as  $i \rightarrow \infty$  for all  $x \in \mathcal{S}$ , where the symbol  $\Rightarrow$  denotes weak convergence (convergence in distribution)).

Recently, there are several approaches that have been proposed for solving this problem. These approaches include (Norkin et al. 1994) who propose using a stochastic version of the Branch and Bound methods. Some authors consider versions of simulated annealing algorithm such as (Gelfand and Mitter 1989), (Gutjahr and Pflug 1996), and (Fox and Heine 1995). All these approaches assume a control sequence (the annealing sequence) that converge to zero with some rate in order to prove that their approaches converge to an optimal solution in probability under some assumption. (Alrefaei and Andradóttir 1999) propose using the simulated annealing algorithm with constant temperature instead of using a decreasing sequence that converges almost surely to an optimal solution under mild conditions. Other random search methods for solving discrete simulation optimization problems can be found in (Alrefaei and Andradóttir 2000), and (Andradóttir 1995, 1996), and (Ólafsson and Shi 1999).

In this paper, we present an approach that uses the standard clock (SC) technique for estimating the objective function values at multiple points, simultaneously, using only one sample path (see Alrefaei (2002)). This method starts by an initial state. At any time  $k$ , we use the SC technique to get estimates of the objective function values for all states in the set of the neighborhood of the current state using only one simulation clock (the standard clock). One of these states is selected as the

next current state with a probability selection that depends on the estimate of the objective function values. States that seems to be good states (have small objective function values in the case of minimization problems) are selected with high probability. However, other states are not neglected (because they may hide good states behind them) they are also considered for selection but with less probabilities.

This paper is organized as follows. In Section 2, we briefly review the standard clock technique of (Vakili 1991). In Section 3, we describe our approach. In Section 4, we provide a numerical example, and in Section 5, we give some concluding remarks.

## 2 THE STANDARD CLOCK SIMULATION TECHNIQUE

The standard clock simulation technique is proposed by (Vakili 1991) to improve the computational efficiency, when a multiple sample paths are needed, one clock mechanism (standard clock) that determine the event type and the timing for the events is shared by all sample paths. Suppose there are  $n$  event types ( $e_1, e_2, \dots, e_n$ ) of exponential rates  $\mu_1, \mu_2, \dots, \mu_n$  respectively, are involved in simulating  $K$  different designs. Let  $\mu = \sum_{i=1}^n \mu_i$  denotes by the *maximal rate*. Then the timing for the next event is generated by generating an exponential random variable  $\tau$ . This represents the epoch for the next event for all the  $K$  designs. To determine the next event type, we generate a uniform random variable  $u \in [0, 1]$ . If  $0 \leq u < \mu_1/\mu$ , then the next event is  $e_1$ . If  $1/\mu \sum_{i=1}^{j-1} \mu_i \leq u < 1/\mu \sum_{i=1}^j \mu_i$  then the next event is  $e_j$ ,  $j = 2, \dots, n$ . After determining the next event and its timing, all sample paths are updated. Note that, it is possible that an event is infeasible for some designs. This could happen when the event is not on the event list (i.e., the event is not scheduled to occur.) e.g., a service completion for a server who is not providing a service. In this case, the event time of this fictitious event is used to update the system time for these designs as if this event were occur. However, no changes to the state will take place for these designs.

## 3 SIMULATION OPTIMIZATION ALGORITHM

We assume that the feasible set of solutions  $\mathcal{S}$  in problems (1) and (2) is well structured in the sense that for each state  $x \in \mathcal{S}$ , there exists a neighborhood subset

$N(x)$  of  $\mathcal{S} \setminus \{x\}$ .  $N(x)$  is called the set of neighbors of  $x$  that satisfies the following technical assumption. For any  $x, x' \in \mathcal{S}$ ,  $x'$  is reachable from  $x$ ; i.e., there exists a finite sequence  $\{n_i\}_{i=0}^l$  for some  $l$ , such that  $x_{n_0} = x$ ,  $x_{n_l} = x'$ , and  $x_{n_{i+1}} \in N(x_{n_i})$ ,  $i = 0, 1, 2, \dots, l-1$ .

Now we describe our approach:

### Algorithm 1

**Step 0:** Select a starting point  $X_0 \in \mathcal{S}$ . Let  $k = 0$  and  $X_k^* = X_0$ .

**Step 1:** Given  $X_k = x$ , use the standard clock technique to draw independent samples  $Y_0^j(z), \dots, Y_{\tau_z^j}^j(z)$ ,  $j = 1, 2, \dots, L_k$  from  $Y_0(z), \dots, Y_{\tau_z}(z)$  for all  $z \in N(x) \cup \{x\}$  simultaneously. where  $\tau_z^j$  is the stopping time for the transient case, and it is large enough for the steady state case. Compute  $\hat{f}_k(z)$  as follows:

$$\hat{f}_k(z) = \frac{1}{L_k} \sum_{j=1}^{L_k} h(z, Y_0^j(z), \dots, Y_{\tau_z^j}^j(z)) \quad (3)$$

**Step 2:** Given  $X_k = x$ , let  $\hat{R}(x)$  be defined as follows:

$$\hat{R}(x) = \sum_{z \in N(x) \cup \{x\}} \exp \left[ -C \left[ \hat{f}_k(z) - \hat{f}_k(x) \right]^+ \right].$$

Set  $X_{k+1} = z$  with probability

$$p'_{x,z}(k) = E \left[ \frac{\exp \left[ -C \left[ \hat{f}_k(z) - \hat{f}_k(x) \right]^+ \right]}{\hat{R}(x)} \right], \quad (4)$$

where  $C$  is a constant used as a scalar.

**Step 3:** Let  $k = k + 1$ , Let  $\tilde{f}_k(X_k)$  be the average function value at  $X_k$  during the previous  $k$  iterations. If  $\tilde{f}_k(X_k) < \tilde{f}_k(X_{k-1}^*)$  then let  $X_k^* = X_k$ ; otherwise let  $X_k^* = X_{k-1}^*$ . Go to Step 1.

Note that the sequence  $\{X_k\}$  is a time-inhomogeneous Markove chain whose transition probabilities  $P'_k(x, x') = P\{X_{k+1} = x' | X_k = x\}$  are given by

$$P'_k(x, x') = \begin{cases} p'_{x,x'}(k) & x' \in N(x) \cup \{x\}, \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

for all  $k \in \mathbb{N}$ , where  $p'_{x,x'}(k)$  is given by equation (4).

Let  $P$  be the transition probability matrix defined by

$$P(x, x') = \begin{cases} \frac{1}{R(x)} \exp[-C[f(x') - f(x)]^+] & \text{if } x' \in N(x), \\ 1 - \sum_{y \in N(x)} P(x, y) & \text{if } x' = x, \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

where

$$R(x) = \sum_{z \in N(x) \cup \{x\}} \exp[-C[f(z) - f(x)]^+].$$

(Alrefaei 2002) shows that the transition probability matrix  $P$  is irreducible and aperiodic and has stationary distribution  $\pi$ , where  $\pi$  is a vector whose entries  $\pi_x$  are given by:

$$\pi_x = \frac{R(x) \exp[-Cf(x)]}{\sum_{x' \in \mathcal{S}} [R(x') \exp[-Cf(x')]]} \quad (7)$$

for all  $x \in \mathcal{S}$ . (Alrefaei 2002) also proves that  $P_k'(x, x') \rightarrow P(x, x')$  as  $k \rightarrow \infty$  for all  $x, x' \in \mathcal{S}$ , where the Markov matrices  $P_k$  and  $P$  are given in equations (5) and (6), respectively.

The proof of the following theorem mimics the proof of Theorem 4 of (Alrefaei 2002).

**Theorem 1** *The sequence  $\{X_k^*\}$  generated by our procedure converges almost surely to a global optimal solution.*

### Stopping Rule

For the purpose of stopping the algorithm, we repeat the procedure several times until we feel that the gap between the estimated optimal value and the actual optimal value is sufficiently small. The following algorithm explains that.

### Algorithm 2

**Step 0:** Choose the maximum number of replications  $M$ .

**Step 1:** For  $m = 1, \dots, M$  do

1. Perform Algorithm 1 and get estimates of the solution  $\hat{x}^m$  and the optimal value  $\hat{v}^m = \tilde{f}(\hat{x}^m)$ .

2. Estimate the optimality gap between the function value of the estimated optimal solution and the actual optimal value and its variance. The estimated optimal solution at replication  $m$  is given by  $\bar{x}^m = \operatorname{argmin}_{i=1, \dots, m} \hat{v}^m$  and the estimated optimal value is given by  $\bar{v} = \frac{1}{m} \sum_{i=1}^m \hat{v}^i$ . If the optimality gap and its variance are sufficiently small, go to Step 3.

**Step 3:** Use the screening and selecting method of (Nelson et. al. 2001) to select an estimate of the optimal solution that is indifferent with the actual solution by a predetermined indifference  $\delta$  with confidence level of  $(1 - \alpha)100\%$ .

## 4 NUMERICAL APPLICATION

In this section, we implement our algorithm for solving an optimization problem involving  $M/M/1$  queuing system. The optimization problem is given by

$$\min_{\mu \in (3,6)} (\mu - 4)^2 + w(\mu) \quad (8)$$

where  $w(\mu)$  is the mean average system time per customer given the service rate is  $\mu$  customers per hour and the inter-arrival rate is  $\lambda$  customers per hour, where  $\mu \in (3, 6)$  and  $\lambda = 3.0$ . In practice, this problem can be found in a queuing system whose stream of arrivals is fixed at  $\lambda = 3.0$  customers per unit time, and the current service rate is  $\mu = 4$  customers per unit time. The decision maker wants to decide the speed of service time given that the cost function is given by  $g(\mu) = (\mu - 4)^2 + w(\mu)$ . Of course, this is a continuous stochastic optimization problem, we make it discrete by considering the feasible set of solutions as  $\mathcal{S} = \{3.0 + .01i, i = 1, \dots, 300\}$ . If one needs a more precise estimated optimal solution, say within 0.001 of the actual solution, one should increase the size of the problem to 3000 alternatives. Note also that the optimization problem (8) has analytical optimal solution as the simulation reaches the steady-state. The analytical optimal solution is  $\mu = 4.3$  with an optimal value of 0.859.

Figure 1 depicts the actual objective function. It is clear that the function is quite flat around the optimal solution which makes it hard to locate the optimal solution.

We perform Algorithm 2 by replicating Algorithm 1 100 replications each with 200 iterations that are increased by 10 iterations for each 10 replications. The

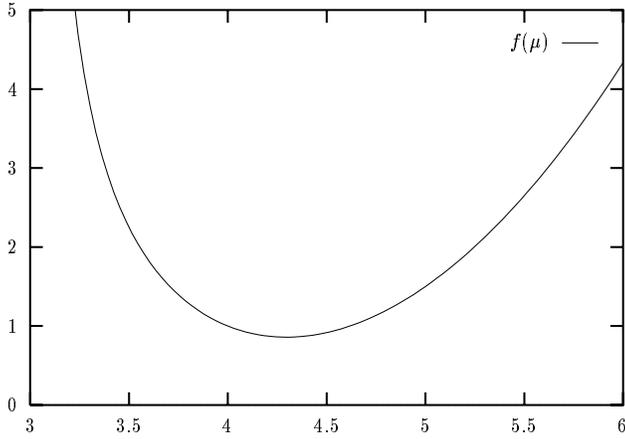


Figure 1: The Objective Function  $f(\mu) = (\mu-4)^2 + w(\mu)$

length of each simulation run varies using the formula  $m = 100,000 + 250k$  service completion, where  $k$  is the iteratin number. We use the indifference zone  $\delta = 0.01$  and a confidence level of 90%. The neighborhood for each feasible solution  $x$  is used as

$$N(x) = \{x | x - 0.1 \leq x \leq x + 0.1\} \cap \mathcal{S}$$

The code is written using Fortran 90 and executed on a Pentium III, 800 MHz. Table 1 shows the estimated error gap between the estimated optimal value and the actual optimal value. Column 3 of Table 1 gives the the variance of the error gap over the 100 replications. It is clear that the error gap variance is decreased as the number of replications is increased. Table 2 shows the estimated optimal solution and its estimated function value. Column 4 shows the estimated optimal value.

There are only 13 feasible solutions survived during the 100 replications; these are 4.23, 4.24, 4.25, ..., 4.36. Table 3 includes the estimated function values for the survived solutions and their variance, the last column indicate how many samples, each survived solution has been sampled during the 100 replications. These are the only states that are carried out for the second stage of the Rinott's procedure (see Nelson 2001). The process of (Nelson 2001) is used to select the state that has the minimum function value. The feasible solution  $\hat{x} = 4.30$  has been selected using this procedure as the final estimate of the optimal solution which is the same as the actual optimal solution and the estimated optimal value is 0.859 with a small variance  $1.0 \times 10^{-5}$ .

Table 1: The Error Gap and Its Variance

Rep. #	Error Gap	The Gap Variance
10	0.00646356	0.00252585
20	0.00552847	0.00157250
30	0.00554021	0.00124703
40	0.00560909	0.00171447
50	0.00550803	0.00154301
60	0.00531977	0.00143791
70	0.00504474	0.00135975
80	0.00509018	0.00128277
90	0.00498923	0.00122340
100	0.00497884	0.00118326

Table 2: The Estimated Optimal Solution and Optimal Value

Rep. #	Estimated Sol.		Estimated Opt. Value
	Sol.	Value	
10	4.33	0.86107848	0.85461493
20	4.33	0.86052668	0.85499821
30	4.33	0.86076393	0.85522372
40	4.27	0.86069904	0.85508995
50	4.27	0.86058867	0.85508064
60	4.27	0.86060326	0.85528348
70	4.27	0.86056373	0.85551900
80	4.27	0.86063837	0.85554819
90	4.27	0.86063608	0.85564685
100	4.27	0.86063970	0.85566086

## 5 CONCLUSION

We have presented an algorithm that locates the optimal solution for the simulation optimization problem when the objective function values can be estimated using transient or steady state simulation. This approach involves using the standard clock technique for estimating the function values at multiple points simultaneously. This algorithm is guaranteed to converge almost surely to a global optimal solution under mild conditions. We repeat the procedure several times until we get a confidence that the optimal solution is reached with high probability. In the future, we are interested in applying this algorithm to more realistic simulation optimization problems.

Table 3: The Survival Solutions and their Estimated Function Value and Variance

Sol	Est. Value	Est. Var.	Samples
4.23	0.86646559	0.00002264	99
4.24	0.86535850	0.00002451	99
4.25	0.86283854	0.00001711	100
4.26	0.86122829	0.00001247	100
4.27	0.86063970	0.00001843	100
4.28	0.85976984	0.00001002	100
4.29	0.85932553	0.00001035	100
4.30	0.85908465	0.00001004	100
4.31	0.85969331	0.00001238	100
4.32	0.86066734	0.00000888	100
4.33	0.86054825	0.00000805	100
4.34	0.86181632	0.00001235	100
4.35	0.86328631	0.00000968	100

## REFERENCES

- Alrefaei, M. H. 2002. "Stochastic Optimization Using the Standard Clock Simulation." *International Journal of Applied Mathematics*, 8: 317–333.
- Alrefaei, M. H., and S. Andradóttir. 1999. "A Simulated Annealing Algorithm with Constant Temperature for Discrete Stochastic Optimization." *Management Sci.*, 45: 748–764.
- Alrefaei, M. H., and S. Andradóttir. 2001. "A Modification of the Stochastic Ruler Method for Discrete Stochastic Optimization." *The European Journal of Operational Research*, 133: 160–182.
- Andradóttir, S. 1995. "A Method for Discrete Stochastic Optimization." *Management Sci.*, 41:1946–1961.
- Andradóttir, S. 1996. "A Global Search Method for Discrete Stochastic Optimization." *SIAM J. Optimization*, 6:513–530.
- Fox, B. L., and G. W. Heine. 1995. "Probabilistic Search with Overrides." *Ann. Appl. Prob.*, 5:1087–1094.
- Gelfand, S. B., and S. K. Mitter. 1989. "Simulated Annealing with Noisy or Imprecise Energy Measurements." *J. Optimization Theory and Applications*, 62:49–62.
- Gutjahr, W. J., and G. Ch. Pflug. 1996. "Simulated Annealing for Noisy Cost Functions." *J. Global Optimization*, 8:1–13.
- Nelson, B. L., J. Swann, D. M. Goldsman, and W. Song. 2001. "Simple Procedures for Selecting the Best Simulated System when the Number of Alternatives is Large." *Operations Research*, 49, 950–963.
- Norkin, V., Y. Ermoliev, and A. Ruszczyński. 1994. "On Optimal Allocation of Individuals under Uncertainty." Appeared in a series *International Institute for Applied System Analysis, A-2361 Luxembourg, Austria*, 1994.
- Ólafsson, S. and L. Shi. 1999. "Optimization via Adaptive Sampling and Regenerative Simulation." In *Proceedings of the 1999 Winter Simulation Conference*, ed. P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, 666–672. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.
- Vakili, Pirooz. 1991. "Using a Standard Clock Technique for Efficient Simulation." *Operation Research Letters.*, 10:445–452.

## BIOGRAPHY

Mahmoud Alrefaei is currently an Assistant Professor at the Department of Mathematics and Statistics and the Department of Computer Sciences and Information Systems at Jordan University of Science and technology. He receives his B.Sc and M.Sc. from Yarmouk University on 1983, 1985, respectively. He also receives an MA. in Mathematics and a joint Ph.D in Mathematics and Industrial Engineering in the field of Operations Research from the University of Wisconsin-Madison (USA) in 1997. He also had the opportunity to work as a research assistant at the School of Industrial and Systems Engineering at Georgia Institute of Technology, which has been ranked number one 12 times during the last 13 years over the USA. He has developed the Simulated Annealing Algorithm for solving discrete stochastic optimization problems. His research interest includes Simulation Optimization, Simulated Annealing, Simulation and Optimization of Queuing Networks.

# **ENVIRONMENTAL SIMULATION**



# ENVIRONMENTAL THEORY OF PROCESSES

Jiří F. Urbánek and \*Bedřich Rýznar

Brno University of Technology and \*Ground Forces University of the Czech Army in Vyškov  
Technická 2, 616 69 Brno; \*V.Nejdlého, 682 01 Vyškov, Czech Republic,  
[urbane@ust.fme.vutbr.cz](mailto:urbane@ust.fme.vutbr.cz)

## KEY WORDS

Theory of Processes, Environment, Integration, Technology

## ABSTRACT

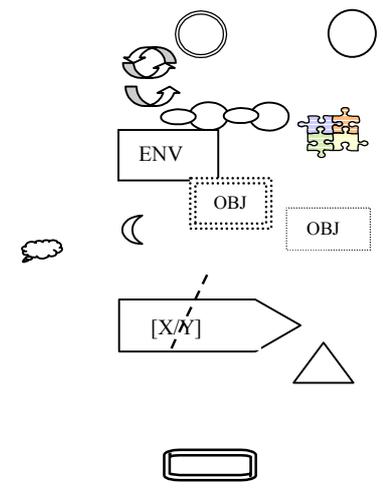
Special process approach is implemented for evaluation, management and change of environmental integrated universal processes. Here mathematical Set Theory is preferred as very friendly and well-arranged graphic apparatus for computerised aid of processes simulation. New Theory of Processes - ToP<sup>®</sup> (Urbánek 2001) is scientific base of complexly first time introduced Czech language book<sup>®</sup>: "Theory of Processes - Management of environments". Here, at interdisciplinary state-of-art, are offered pure original, world-new methods and approaches, which outline environmental/ technology/ production/ anthroposophy problem solutions. Main special entities of the ToP<sup>®</sup> are following keywords (alphabetically): Attribute, Blazon, Cell, Cycle, Environment, Fragment Function, Management, Modality, Operation, Process, Productivity, Relation, System, Technology. Here the process functions create a chain of five modalities – principal, fundamental, subsidiary, possible and supplementary. The cycle is defined reliably, if terminal – environmental connected (i.e. principal & supplementary) functions come through defined environments. The ToP<sup>®</sup> methodology is easy and comprehensive engineering & management tool of very complicated environmental relation. A process is not absolute entity. It is function structure of any activity. Natural and/or anthropogenic process characters are relativity, dynamics and integration to environments. The procedure is given way of transaction activity. The procedure, especially formalized (written) one is the mean for a reaching of process aims only. The technology is process chain model, which expresses the characters, tools, agents and function incidences during process object transformation. At operation discriminative level the technology is integrated to process cell within environmental input/output. The environment isn't only pure surroundings (inversion) of process system, but it is incidental to the process system, which includes also.

## INTRODUCTION TO THE UNIVERSAL THEORY OF PROCESS RELATIVITY

The ISO 8402 defines the process: "The set of mutually interconnected sources and activities, which change an inputs to an outputs." But this definition needs revision towards environmental integration. It is necessary to build integrated, complex and interdisciplinary branch, based on chain functional structure and integrity way of the

processes and their environments. By this paper author's opinion (Urbánek and Skála 2001) is, such above branch, just represented by means a "Theory of Processes" (ToP<sup>®</sup>),

Table 1. The ToP<sup>®</sup> in the synopsis

PROCESS ENTITIES		
struc ture	Named (the examples)	Expressed, symbolised, modelled, specified,
<b>VARIABLES</b>	independent	An Agent An Information A Space A Time Man, machine, tool i- data, symbol, bit s- (three-dimensional) t- (time units)
	dependent	Price Coherence Productivity Value Added Transformation Simultaneity Synergy Influence P (money unit) κ (unit-less) PR (fraction) VA (unit-less) τ (unit-less) S (unit-less) σ (unit-less) ψ (unit-less)
<b>FUNCTIONS</b>	qualitative	<i>to act</i> <i>to add</i> <i>to change</i> <i>to create (design)</i> <i>to decide</i> <i>to implement</i> <i>to operate</i> <i>to optimise</i> <i>to organize</i> <i>to tax</i> <i>to transform</i> by function function function function function function function by price τ
	quantitative	
linkage	modality to ENV chains	Terminal: principal (Pri) fundamental (Fun) subsidiary (Sub) possible (Pos) Terminal: supplementary (Sup)

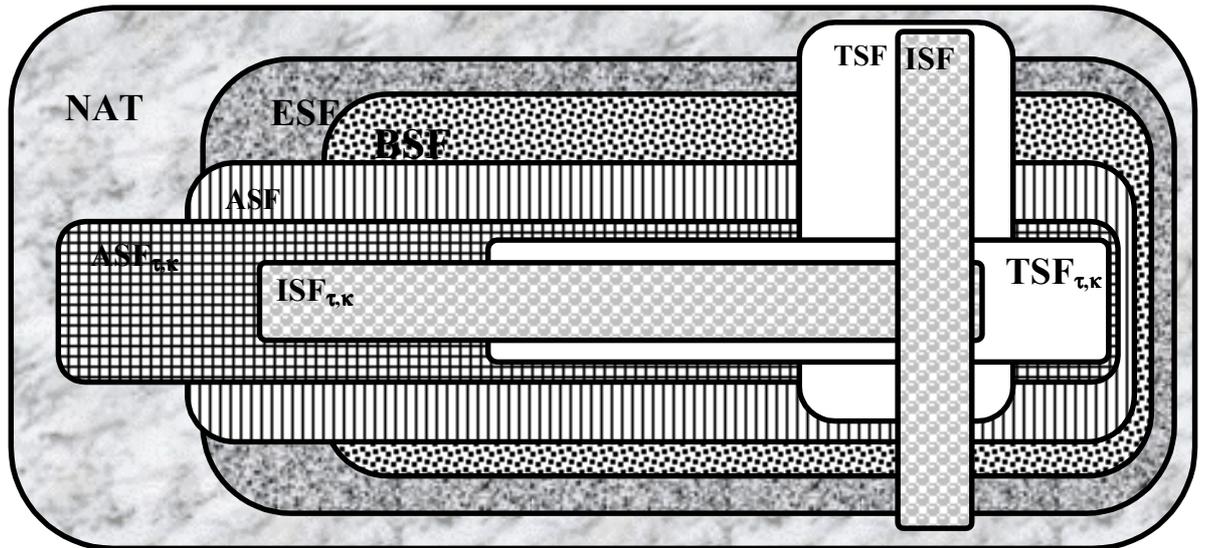
Tab 1. The ToP<sup>®</sup> in the synopsis (continue)

RELATIONS		modal and semantic properties and relation	
		the Accuracy the Adequacy the Automation the Coherence the Connection the Dynamism the Effectiveness the Flexibility the Importance the Liability the Necessity – “Nec” time Orientation the Possibility – „Pos“ the Probability the Recency the Relevance the Reliability the Trustworthiness the Universality the Validity	modality of field semantic field field κ (semantic magnitude) interaction modality of vector field incidence of field modality of field modality of field semantic field semantics vector field semantics semantic field semantic field semantic field semantic field semantic field semantic field
BLAZONS		syntactic	
		the Cycle the Recycling the Fragment the Chain the Environment the Object the Phase the Batch the Interface the Productivity the Subject the Operation Process Quantum The System	waste                  product  waste                  product  X/Y  consumer, market Process cell (Ucell) PrQ
		pragmatic	
		the Automation the Blazon the Clip the Decision the Layout the Outline the Ucell the Plan the Production the Program the Project the Script the Sketch	the Multimedia Means of an Expression
OTHER ENTITIES		process	
		the Attribute the Agents the Rank the Implication the Manager the Management the Operation the Operator the Organization the Mechanism the Method the Prescription the Procedure the Program Technology System the Technology	epithet, symptom factors, tools, apparatus, R, a degree of change effect of influence → decides off-line anthropogenic control a process at defined ENV operates on-line process system principle of incidence model of procedure plan of process model of proceeding virtual prescription model of process system model of process

which is constructed on a basis of up-to-date level process system (production) knowledge. The ToP<sup>®</sup> offers common "language" - apparatus, arguments, criteria, evaluative

tools, technology and approaches for anthroposophy, ecology and economy management of industrial process systems. This "language" was born from technology father's and set mathematics mother and it sight of world's light recently. It is clear, comprehensible, well arranged, narrowly logical, applicable and easy transferable to computer aid. However, it is "freshly birth child" and it certainly will be predisposed to all child's diseases. Now, it is a turn at technologic community, to catch it and to begin its training (exercising, using, applying, elaboration, deepening, correction, reparation and upgrade). The ToP<sup>®</sup> is suitable for any process (Rýznar and Urbánek 2002) - biological, chemical, physical, technological, production, creation, design, evaluation, management, decision making, engineering, etc. The implementation of time-space-information dynamic parameterised entity structure makes possible to apply process approach in different branches at universal axioms and with applying of ToP<sup>®</sup> rules, which are inaugurated in Table 1. Specific process approach creates a quite new ways for the virtual reception of the actions, which proceed in natural and/or anthrop environments. Here process environmental relativity is introduced by means of fundamental mathematical definitions of new entities and functions chains of five modalities. It is susceptible to express very complex, multi-levels relations, influences and integration of arbitrary process chains within their environments. Incoming new era of process approach gives a necessity to express process evaluation, management, design and reengineering (Hammer and Champy 1993) by the help of multimedia possibilities of the computers. It is necessary to say, that the "golden age" of a philosophy about state (about system) was finished in nineteenth century. The Physics and the Technology (leading to the epoch of nuclear bomb, laser and computer) are about the processes rather than about a state, about the metamorphoses rather than about an existence. It seems, that three things will characterise a science of twentieth century: "relativity, quantum mechanics and theory of a chaos" (Mandelbrot 1982). How they are different. But all three are about the processes (Hammer and Champy 1993). Similarly is possible go to another philosophies, where revolutionary changes are happened in past century, for example to the biology, sociology, economics, etc. The shift is happened from the system to the process in their investigation. Cognition sources of anthropogenic processes and their environments are – the technology (TSF), anthroposophy (ASF), industrial production (ISF), which can be monitored by a logistics (flows of material substances); and an informatics (transmutation & flows of the data to the anthrop-pragmatic information in the frame ASF). These expressive sources are derived at computer created topology model at the Figure 1. Here the necessity of anthrop sustainable development regulation (see τ, κ indexed entities) is expressed in relation to a Nature (NAT) and its biosphere (BSF) and ecologic sphere (ESF). This model is expression of "relative field (blazon)", which is vector composed from two-dimensional directions (Production and/or Cognition). For future anthrop sustainable development is desirable "horizontal extension" (on behalf of the cognition) together with "vertical compression" of regulated (τ, κ) spheres (Urbánek 1999).

Extension  
a  
x  
i  
s  
of  
P  
R  
O  
D  
U  
C  
T  
I  
O  
N  
field



Extension axis of COGNITION field

Figure 1: A model of environmental spheres in a Nature.

A waste becomes from all products always and/or the product could arise from all waste, if we try one's utmost. The waste cannot be taken separately, but it is a native product (P&W) of processes technologies (Urbánek 1999). Without-waste technology doesn't exist; only bad discriminative level of process evaluation can characterize the technology like without waste. Discriminative level relates with process environs definition. If anthroposophy approach is applied to process system and its environs (include e.g. market, customer and other subjects), a complex image - an environment (ENV) is obtained. Than the TOP<sup>®</sup> can be defined as a science discipline covers functional dependences, relations and integration of the processes and of their environments. A recycling is defined reliably if environmental integrated (terminal - principal & supplementary) functions come through defined ENVs. The fragments are such the cycles, which don't create closed loop, or which don't pass through defined ENV. The environment (ENV) isn't only pure surroundings (inversion) of process system, but it includes process system integration to various ENVs, see Figure 2 (Urbánek 2002).

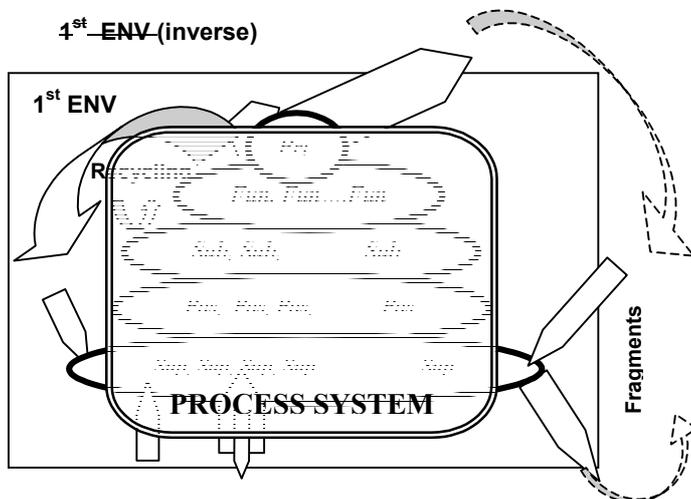


Figure 2. The blazon of modal function's chain

## CONCLUSIONS AND EXAMPLES

### 1st example

The best is to introduce ToP terminology at the example (Urbánek 2002). The lovemaking of two persons of opposite sex is ten thousands years repeated *process*. And nearly every one understands it. He is Adam. She is Eve. And the both are fundamental (Fun) entities of lovemaking process. It is routine to divide a system to the subsystems or elements, and elements to the objects and subjects, in the Theory of System. But the processes are *indivisible* basically, its structure create *function chains* of the *entities*. It is possible to show however, that all dividing is highly dependent on discriminative level. Therefore, the Adam is a subject (from his subjective point of view) and the Eve is a object. But it is contrariwise exactly from the point of view of the Eve. Their joint object is a love at first (later it may be children, substance, money etc. - the time relativity of object/subject dividing follows from this fact). Then it will be better to term the love as the most often *attribute* of the lovemaking. The attribute expresses a process state. Further entity of a drama (the process) named lovemaking is an *environment*. Here the environment is formulated at Judaic - Christian anthroposophy theme. But other themes (Islamic, Buddhism etc.) are possible to imagine here. An act of Adam and Eve lovemaking follows from their *relationship*. Adam with Eve correlate (they have a *liaison*) at the lovemaking, but long term-married couple can have pure *bond* already only (marriage certificate, children, common house, common apartment, a thrift etc). Their relation affects at the fields of different callings, compulsions, predispositions, appeals, jealousies, violence, manias, nostalgias and different modalities of the love. They all are different strong and oriented in dependencies on time, space and information - so on *relation fields*. Relation fields are in ToP named (modelled and symbolized) - the *Blazons*. Technologic-physical act (technology) of the lovemaking (coitus) is an *operation*, because it courses in real-time (on-

line) in real space and with real factors (the Agents) - and their instruments, in some case with aids, prophylactics etc. One from the agents is process **control programme** also, which operates by means of chemical-biological-physiological attributes (the hormone) and with psychological modality (the hunger). This programme is a realization of potential (virtual) **plan**, which is recorded at genetic data media (in chromosomes). Minimally one from process agents starts this programme by means of cross-section function – to **decide**. The **modality** is expressed in relation to all entities of lovemaking process. For example: **environmental modality** - (environment can be puritan, unfriendly, friendly, promiscuous etc.); the modality of process technology - (the lovemaking can be passionate, bashful, violent etc.); the modality of most often attribute of lovemaking - (the love can be early, unrequited, smashing, exalting etc.). An **icon** symbolizes of process function in the ToP©. The functions accomplished during lovemaking are structured to five modalities in function chain. **Principal** function of lovemaking can be only one (not fungible) and its finding isn't so simple way. Basically, it must be spare from excessive influence of discriminative level and in addition it would to give something to the environment. Principal function of the lovemaking offers - to **reproduce** (so how it's in the case of all natural creature). But the people are considerably diverse in this. It causes the anthroposophy modality of Adam and Eve relation, which are simultaneously living in human civilization. It causes the culture, technology and intellect. It is possible to say, that expressive majority of human couples make the love for principal function of this process - to **associate** (to make a couple).....(This example can continue at the Conference.)

## 2<sup>nd</sup> example

Author's research activities (following the ToP© creation) contribute to competitiveness and sustainability of the development of SME industry. Here industry's role is not only in the collaboration but also in its bringing and integration of Project proceedings. Typically cross-sectored Project arise along the industrial value chain, so that technology uptake and innovation are more effectively ensured across European industry. The core of Project was recycling of all waste, which produces the technology of Abrasive Waterjet (AWJ) Cutting System. Strategic contributions of the Project significantly include: (a) modernization of industry and adaptation to change; (b) substantially improve overall quality and environmental impacts; (c) minimize resource consumption. The former solutions were logistic cost ineffective and may cause environmental risk. The Project realization will offer a modular under-one-roof Device (AQUArec®) – see Figure 3) as a recycling system, which will be able to carry-out recycling not only abrasive but even a water and other waste component, include others pollutant and human health harmful substances (noise, dust, bacteria etc.). In this way the AWJ together with Device will be promoted as a clean (without waste) technology.

AWJ process environmental integration, behaviour and influence is tested and evaluated by means of process logistic flows, relations, interactions, inputs, duration, outputs, integration and a consequences at Figure 3

(Urbánek and Koutný 2000). It serves to computer aided environmental evaluation of AWJ processes. The core is abrasive recycling, which is made with the help of patented AQUAscr®system. Graph at Fig.3 is used for AWJ waste recycling analysis, evaluation and management. A synthesis of environmental models of AWJ recycling processes is an instrument for optimised processing way of the abrasive here. Complications of environmental relations and their consequence are expressed by material flow productivity in real time, space and information relation vector fields – the Blazons (Urbánek 2002). The information is shown at graphic simulation model - blazon, very well arranged and computer-able at mathematical set theory.

## REFERENCES

- Hammer, M. and Champy, J. 1993. "Reengineering the Corporation: a Manifesto for Business Revolution". Harper Collins Pb., Inc. USA.
- Mandelbrot, B.B. 1982. „The fractal geometry of nature”. W.H. Freeman, San Francisco.
- Rýznar, B. and Urbánek, J.F. 2002. "Combat Simulation Terminology Shift to Process Crisis Management", MESM'2002 International Conference, Sharjah, U.A.E., in print
- Urbánek, J.F. 1999. "The Recycling of Waste on Multi-grade PLC Structure", 1st International Conference on Solid Waste, Technology, Safety, Environment, Rome, Italy, ICSW –99-1803.
- Urbánek, J.F. 1999. "New Instrument of Integrated Waste Management – DYVELOP", 15th International Conference on Solid Waste Technology and Management, December 12-15, 1999, Philadelphia, PA U.S.A., ISSN 1091-8043.
- Urbánek, J.F., and Koutný, O. 2000, "Environmental System of AWJ After cut-process abrasive recycling," 15<sup>th</sup> International Conference on Jetting Technology, British HR Group, Ronneby, Sweden, ISBN 1 86058 253 2.
- Urbánek, J.F. and Skála, Z. 2001. "Theory of Processes® - new Language of Human Sustainable Development", *International Containment & Remediation Technology Conference and Exhibition*, USA, Orlando.
- Urbánek, J.F. 2002. „Theory of Processes – a Management of Environments“, Czech language book, edition of the CERM Brno, ISBN 80-7204-232-7.

## AUTHOR BIOGRAPHY

JIRÍ URBÁNEK (J.F. Urbanek – English pseudonym) was born in Pelhřimov (Czech Republic) and went to the Technical University of Brno. Here he studied mechanical technology and obtained his degrees in 1972 and PhD in 1983. 15 years he worked in industry as a technologist mainly. Since 1990 he teach as the Associate Professor at Brno University of Technology in the branches Mechanical Technology. His specialised fields of scientific qualification are Mechanical Technology, Non-conventional Technology of Machining, Production Logistics, Automation, Methods of Industrial Engineering

and Management, Environmental Management, Computer Aided Process Engineering & Management, Forensics Engineering. His professional memberships are “Association of Czech Engineers”, CZ; “International Society of Water Jet Technology”, A Federation of National Associations and Societies, Gloucester, ON, Canada; “International Society of Solid Waste Management, Research”, at Pontificia Universitas Urbaniana, Rome, Italy. His contemporary pedagogical activities course at 3 Czech Universities:

(1) Brno University of Technology, faculty of Mechanical engineering: Member of the Committee for State Examinations; Lectures of engineer’s study: Mechanical and Non-conventional Technology, Automation of Production Processes, Production Logistics, Methods of Industrial Engineering and Management. Managing of 6 doctoral students: Non-conventional Technologies (laser, plasma, AWJ, ultrasound...), Industrial Management, Production Logistics, Methods of Industrial Engineering, Environmental Engineering, Process Engineering & Management & Design, Computer Aided Process Management.

(2) Military University of Army Ground Forces Vyškov, Faculty of Military Control Systems, CZ: Member of University Scientific Board; Chairman of the Committee for State Examinations; Lectures of engineer’s study:

Process Engineering & Management & Design, Industrial Executive Information Systems, Enterprise Resource Planning. Computer Aided Process Management.

(3) Brno Business School, CZ: Lectures of engineer’s study for foreigner - Industrial Process Planning.

Managing of significant scientific projects:

Since 1999 to 2002; “Abrasive Water Jet Cutting - a clean technology”, responsible participant of grant project No. ERB IC15 CT98 0821, Grant of European Community “INCO – Copernicus”, Head solver – Universität Hannover, Institut für Werkstoffkunde, Jet-Cutting, Appelstrasse 11, 30 167 Hannover, Germany

Since 2000; “An Implementation of Eco-environmental Management System to Machining and Metallurgical Companies“, Number of Project 101/00/0890, leder of scientific grant project of Czech Grant Agency, Národní 3, 110 00 Praha 1, Czech Republic

Special qualification: forensic specialist – state, price, value, economic and technical estimation of movable, machines, accessories, plants. Environmental engineering.

Publication activities: 115 specialised papers at International Conferences and in the journals. 4 books. Last his book (Urbánek 2002) outlines revolution approach to universal process relativity. Its overwrought English version will be print during 2003 year.

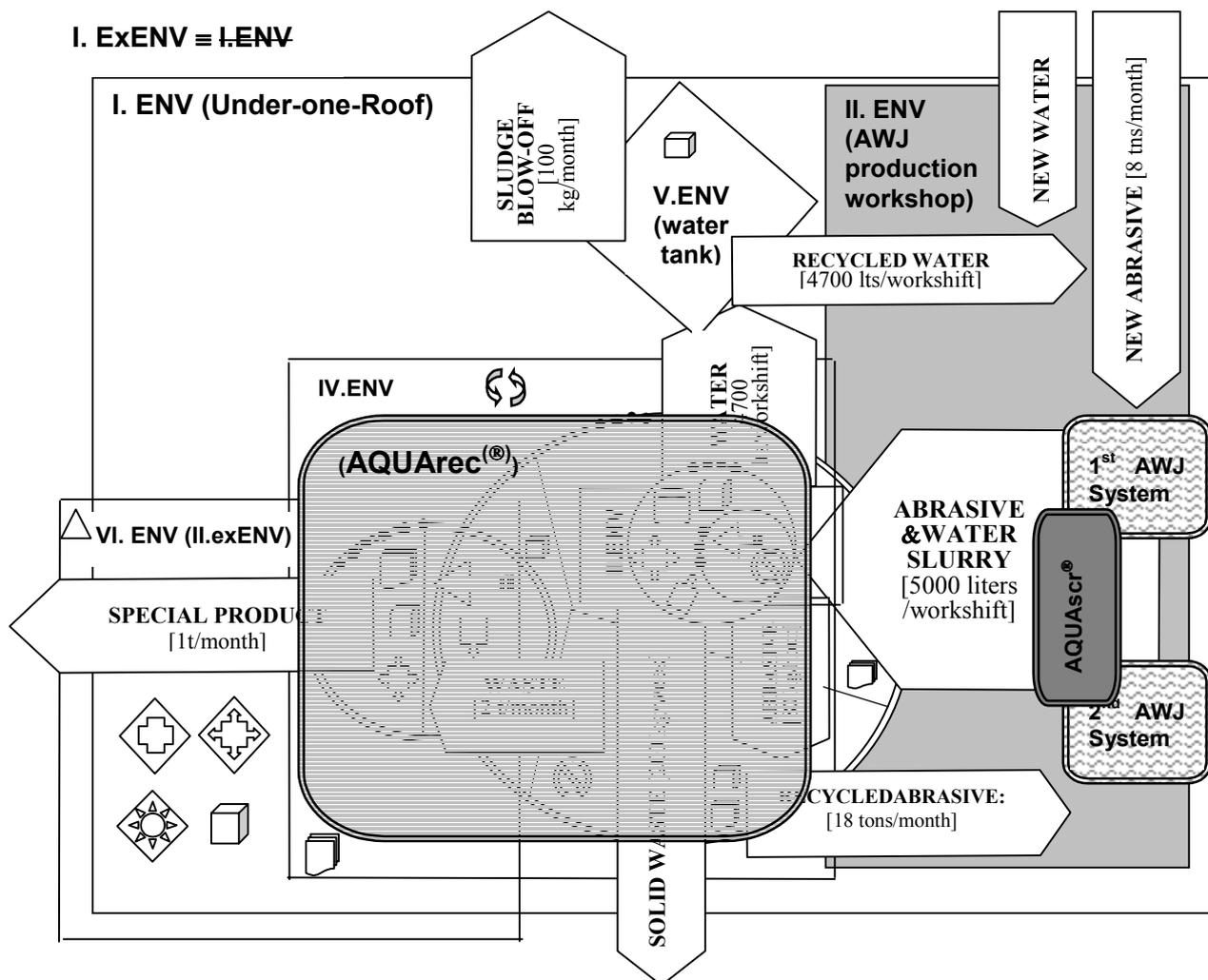


Figure 3. Under-one-Roof (indoor) recycling processes blazon.



# **LATE PAPERS**



# CALCULATING THE CURRENT IN CONDUCTING SEMI-CONDUCTORS MODELLED BY PERFECT SWITCHES.

Faouzi BOULOS  
Department of physics  
University Lebanese of Science II  
Fanar-Lebanon BP 90656, Jedeidet-El-Maten, Lebanon  
Email: faboulos@ul.edu.lb

## KEYWORDS

Semi-conductor modeling, power electronic circuits, numerical simulation.

## ABSTRACT

The aim of this paper is to describe a numerical simulation method of power electronic circuits where the semi-conductors are simulated by perfect switches. The method consists in simplifying the initial schema of the circuit and setting up of the mathematical formulation and development of the automatic simulation. A method is exposed based on a topological analysis of the circuit allowing the extraction of the relations and the calculation of the current in a conducting semi-conductor.

## NOTATIONS AND ABBREVIATIONS USED.

K Number of passive branch, X Sequence number of the blocked semi-conductor, Z Variable equalizes to 1 for the side of the node of output and to 2 for the side of the node of input of the blocked semi-conductor.

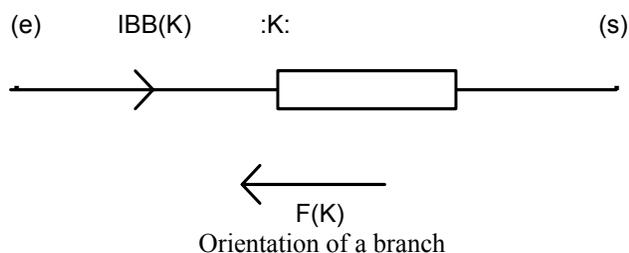
SCE(K) Vector containing the maximum number of branches connected to the node of input of the conducting semi-conductor number K.

SCS(K) Vector containing the maximum number of branches connected to the node of output of the conducting semi-conductor number K.

ISCE(I) Vector of dimension representing the total number of branches connected to the node of input of the conducting semi-conductor number I.

ISCS(I) same definition as ISCE(I) but for the side output.

BNC(I) Contains the new command of the conducting semi-conductors to know the current in all the conducting semi-conductors. (n) Number of node and :n: Number of branch. In all that follows, one needs to direct the branches. To direct a branch we invite (e) the node of input of the branch and (s) the node of output of the branch. We always indicate the orientation by an arrow going from the node of input towards the node of output.



## INTRODUCTION

During the Seventies and Eighties several publications presented methods for the simulation of power-electronics circuits but none used a method of total simulation with variable topology. In 1985, the study "Synthesis of the methods of simulation of the static inverters" (Bordry 1985), stated that a method of simulation with variable topology did not exist and was very delicate to realize. This problem was addressed in my thesis in 1988 "Study and realization of a simulation program of static inverters" (Boulos 1988), where a simulation method based on variable topology was proposed and successfully used.

The choice of the electric model of a semi-conductor is very important. There are several ways to represent a semiconductor which are all electrically equivalent (Antognetti and Massobrio 1988). Similarly, all semiconductors may be represented by equivalent electric circuits more or less complex, which can be used in the electronic circuit simulation as for example in SPICE (Nagel and Pederson 1973). In the field of power electronics it is important to know the global function of the elements as switches, either "on" or "off", and the used model is that of binary impedance. Some of the most frequently used models are: binary variable resistance; depending whether the semi-conductor is "off" (high resistance) or "on" (low resistance). This method is mostly used in the simulators (Schonek 1977), controlled voltage source or current source (Lakatos 1979), inductance in series with a parallel RC circuit (Rajagopalan 1978). In all these cases the topology is fixed and the circuit and the size of the calculating matrix are fixed. Another method to represent the semi-conductor is to replace it by an open circuit when it is "off" and by a short-circuit when it is "on", thus, the topology is variable. Each condition has its set of equations, which reduces the size of the calculating matrix considerably (Boulos 2001). While in the variable topology method the semi-conductors are simulated by perfect switches, the graph and the equation system are variable (Boulos 2001). The difficulty is due to the fact that in such a method the semi-conductors are simulated by perfect switches, that is two nodes connected by a conducting semi-conductor are merged and two nodes connected by a turned-off semi-conductor have their branch removed, sometimes leaving some nodes pending. The branches ending by pending nodes are also removed. This makes the knowledge of the current in a conducting semi-

conductor and the knowledge of the voltage across a turned off semi-conductor difficult to obtain. This point is resolved by the algorithm named TOPOVAR "TOPOlogy VARIABLE". I have exposed the algorithm at MESM 2001 (Boulos 2001), it is possible to modify the complete topology of the circuit at each sequence of operation. In this paper I expose my contribution for the method based on a topological analysis of the circuit allowing the extraction of the relations calculating the current in a conducting semi-conductor.

### THE VARIABLE TOPOLOGY METHOD

In the method of variable topology, the semi-conductors are simulated by perfect switches. Shown in fig.1. When the semi-conductor is "on", we join the two nodes linked together by a semi-conductor, when the semi-conductor is "off", we take off the link between the two nodes. At each step, the topology of the electric circuit is simplified compared to the original graph. The number of differential equations is lower than the one obtained by the fixed topology simulation.

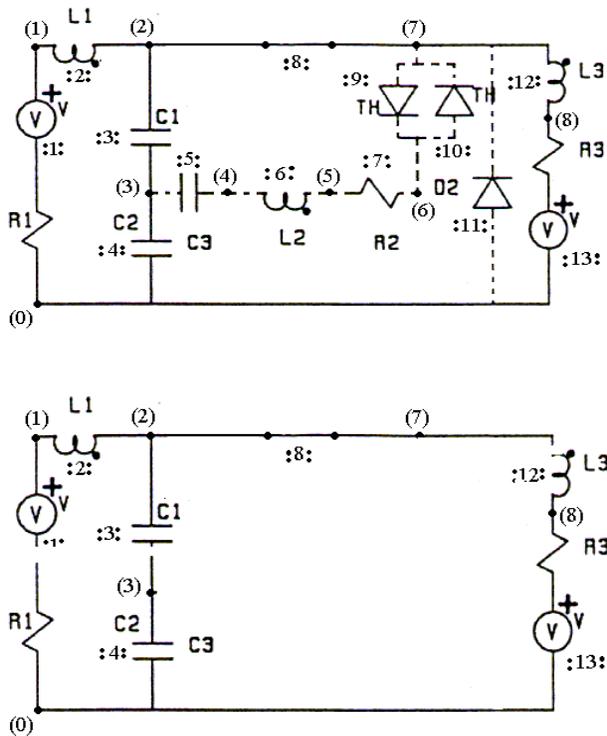


Fig. 1

### THE FIXED TOPOLOGY METHOD

In this type of simulation, the graph is fixed and the equation system is unique, yet, some values of the coefficients may change according to the operating point of the semi-conductors. They are considered high resistance when they are "off" (R) and low resistance when they are "on" (r) fig(2).

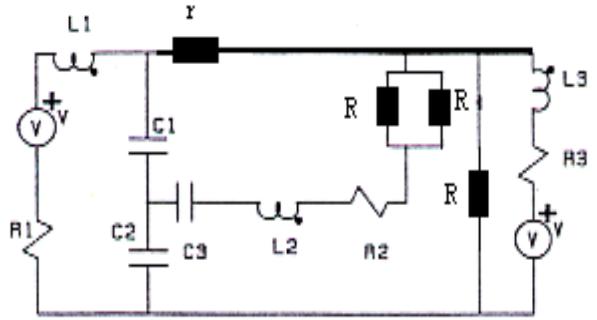


fig. 2

### STATE PROBLEMS FOR THE REMOVAL OF THE CONDUCTOR SEMI-CONDUCTORS.

Each conducting semi-conductor is short-circuited. The branches which connect the nodes of input and output are then attached to same the node (fig. 3).

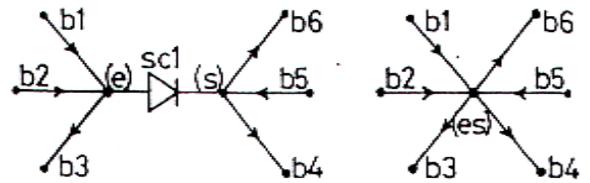


Fig.3: Removal of a conducting semi-conductor.

The principal problems are:

- 1 determination of the new form of minimal topology, the modifications in front to be carried out on new the nodes blocked semi-conductors.
- 2 determination of the value of the current which crosses a conducting semi-conductor in short circuit. To facilitate the explanation of the followed step, we prefer to study each problem separately.

### New form of topology and modification of the numbers of nodes.

The algorithm of the method is as follows (fig.4):

- 1 We traverse the circuit connects by branch. For each conducting semi-conductor met, we decrease by a unit the total number of branches and nodes.
- 2 We seek among the blocked semi-conductors those which have a node of input or output linked to the nodes of this conducting semi-conductor. For each blocked semi-conductor concerned, we modify the new value of the node of input or output which takes the value of the one of both nodes of the conducting semi-conductor.
- 3 We seek the number and the number of all the branches connected to the node of input and we store these values in a matrices MATENT(X,X1) and SCE(K) to know the current which cross a conducting semi-conductor. We do similarly for the nodes of output as we store in matrices MATSOR(X,X1) and SCS(K).

The parameters are:

- K = number of the conducting semi-conductor.
- X sequence number of the conducting semiconductor.

X1=varying variable integer of 0 until the maximum number of branches connected has this node.

-4 We link both nodes of input and output, then we modify the numbers nodes of all the branches connected to these nodes. We preserve the node of output if this semiconductor is not connected to the node a reference. In the opposite case, it is the node of input of this conducting semiconductor which is concerned.

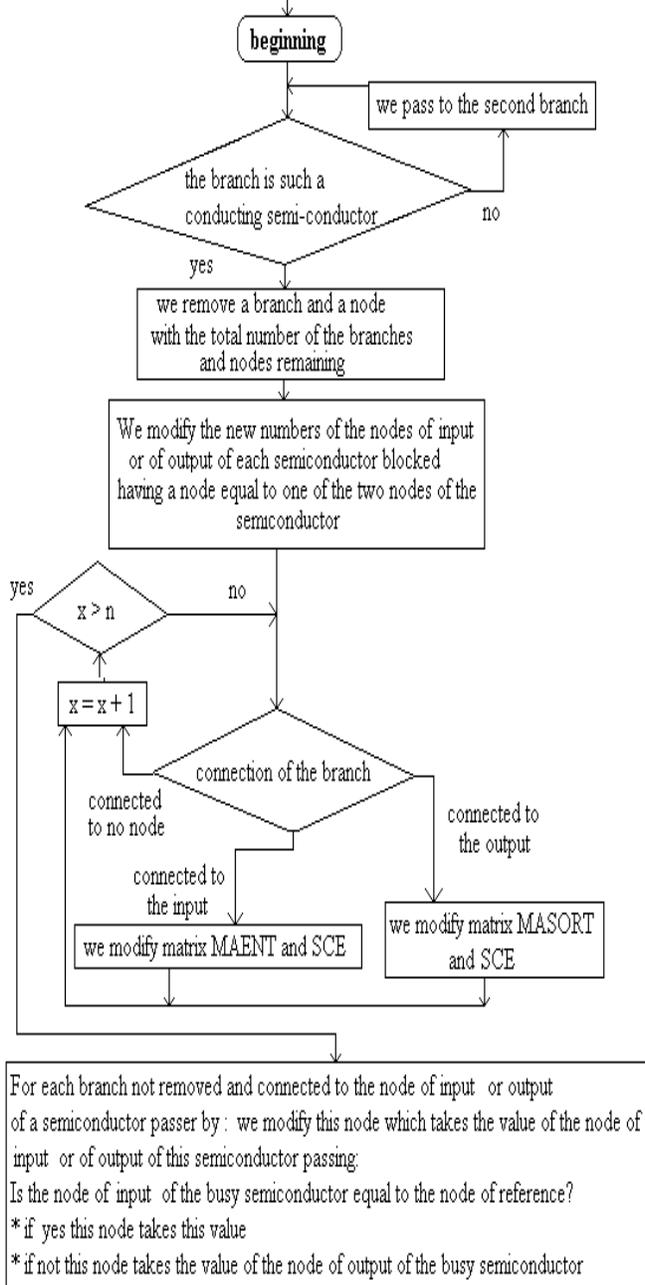


Fig.4 Flow chart of required of information concerning the semiconductors conducting and them branches connected to the nodes.

**Automatic determination of the value of the currents which cross each removed conducting semiconductor.**

The current in a semi-conductor is the algebraic sum of the currents of the branches which connect the node of input or the node of output. For that, we use the numbers of the branches stored in matrices MAENT(X,Y) and SCE(X) for

the side input and matrices MASORT(X,Y) and SCS(X) side output. This determination is more delicate if the branches which connect the node of input and/or output are common to another conducting semi-conductor for one does not know, a priori, the current which cross it (fig.5). It is this problem which we solve now.

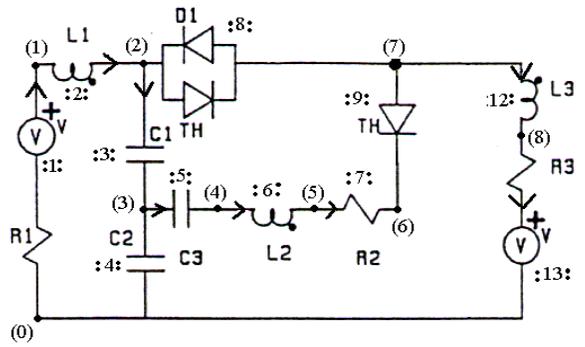


Fig. 5: Semi-conductors of the branches :10: and :11: blocked and branches :8: and :9: conducting.

We choose among the conducting semiconductor a branch for which the current can be calculated before the others. This choice is carried out by traversing all the conducting semiconductor, by sorting matrices MAENT(X,X1) and MASORT(X,X1). For each conducting semiconductor, we seek if one from its two nodes is connected only to passive branches or active branches already put in vector BNC(X).

- If so, this conducting semiconductor is treated. One stores it in BNC(X). Moreover we fill the vector choice such as:

ISCE(X) = 1 ---> if this side is that of the input.

ISCS(X) = 1 ---> if this side is that of the output.

- If not, we pass to another conducting semiconductor.

One repeats the preceding process until all the conducting semiconductor are treated. The command is then such as the calculation of the current in each conducting semiconductor becomes possible.

To carry out the algebraic sum of the currents, we define a matrix containing the sign of the current of the branches connected to the node of input or output of the conducting semiconductor. This matrix is SIGC(X,X1):

X is a specific coefficient of the sequence number of the conducting semiconductor.

X1 a specific coefficient of the sequence number of the branch connected to the nodes of input or output of the semiconductor.

**Example.**

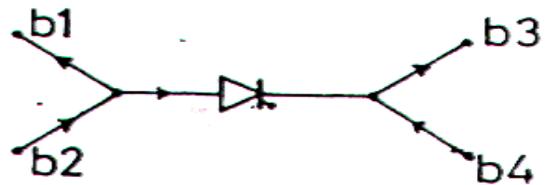


Fig.6: Reconstitution of a current in a semiconductor.

That is to say the circuit (fig. 6) including two passive branches :b1: and :b2: for which one knows the value of the current. Then, with:

$X = 1$  (sequence number of the semi-conductor).  $X1 = 1$  sequence number of the branch :b1:. and  $X1 = 2$  sequence number of the branch :b2:.

$SIGC(1,1) = -1$  and  $SIGC(1,2) = +1$ . It is to be noticed that the knowledge of the information stored in matrices  $MAENT(X,X1)$ ,  $MASORT(X,X1)$ ,  $SIGC(X,X1)$ ,  $ISCE(X)$ ,  $SCE(X)$ ,  $SCS(X)$  and  $BNC(X)$  enable us automatically to know the value of the currents which at every moment cross all the conducting semi-conductors of calculation. For that, we apply one of two following equations:

$$I_s ISC(X) = ISC(X) + SIGC(X,X1) * IBB(MASORT(X,X1)) \quad (1)$$

for the side of output of the conducting semi-conductor (fig.7).

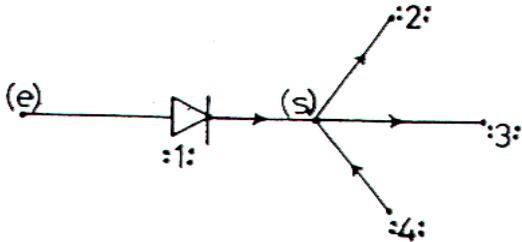


Fig.7: Reconstitution of the output current.

$$I_s ISC(X) = ISC(X) + SIGC(X,X1) * IBB(MAENT(X,X1)) \quad (2)$$

for the side of input of the conducting semi-conductor (fig.8).

### RENUMEROTATION OF MINIMAL TOPOLOGY.

After having removed all the blocked semi-conductors and all the branches connected to a nodes of degree 1 like all the conducting semi-conductors, we connected the nodes of input and output and modified the numbers of the nodes of all the branches connected to this conducting semi-conductor. It is now a question of renumbering the branches and the nodes which form minimal topology. We represent the graph by means of the two matrices line  $ENN(X)$  and  $SII(X)$ . Terms  $ENN(X)$ , and  $SII(X)$  are the numbers of the nodes of input and of output of the x-ième branch. The total capacity of memory necessary for this representation is very low ( $2NN$  integers to store where  $N$  is the number of branches). Knowing that vectors  $BB(I)$ ,  $ENN(I)$  and  $SII(I)$  contain the numbers of the branches and the new classification of the nodes of input and output,  $MM$  and  $NN$  the number of nodes and branches which represents minimal topology.

### FORMING AND SOLVING THE EQUATIONS

In order for the program to be automatic, it must establish the equations by itself. This is achieving in two steps: The first step consists of the topological study of the circuit, and the second step consists of establishing the equations. An electronic circuit is representing by a group of branches connected together and the equations are establishing using mathematical techniques. This enables the analysis of a great number of different circuits, for

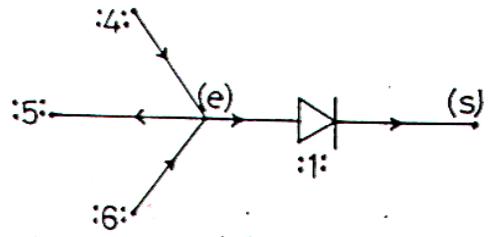


Fig.8: Reconstitution of the current of input.

In these equations:

$ISC(X)$  contains the value of the current which cross the conducting semi-conductor number  $X$ .

$SIGC(X,X1)$  contains the sign of the current which cross the branches connected to the conducting semi-conductor for the side of input or output.

$X$  is a coefficient which varies value 1 until the maximum number of conducting semi-conductors.

$X1$  is a coefficient which varies from 1 until the number of branches connected to the nodes of input or output.

$MASORT(X,X1)$  is a matrix which contains the number of the branch connected to the node of output of the semi-conductor number  $X$ .  $MAENT(X,X1)$  is a matrix which contains the number of the branch connected to the node entered of the semi-conductor number  $X$ .

$IBB(Y)$  is a matrix which contains the value of the current of the branch number  $Y$  calculated with each step of calculation.

which the formation and solving of the equations would otherwise require long and tiring calculations. It is not possible to list all mathematical techniques and numerical calculations in this article. For the application, we chose the setting in equation by the nodal method for it asks less memory.

### Nodal analysis method.

Several simulators use this method, which is based on the solution of the matrix associated to the equations representing the electronic circuit. The matrix is solved by the method of the pivot. SPICE (Antognetti and Massobrio 1988). The nodal analysis method is used by several simulators because it is simple and powerful.

In this method, the equation system is in the form of:

$$(I) = (G) * (V)$$

(I) Vector of electrical current.

(V) Vector of the voltages of the nodes compared to the reference nodes.

(G) Matrix of the admittance.

### RESULTS

In order to illustrate the mentioned methods, the results obtained by simulating a chopper fig.9 are shown.

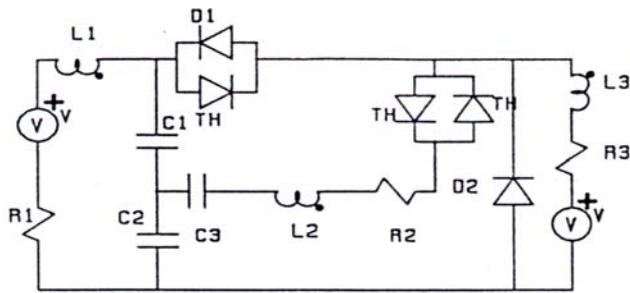


fig.9

The states of the semiconductors are represented by 0 for the semiconductors which are "off" and 1 for the semiconductors which are "on".

For: D2 TH1 TH2 THP

$$T=0.000E+0 \quad 1 \quad 0 \quad 0 \quad 0 \Rightarrow VD=8, \overline{W} (4,5)$$

$$T=1.000E-5 \quad 1 \quad 0 \quad 0 \quad 1 \Rightarrow VD=9, \overline{W} (4,5)$$

$$T=1.376E-5 \quad 0 \quad 0 \quad 0 \quad 1 \Rightarrow VD=1, \overline{W} (6,7)$$

$$T=3.000E-5 \quad 0 \quad 0 \quad 1 \quad 1 \Rightarrow VD=3, \overline{W} (4,5)$$

$$T=5.927E-5 \quad 0 \quad 0 \quad 0 \quad 1 \Rightarrow VD=1, \overline{W} (6,7)$$

$$T=1.200E-4 \quad 0 \quad 1 \quad 0 \quad 1 \Rightarrow VD=5, \overline{W} (7,8)$$

$$T=1.513E-4 \quad 0 \quad 1 \quad 0 \quad 0 \Rightarrow VD=4, \overline{W} (6,7)$$

$$T=1.733E-4 \quad 1 \quad 1 \quad 0 \quad 0 \Rightarrow VD=12, \overline{W} (4,5)$$

We regard a thyristor and an antiparallel diode as only one switch THP.

Fig. 10 represents the results of the simulation using the variable topology.

We present the form of the currents and the tensions for the capacitive branch C3 and the diode D1 and the antiparallel thyristor TH.

Table I shows the time ratios obtained when using the example of the chopper fig.9 The gain is between 2 and 3.

$\overline{W}()$	$\delta_n\%$	$\delta_t \mu_s$	$\frac{\delta}{\mu=TT\%}$
(20)	27%	95 $\mu_s$	47,5 %
(72)			
(42)	58%	75 $\mu_s$	37.5%
(72)			
(52)	77%	30 $\mu_s$	15 %
(72)			

Table I

Where:

$\overline{W} (n-1,n)$ : calculating matrix corresponding to the minimal topology.

$W(n-1,n)$ : calculating matrix corresponding to the fixed topology.

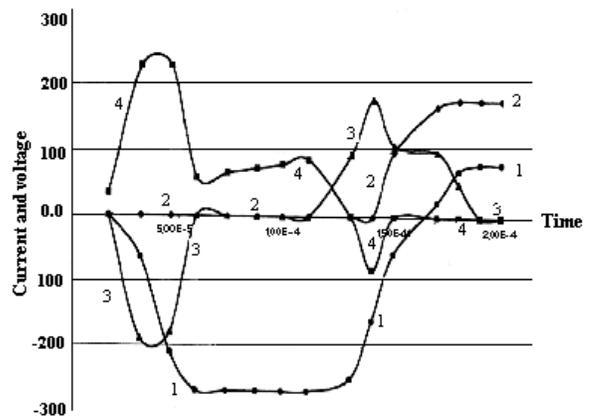
$\delta_n\%$ : Percentage of the number of elements used for the variable topology.

$\delta_t$ : Is the simulation time used for  $\overline{W} (n-1,n)$  during a simulation period  $2.10^{-4}$  s.

$\mu$ : Percentage of the simulation time.

TT: Simulation period.

## CONCLUSION.



- 1- Voltage across a condensator C3.
- 2- Voltage across a diod D1 and the antiparallel TH.
- 3- Current in a condensator C3.
- 4- Current in a diod D1 and the antiparallel TH.

The work presented in this article has as an aim to show the difficulties and the solutions suggested for the determination of the current crossing the conducting and remove dsemi-conductors. The characteristic of this method is there presentation of the semi-conductors in the form of perfect switches. The advantages which this type of modeling brings are a notable simplification of the system of equation to be solved, a precision of calculation of the current in a conducting semi-conductor and a greater speed of simulation compared to the methods with fixed topology. The secondary times constants by the modeling of semi-conductor ( $r$  one / R of) by fixed topology are avoided.

Work was carried out in two stages: determination of minimal topology according to the state of the semi-conductors then the reconstitution of the currents which cross the conducting semi-conductors. The second stage is the setting in automatic equation and the resolution of these equations. And as a result the application to an assembly chopper is presented.

## REFERENCE

- BOULOS Faouzi 1988 "Etude et réalisation d'un programme de simulation de convertisseurs statiques. Utilisation d'une méthode à topologie variable". Thèse de docteur de l'Université Paris VI.
- BOULOS Faouzi 2001 "Numerical simulation of power electronic circuit. "Use of variable topology method. Modelling of the semiconductor as perfect switches". MESM2001.
- Schonek J 1977. " Simulation numérique des convertisseurs statiques application à la conception et à l'optimisation des convertisseurs". Thèse INP Toulouse
- Lakatos L. Feb. 1979 " A New method for simulating power semi-conductor circuits". IEEE Trans. Ind. Electro. Vont. Inst. Vol IECI - 26 n 1.
- Rajagopalan 1978 "Simulateur digital des convertisseurs de puissance à thyristors" Can. Elec. Eng. J. n 1 PP 5-10
- L. N. Nagel, D-O. Pederson 1973 "SPICE" Simulation program with integrated circuit emphasis" Berkeley california : Memorandum ERL, M 382 Apr 12.
- Paolo ANTOGNETTI, Giuseppe MASSOBRIO 1988 "Semi-conductor Device Modeling with SPICE"

McGraw-Hill Book.

P. PELLETIER. 1971 "Techniques numeriques appliquées au calcul scientifique". Masson .

F.BORDRY. 1985 Thèse de docteur ès-sciences ENSEEIHT Toulouse.

# ON DCT-BASED VS. WDCT-BASED IMAGE COMPRESSION

Ljupcho Panovski  
Zoran Ivanovski  
Aleksandar Velev

Institute of Electronics, Faculty of Electrical Engineering,  
University "SS. Cyril and Methodius" of Skopje, Karpos II bb,  
1000 Skopje, Republic of Macedonia.  
E-mail: panovski@ukim.edu.mk

## KEYWORDS

Image processing, Data compression, Filtering,  
Transforms, Signal processing.

## ABSTRACT

The paper addresses a recently proposed image compression algorithm that uses the Warped Discrete Cosine Transform and presents the results of the experimental verification we have performed. The results of the experiments have demonstrated that the actual performance of the algorithm is quite lower when quantization effects are taken into account, and that the proposed algorithm has a very limited field for successful application.

## INTRODUCTION

Thanks to its excellent features of energy compaction and coefficient decorrelation, the discrete cosine transform (DCT) has been the basis of many standard and non-standard image and video compression techniques (Bhaskaran et al. 1997), (ISO/IEC 1993), (CCITT 1992). By having an image-independent kernel, the DCT offers the advantage of fast computation, but also the disadvantage of not being able to perform optimal energy compaction for certain images having dominant high-frequency components. An effort to overcome this was recently proposed by the introduction of a modification of the DCT, named the warped discrete cosine transform (WDCT), along with a corresponding WDCT-based image compression algorithm (Cho and Mitra 2000).

## THE WARPED DISCRETE COSINE TRANSFORM

WDCT is in fact a cascade of an all-pass filter and DCT, the role of the all-pass filter being to warp the frequency distribution of the input signal in such a manner as to make it more suitable for the DCT. More precisely, if the 8-point DCT is implemented as in Figure 1, then the 8-point WDCT is implemented as in Figure 2.

The all-pass filter  $A(z)$  is the well-known Laguerre filter (Masnadi-Shirazi 1991):

$$A(z) = \frac{-\alpha + z^{-1}}{1 - \alpha z^{-1}} \quad (1)$$

where the value of parameter  $\alpha$  can be chosen to adjust the direction and the extent of frequency warping. Positive

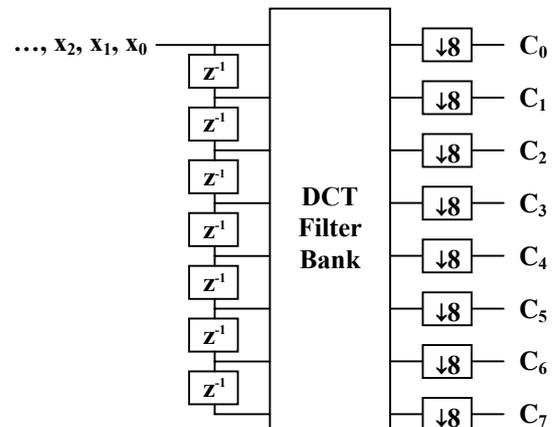


Figure 1 Implementation of DCT

values of  $\alpha$  warp a low-frequency input so as to make it more suitable for the DCT, negative values of  $\alpha$  have the opposite effect, and  $\alpha = 0$  reduces WDCT to DCT. This

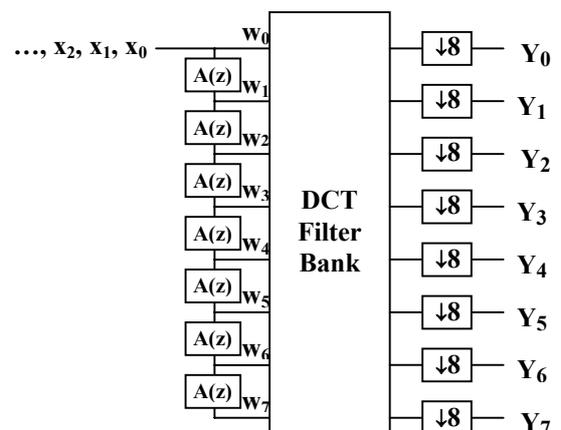


Figure 2 Implementation of WDCT

can be seen in Figures 3, 4 and 5, where, for illustration purposes, the coefficients of a 4-point WDCT for  $\alpha = 0$ , i.e. of DCT, and for  $\alpha = 0.5$  and  $\alpha = -0.5$  are given, respectively.

## THE WDCT-BASED IMAGE COMPRESSION ALGORITHM

The WDCT-based image compression algorithm follows in general the steps of the well-established JPEG processing:

- the image to be compressed is divided into blocks of 8x8 pixels each,
- each 8x8 block of the image is subjected to transformation,
- the transform coefficients are quantized to reduce the amount of data to be stored or transmitted (by dividing each coefficient by the corresponding element of a quantization matrix),
- the sequence of quantized coefficients is encoded by an entropy coder (Huffman) to obtain a compressed representation of the image.

The distinctive feature of WDCT-based compression scheme is that, instead of using a fixed DCT on each block, as is done in JPEG, each image block undergoes 16 WDCT and 16 IWDCT for 16 different values of  $\alpha$  between -0.1 and 0.1. The value of  $\alpha$  yielding the best fit is being sent to the decoder along with the WDCT coefficients for that particular value of  $\alpha$ . Thus, there is an overhead of 4 bits/block, which prevents the algorithm to be effective for low bit rates, and the execution time is roughly sixteen times the execution time of the DCT-based JPEG algorithm.

The quantization of the WDCT coefficients in the WDCT-compression scheme proposed in (Cho and Mitra 2000) is performed by using a uniform quantization matrix (hereinafter: the constant QM), while the JPEG compression algorithm that has been chosen as a reference employs the well-known quantization matrix that had been so defined as to provide the least visibility of the image artifacts introduced by the quantization (hereinafter: the visually optimized QM). The authors in (Cho and Mitra 2000) justify their using of two different quantization matrices when comparing the performance of the two different compression algorithms by claiming that the effects of the visually optimized QM in JPEG are equivalent to the effects of the frequency weighting performed by WDCT in the proposed WDCT-based algorithm.

On the basis of experiments performed by compressing images by the WDCT-based compression algorithm (which has been implemented with the constant QM) and by the DCT-based JPEG compression algorithm (which has been implemented with the visually optimized QM), the authors in (Cho and Mitra 2000) draw the conclusion that, thanks to the frequency warping feature of the WDCT, their proposed WDCT-based image compression algorithm, although slower, performs better than the conventional DCT-based JPEG image compression algorithm for high bit-rate applications and natural images

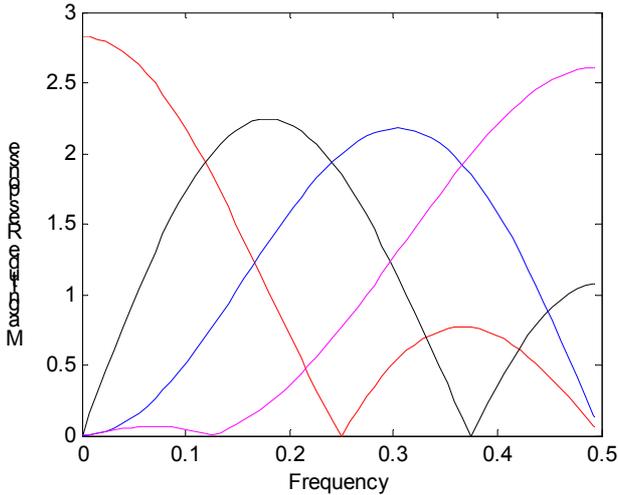


Figure 3 Frequency responses of the WDCT filter bank for  $\alpha = 0$

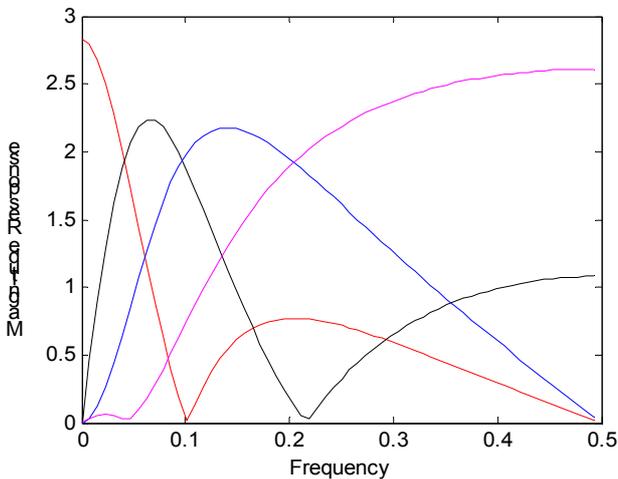


Figure 4 Frequency responses of the WDCT filter bank for  $\alpha = 0.5$

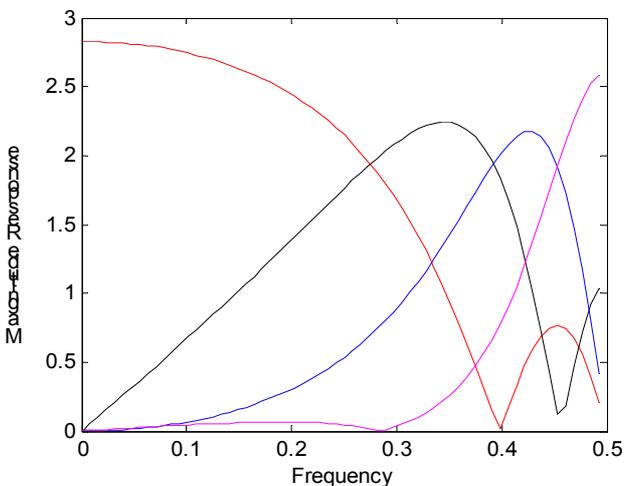


Figure 5 Frequency responses of the WDCT filter bank for  $\alpha = -0.5$

containing significant high-frequency components, the improvement being 1.1 to 3.1 dB better PSNR at 1.5 bpp.

### THE EFFECTS OF THE QUANTIZATION MATRIX

The higher the values of the elements of a quantization matrix, the higher the compression ratio (i.e. the lower the bit rate), and the lower the visual quality of the reconstructed image (i.e. the lower the PSNR or the SNR). Lacking a practical measure for the visual image quality, it is common to use PSNR or SNR instead.

One has to bear in mind that the compression scheme of JPEG has been developed so as to show sufficient performance even at rather high compression ratios (up to 32:1), i.e. low bit rates (down to 0.25 bpp). In order to have sufficient visual image quality at such low bit rates one needs a quantization matrix that tends to "sacrifice" the high-frequency and to "preserve" the low-frequency transform coefficients of the image i.e. that suppresses the coarse image artifacts at the expense of emphasizing the fine image artifacts. This is exactly what is being achieved when the visually optimized QM is used. And because JPEG is normally used for lower bit rates, the use of the visually optimized QM has become an unwritten and unquestioned standard in everyday image compression practice.

On the other hand, if one is to use JPEG compression at higher bit rates (of the order of 1.5 bpp), one could use the constant instead of the visually optimized QM. In doing so one could expect the visual image quality in both cases to be more or less the same, since the visual image quality is in general affected mainly by the coarse image artifacts. At the same time, one could expect the image obtained with the constant QR to have better SNR, since the fine image artifacts, although being below the threshold of the human visual system, should have higher values than in the other image.

Although at first sight the approach to compare two different compression schemes employing two different matrices, as is done in (Cho and Mitra 2000), seems to have a well-grounded justification, the reasoning of the previous two paragraphs has led us to check this approach and the corresponding conclusions empirically.

### EXPERIMENTAL RESULTS AND DISCUSSION

We have implemented both the WDCT compression scheme and the JPEG compression scheme in MATLAB and have tested them on the set of typical images, given in Figure 6, at different bit rates.

Since the two-dimensional transformation is performed as two consecutive one-dimensional transformations, the WDCT-based image compression allows two slightly

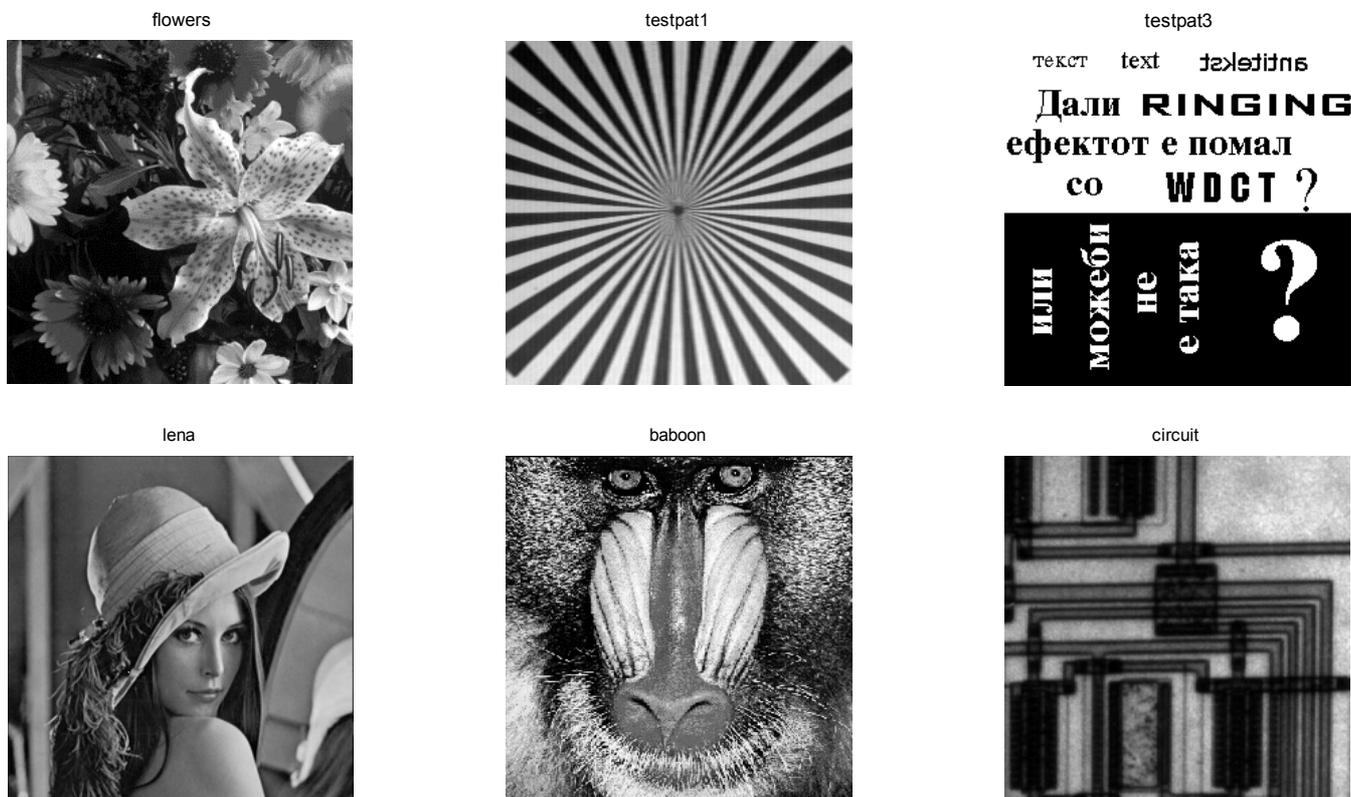


Figure 6 The set of typical images used in the experiments

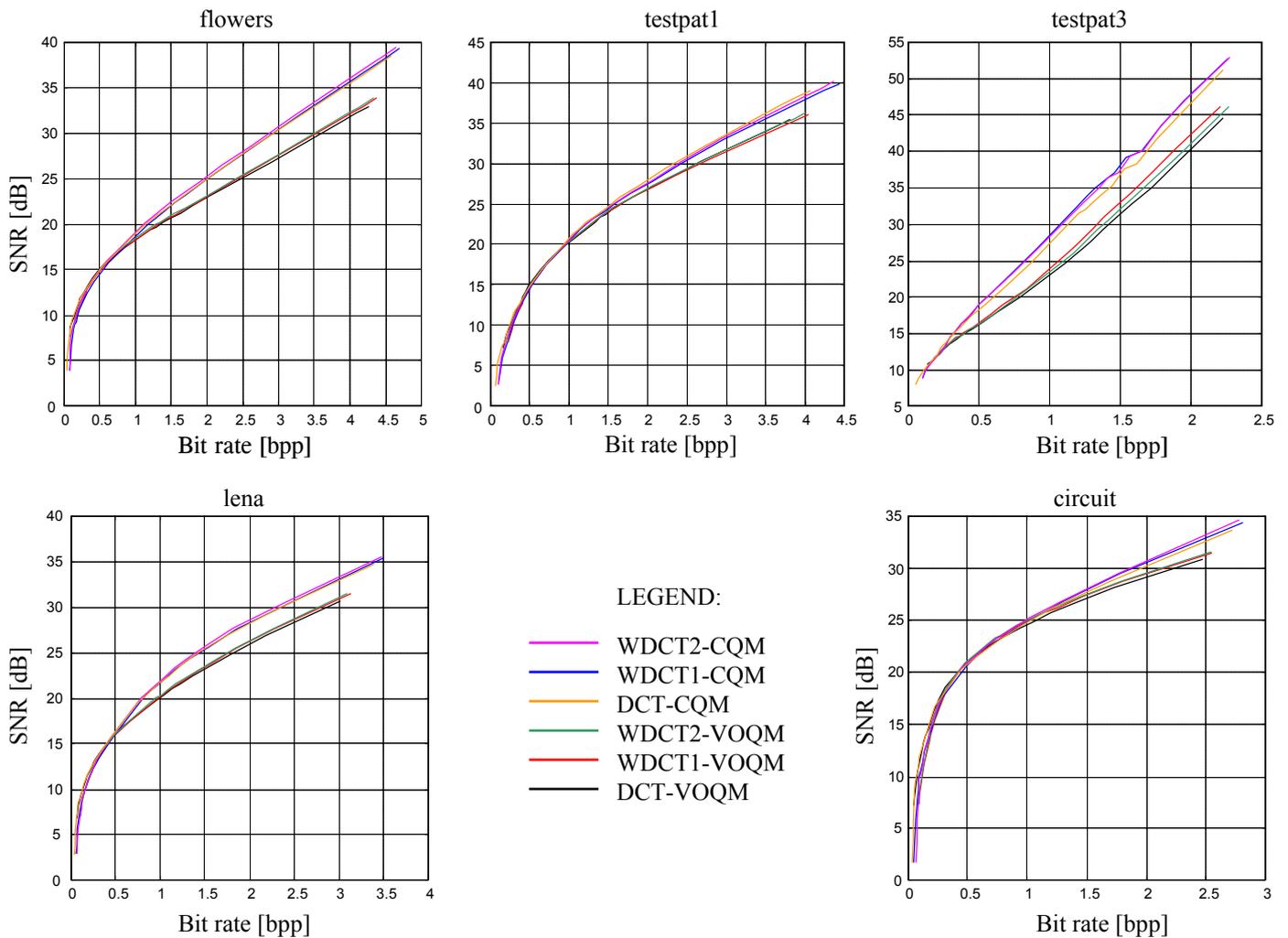


Figure 7 Experimental results

different implementations: by using the same value of  $\alpha$

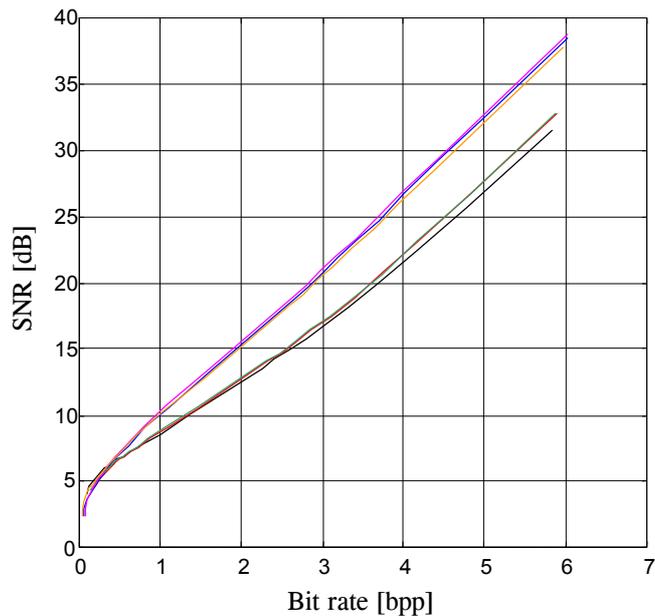


Figure 8 Experimental results for the 'baboon' image

for both dimensions (hereinafter: WDCT1) and by using different values of  $\alpha$  for the horizontal and vertical directions (hereinafter: WDCT2). In contrast to the experiments performed in (Cho and Mitra 2000), both WDCT1 and WDCT2 were performed by using the constant (hereinafter: WDCT-CQM) and the visually optimized QM (hereinafter: WDCT-VOQM). The JPEG compression scheme was also performed by using both matrices (hereinafter: DCT-CQM and DCT-VOQM, respectively).

The results are given in Figures 7 and 8. Because they all have the same general characteristics, we shall discuss them while referring to the results for the 'baboon' image shown in Figure 8.

It can be readily seen that all the curves are grouped into two groups according to the type of quantization matrix and not according to the type of transformation. This means that the improvement in SNR (which is the same as the improvement in PSNR) of over 2 dB for bit rates higher than 1 bpp is almost entirely due to the use of the constant instead of the visually optimized QM, while the contribution of using either WDCT1 or WDCT2 instead of DCT is quite modest.

Our results demonstrate that a gain of a few tenths of dB when using WDCT instead of DCT is a much more realistic figure. Only for special cases of images containing very high frequency components one can expect higher gains over DCT-based compression, but even then they can come nowhere near 3 dB, as claimed in (Cho and Mitra 2000). The maximum gain that we have been able to obtain by specially selecting the suitable 'testpat3' image was 1.5 dB, and it was obtained at a high bit rate of over 2 bpp.

## CONCLUSIONS

After having performed the experiments, we have come to a quite different conclusion than the authors of (Cho and Mitra 2000). Namely, that the 1.1-3.1 dB gain in PSNR at 1.5 bpp of the WDCT-based compression algorithm over JPEG compression is a far too optimistic figure and that it had resulted because the effects of quantization were neglected. When quantization of the coefficients with a constant quantization matrix is performed in JPEG as it is in WDCT-based image compression, the SNR values in both cases become very similar. The same, of course, happens if quantization of the coefficients is implemented in the WDCT-based image compression algorithm as it is implemented in JPEG.

The results of the experiments have demonstrated that the actual performance of the algorithm is quite lower when quantization effects are taken into account, and that the proposed algorithm has a very limited field for successful application.

## REFERENCES

- Bhaskaran V. and Konstantinides K. 1997. *Image and Video Compression Standards 1997*, Kluwer Academic Publishers, Boston.
- CCITT 1992 Recommendation T-81 (1992 E)
- Cho N.I. and Mitra S. 2000 "Warped Discrete Cosine Transform and Its Application in Image Compression", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 10, No. 8, 1364-1373.
- ISO/IEC 1993 10918-1 (1993 E)
- Masnadi-Shirazi M.A. 1991 "Optimum Laguerre networks for a class of discrete-time systems", *IEEE Transactions on Signal Processing*, Vol. 39, No. 9, 2104 - 2108

**LJUPCHO PANOVSKI** was born in Skopje, Republic of Macedonia. He obtained his diploma in electronics engineering and the titles of M.Sc. and Ph.D. from the University of Ljubljana, Republic of Slovenia, in 1970, 1981 and 1991, respectively. Until 1975 he worked for the Macedonian Broadcasting Company before moving to the Faculty of Electrical Engineering at the "SS. Cyril and Methodius" University in Skopje, where he has been working ever since. Currently he is a professor and Head of the Institute of Electronics.

# THE EFFECT OF DISTURBANCE MODEL ON THE CONTROL PERFORMANCE OF A SIMULATED ACTIVATED SLUDGE BENCHMARK PROBLEM

Farideh Ghavipankeh, Green Research Centre, Iran University of Science and Technology,  
Tehran, Resalat Square, Hengam St., University St.,  
f.ghavipankeh@greenrc.iust.ac.ir

## ABSTRACT

Wastewater treatment systems are unreliable in terms of their performance. Many parameters could affect their efficiency in which, amongst them, the influent flow variations could considerably change the performance of the system. In this paper the model of influent flow disturbance has been added to the control system structure in order to alleviate the performance problem. The control system is PIP/TDC, which is published in numerous papers and also described briefly here. All experiments have been applied to an activated sludge simulation benchmark and the results show improvements in the control objectives.

## INTRODUCTION

The nonlinearity of wastewater treatment systems together with variable environmental conditions makes their performance highly variable. This leads to under standard run-times of the systems, which in turn, causes environmental damages as well as legislation charges. In order to alleviate the problem and also cope with new-hardened regulations, many researches have been conducted to enhance the reliability of wastewater treatment systems and improve their performance. Some have focused on upgrading the systems and alternative systems layout as well as searching for new technology (Pochana, *et al.*, 1999, Stams, *et al.*, 1997). However, some researches concentrate on improving the control strategies and control systems (Ayasa, *et al.*, 1998, Barnett, 1992). Here in this paper we have focused on the control system improvement by adding influent flow disturbance model into the control structure.

A data-based modelling approach that was developed in 1987 within the more general context of True Digital Control (TDC) system design (see, Young *et al.*, 1987, 1988) is being used here to produce a linear transfer function model. This approach exploits the latest methods of time series analysis technique. In this technique, a reduced-order transfer function model is directly identified from the data obtained either from a real system or from a complex simulation model of the system. The methodology is applied to a simulated wastewater treatment benchmark problem. Since biological wastewater treatment systems are nonlinear processes with enormously complex reactions, and also due to the problems in identifying an accurate mechanistic model and unavailability of all the state variables of biological treatment systems, the data-based modelling technique could be favourable in the viewpoint of model

identification and parameter estimation. Indeed, this technique seems a practical, simple, cheap and less time-consuming method to identify a proper model.

Statistical modelling has also been used by other researchers such as McMichael and Vigani (1972), Berthouex *et al.* (1978), Berthouex and Box (1996), Van Dongen and Geuens (1998), but these models were mainly developed for the purpose of forecasting and fault diagnosis, rather than utilisation directly in the control loop.

The *WORKING GROUP No.2* within the framework of the European COST Action 682 "Integrated Wastewater Management" has developed a benchmark simulation for evaluating different control strategies for activated sludge plants (e.g. Alex *et al.*, 1999). The benchmark is a platform-independent simulation environment defining wastewater treatment plant layout, simulation model, influent data, test procedures and evaluating criteria for comparing control performance (Pones *et al.*, Alex *et al.*). .

Here, the benchmark problem is used to introduce the load disturbance model, mainly the influent flow rate into the control loop. The dissolved oxygen (DO) control of the benchmark is being investigated here. The controller that is used for the dissolved oxygen in the last zone of the system is a Proportional-Integral-Plus (PIP) controller (Young *et al.*, 1987; Young *et al.*, 1991) within the context of TDC methodology. The PIP DO controller settings and also the results of that controller on the benchmark have already been published in Ghavipankeh, *et al.*, 2000b.

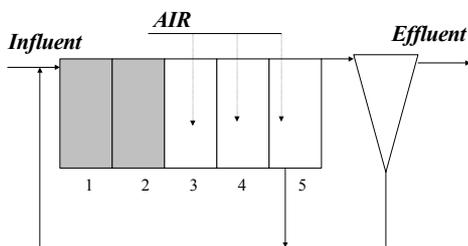
There are two main forms of PIP controller; *standard* and *forward path* structures. The research implemented by Taylor *et al.* (1996) implies that the forward path structure is useful when the output measurements are noisy so that the noises are not amplified within the controller function thus, providing a smoother control input. However, when the system is suffering from high amplitude and frequency of load disturbances, which is the case in wastewater treatment systems, the forward path PIP is unable to provide a good closed loop performance. Since the data obtained from wastewater treatment systems are often noisy, the forward path structure seems to be appropriate if somehow tackle with the input load disturbances. In this paper, an improvement in the forward path structure of PIP is being suggested by adding the disturbance model into the control loop in order to anticipate the output variations more accurately.

## THE BENCHMARK PROBLEM

The benchmark system is based on the layout of a typical activated sludge plant for nitrogen removal. It is composed of a five-compartment bioreactor for pre-denitrification (the first two anoxic zones) and nitrification (the last three aerobic zones), together with a settler at the end of the process. The benchmark was downloaded from the following web site, from which further information may be obtained: [www.ensic.u-nancy.fr/COSTWWTP/benchmark.html](http://www.ensic.u-nancy.fr/COSTWWTP/benchmark.html)

Figure 1 shows the general layout of the benchmark system. The first two zones have no aeration and provide denitrification processes while the last three zones are aerated and provide nitrification reactions. The benchmark simulation is based on the two well-known models; the Activated Sludge Model no. 1 (ASM no.1 or IAWQ model) (Heinze *et al.* 1987) for modelling the biological part of the process; and, the Tackas model (Tackase *et al.* 1991) for modelling the settler dynamics. The ASM model consists of 13 state variables and 19 parameters. In this model, the wastewater and biomass are structured into a number of different variables to represent the effect of different groups of bacteria on the different component of the wastewater.

The benchmark model is simulated using the MATLAB/SIMULINK™ software system. A verification routine was applied in order to check the consistency of the simulation with the benchmark model. This involved the line by line check of the SIMULINK diagram with the model and also comparing the static and dynamic response of the simulation subject to constant and dynamic inputs with the results released by the *WORKING GROUP*. The dynamic input files provided by the *WORKING GROUP* contain the diurnal variation of some of the input components such as influent flow rate for a period of 14 days and under different weather conditions. These files were supplied in order to verify the simulation and also to evaluate the proposed operation or control strategy.



## CONTROL OBJECTIVES

There are two common objectives for the benchmark system; control of dissolved oxygen in the last zone and control of nitrate in the second zone by means of the aeration rate to the last zone and the returned flow rate from zone 5, respectively. The application of PIP/TDC methodology on the benchmark has already been released in Ghavipankeh and McCabe, 2000, Ghavipankeh, *et al.*,

2000a, Ghavipankeh, *et al.*, 2000b. In Ghavipankeh, 2001, it is shown that the forward path PIP control of DO and nitrate produces a much less noisy control input than the standard form. However, the output variation is higher and less stable (see also, Taylor, *et al.*, 1996). In this paper it is shown that the dissolved oxygen control loop with the forward path PIP controller on can be improved by introducing the disturbance model into the control loop. Hence, the control objective here is to maintain the DO level at 2 mg/l, which is a previously defined value of DO controller in the benchmark.

## DISTURBANCE MODEL

The influents of wastewater treatment systems are usually highly variable in terms of their rate, concentration and composition. It has shown that the influent flow rate is mainly the most important variable that largely affects the systems performance than the concentration and composition of influents (see, Ghavipankeh, 2001). Here, the relationship between the influent flow rate and dissolved oxygen is modelled using a data-based modelling technique. We will show that this relationship can be well approximated by a simple linear transfer function (TF) model.

For a linear, single-input, single-output (SISO) discrete-time system,

$$y(k) = \frac{b_1 z^{-1} + \dots + b_m z^{-m}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}} u(k) = \frac{B(z^{-1})}{A(z^{-1})} u(k) \quad (1)$$

where  $A(z^{-1})$  and  $B(z^{-1})$  are appropriately defined polynomials in the backward shift operator  $z^{-1}$ ; i.e.,  $z^{-i}y(k) = y(k-i)$ . Here,  $y(k)$  and  $u(k)$  are the dissolved oxygen concentration (mg/l) and influent flow rate ( $\text{m}^3/\text{d}$ ) respectively.

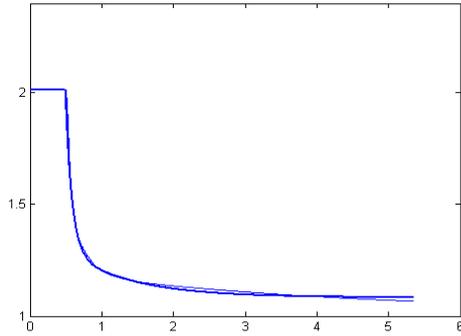
In the present benchmark example, the data are obtained from simulation experiments conducted on the high order nonlinear model of the system. For the present analysis, an equilibrium point is first obtained by running the simulation model for a long period of time (over a day) with constant inputs and no disturbances. Secondly, the simulation model is perturbed about this operating point by a step applied to the control input signal. Finally, the resulting input-output data are used as the basis for statistical model identification and estimation.

Here, the input-output data are collected in one-minute sampling rate. The statistical Simplified Refined Instrumental Variable (SRIV) identification and estimation algorithm (Young 1991) is used here and result is a second order model as follows;

$$y(k) = \frac{-0.1935 * e^{-5 * z^{-1}} + 0.1933 * e^{-5 * z^{-2}}}{1 - 1.9914 z^{-1} + 0.9914 z^{-2}} u(k) \quad (2)$$

The Coefficient of determination,  $R_T^2$ , together with the Young Identification Criterion (YIC) (Young 1991) determines the best-identified model. A value of  $R_T^2$  close to unity and a large negative YIC are normally desired that

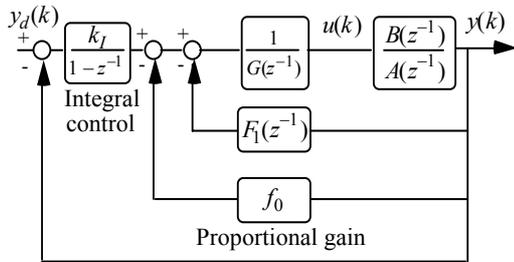
guarantee goodness of fit and not over-parameterising of the model. The values of  $R_T^2$  and YIC for the obtained second order model are 0.998 and  $-16.53$  respectively. As clear also from figure 5, a very good model fit is obtained.



**Figure 2.** Simulation output (thick trace) compared with the response of the TF model (thin trace); DO (mg/l) is plotted against time days.

### PIP CONTROL SYSTEM

The TDC control design methodology is based on the definition of a suitable non-minimum state space (NMSS) form for the discrete-time model. Here the state variables are defined as the present and past sampled values of the input and output variables, together with the ‘integral of error’. The control methodology is called Proportional-Integral-Plus (PIP) control design, because the structure is the same as a PI controller for a first-order model, but involves additional feedback terms for higher-order models. Figure 3 shows the structure of the PIP controller.



**Figure 3.** The PIP/NMSS controller structure in standard feedback form.

Here  $\frac{B(z^{-1})}{A(z^{-1})}$  is the process transfer function model of the

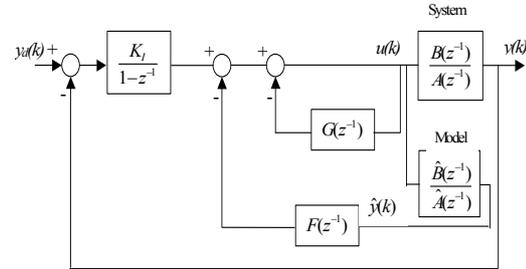
system in terms of backward shift operator ( $z^{-1}$ ),  $K_I$  is the integral action of the controller and  $F_1(z^{-1})$  and  $G(z^{-1})$  are defined as follows:

$$F(z^{-1}) = F_1(z^{-1}) + f_0 = f_0 + f_1 z^{-1} + \dots + f_{n-1} z^{-(n-1)} \quad (3)$$

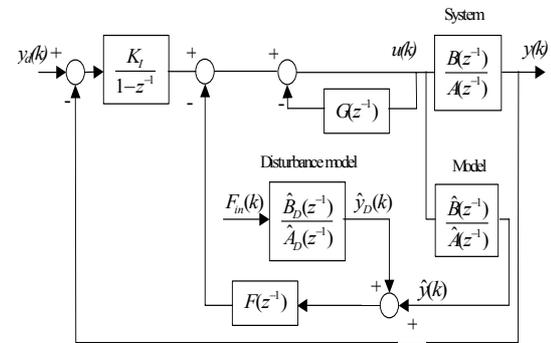
$$G(z^{-1}) = 1 + g_1 z^{-1} + \dots + g_{m-1} z^{-(m-1)}$$

The forward path structure is shown in figure 4 in which the output is being fed back from the model of the system instead of the real system. This avoids the output noise being amplified within the controller function. But, when

there are load disturbances the model of the system in the forward path control loop, do not take into account the effects of the load disturbances which affect the real output. In other words, these effects are only being fed to the integrator, not the other compensators of the forward path control loop. In order to contribute the disturbance model to the control structure the diagram of figure 4 is modified as shown in figure 5 by adding the load disturbance model into the control loop.



**Figure 4.** Forward path structure of the PIP controller.



**Figure 5.** Forward path PIP control structure with disturbance model.

In order to develop a PIP control algorithm, a linearised representation of the plant is required. In the NMSS/PIP approach to control system design, a data-based approach to combined linearisation and model reduction is employed. As mentioned already, this methodology for control of DO in the last zone of the benchmark by mean of aeration rate to this zone, has been implemented in Ghavipankeh *et al.* 2000b, where detailed information can be obtained. Here the DO PIP control settings mentioned in that paper is used to investigate the modified forward path PIP control of DO.

The PIP control design algorithm utilises the State Variable Feedback (SVF) design strategies within the NMSS setting such as closed loop pole assignment or optimisation in terms of a Linear-Quadratic (LQ) cost function.

The PIP-LQ control polynomials for DO as obtained in Ghavipankeh *et al.* 2000b are as follows:

$$F(z^{-1}) = 281.0 - 176.1z^{-1}$$

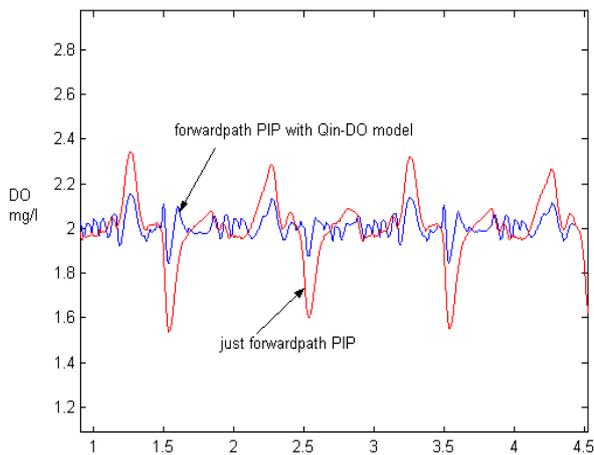
$$G(z^{-1}) = 1 - 0.9878z^{-1}$$

$$K_I = 26.35 \quad (4)$$

These PIP control settings are used to investigate the performance of the forward path PIP controller while, having the disturbance model in the control loop and system is suffering from influent flow rate disturbances.

## RESULTS AND DISCUSSION

The performance of the above PIP controller (4) in the modified forward path PIP control loop (figure 5) is investigated using the dry weather data file as in the benchmark procedure. The result is then compared with the PIP control loop with no disturbance model as shown in figure 6.



**Figure 6.** PIP control of DO with disturbance model (blue trace) and no disturbance model (red trace). DO (mg/l) is plotted against time in days.

As clear, the control output variation of the modified PIP forward path control loop (figure 5) has decreased notably showing a better performance of the controller.

## CONCLUSIONS

This paper has presented an improvement in the application of the univariate PIP methodology to control Dissolved Oxygen (DO) concentration in a well-known benchmark system. A modification is being made in the forward path PIP control algorithm that successfully maintains DO concentration at the required level, despite the presence of significant disturbances to the influent. This has involved the introduction of the disturbance model into the forward path PIP control loop. The disturbance model is a data-based model, identified and estimated using the well proven SRIV statistical identification and estimation algorithm.

## REFERENCES

Alex, J., Beteau, J.F., Copp, J.B., Hellings, C., Jeppsson, U., Marsili-Libelli, S., Pons, M.N., Spanjers, H., and Vanhooren, H., (1999c), Benchmark for evaluating control strategies in wastewater treatment plants, *European Control Conference (ECC) 99*, Karlsruhe, September 1999.

Ayasa, E., Goya, B., Larrea, A., Larrea, L., Rivas, A., (1998), Selection of operational strategies in activated sludge processes based on optimisation algorithms., *Wat. Sci. Tech.*, vol.37, no. 12, pp.327-334.

Barnett, M.W., (1992), Application of artificial intelligent (AI) techniques in activated sludge process design, operations, and control., In: *Dynamics and control of the activated sludge process* (ed. J.F. Andrews), Technomic Publishing Company, Inc., USA.

Ghavipankeh, F., (2001), Modelling and control of wastewater treatment systems., *PhD. thesis*, Lancaster University, UK.

Ghavipankeh, F., McCabe, A., (2000), Data based modelling for the true digital control (TDC) methodology: applications in control of wastewater treatment systems., *ESS' 2000*, Hamburg, Germany.

Ghavipankeh, F., Taylor, C.J., Young, P.C., Chotai, A., (2000a), Data-based modelling and Proportional Integral Plus (PIP) control of nitrate in an activated sludge benchmark., *Wat. Sci. Tech.*, vol. 44, no.1, pp. 87-94.

Ghavipankeh, F., Taylor, C.J., Chotai, A., Young, P.C., (2000b), Modelling and PIP control of dissolved oxygen in a wastewater treatment benchmark system., *UKACC 2000*, Cambridge, UK.

Henze M., Grady Jr, C.P.L., Gujer, W., Marais, G.v.R. and Matsuo, T., 1987, *Activated sludge model No. 1*, Report of IAWPRC, London.

Pochana, K., Keller, J., Lant, P., (1999), Model development for simultaneous nitrification and denitrification., *Wat. Sci., Tech.*, vol. 39, no. 1, pp. 235-243.

Stams, A.J.M., Elferink, S.J.W.H.O., (1997), Understanding and advancing wastewater treatment., *Current opinion in Biotechnology*, vol. 8, pp. 328-334.

Takacs, I., G.G. Patry, and D. Nolasco. 1991, A dynamic model of the clarification –thickening process. *Water Research*, vol.25, no.10: 1263-1271

Taylor, C.J., Young, P.C., Chotai, A., Tych, W., Lees, M.J., (1996), The importance of structure in PIP control design. *IEE*.

Young, P.C., M.A.Behzadi, C.L.Wang and A.Chotai. 1987. "Direct Digital and Adaptive control by input-output, state variable feedback pole assignment." *Ini. Jul. of Control*. no. 86: 1861-1881.

Young, P.C. 1991. "Simplified Refined Instrumental Variable (SRIV) estimation and true digital control (TDC): a tutorial introduction." *Proceeding of the first European Control Conference*, Gernoble, 1295-1306.

Young, P. C., Chotai, A., Tych, W. (1991) True Digital Control: A unified Design Procedure for Linear Sampled Data Systems, In *Advanced methods in adaptive control for industrial applications, Lecture notes in Control and Information Sciences*, K.Warwick, M.Karney and A.alouskova, eds. Vol. 158, springer-verlag: Berlin, 71-109.



# **AUTHOR LISTING**



## AUTHOR LISTING

Abd Rahim R. ....	85	Hertkorn G. ....	183
Abdulaziz N. ....	104/146	Hilal O. ....	88
Addasi J.S.M. ....	172	Hj Man N. ....	85
Adel M.B. ....	35	Ibrahim H. ....	85
Ahmed S. ....	151	Ivanovski Z. ....	223
Al Moudni A. ....	17	Javoronkov M. ....	131
Al-Ahmad H. ....	142	Krajzewicz D. ....	183
Al-Akaidi M. ....	168	Krcum M. ....	121
Alani O. ....	168	Mat Tahar R. ....	85
Al-Bassam M.A. ....	35	Mehic N. ....	69
Al-Hammadi A. ....	30	Mignot B. ....	17
Al-Huwaihi A.S. ....	69	Miller L.K. ....	93
Aljunaibi M. ....	65	Moench L. ....	192
Allagui H. ....	131	Mojallal-A S. ....	126
Al-Mualla M.E. ....	101/142	Montazeri-Gh M. ....	126
Alnsour D. ....	168	Munitic A. ....	121/136
Alrefai M.H. ....	202	Najjar M. ....	161
AL-Rousan M. ....	151	Nielsen E.H. ....	188
Antonic R. ....	136	Oršulic M. ....	121
Awad M.S. ....	176	Otto P. ....	192
Bastaki E. ....	146	Ouni A. A. ....	131
Boulos F. ....	217	Ozkul M.H. ....	109
Chiari F. ....	25	Ozkul T. ....	109
Daley G.C. ....	93	Pang K.K. ....	104
Dawood A.M. ....	115	Panovski L. ....	223
de Lara J. ....	5	Rajpoot N.M. ....	165
Delhom M. ....	25	Rawabdeh I.A. ....	35
Derbel N. ....	43	Razzak M.A. ....	93
Dismukes J.P. ....	93	Rekik C. ....	43
Djemel M. ....	43	Rössel C. ....	183
Dvornik J. ....	121	Rýznar B. ....	73/209
El-Khazali R. ....	142	Saïdane L.A. ....	197
Filipas I. ....	17	Santucci J.-F. ....	25
Ghavipanjah F. ....	228	Schormans J. ....	30
Glass A. ....	104/146		
Hanako-Matcovshi M. ....	88		

## AUTHOR LISTING

Sharar S.R.....	101	Wack M.....	17
Shatnawi A. ....	151	Wagner P.....	183
Sihn W. ....	79	Wellen U. ....	156
Stoklosa J.....	161	Yakoub Z.H.....	35
Störzel M. ....	156	Zainon F. ....	85
Urbánek J.F. ....	73/209	Zarei B.....	61
Van Welden D. ....	53		
Vangheluwe H. ....	5		
Vansteenkiste G.....	5/53		
Velev A. ....	223		
Vitanov V. ....	65		
von Briel R. ....	79		
Vukic Z.....	136		