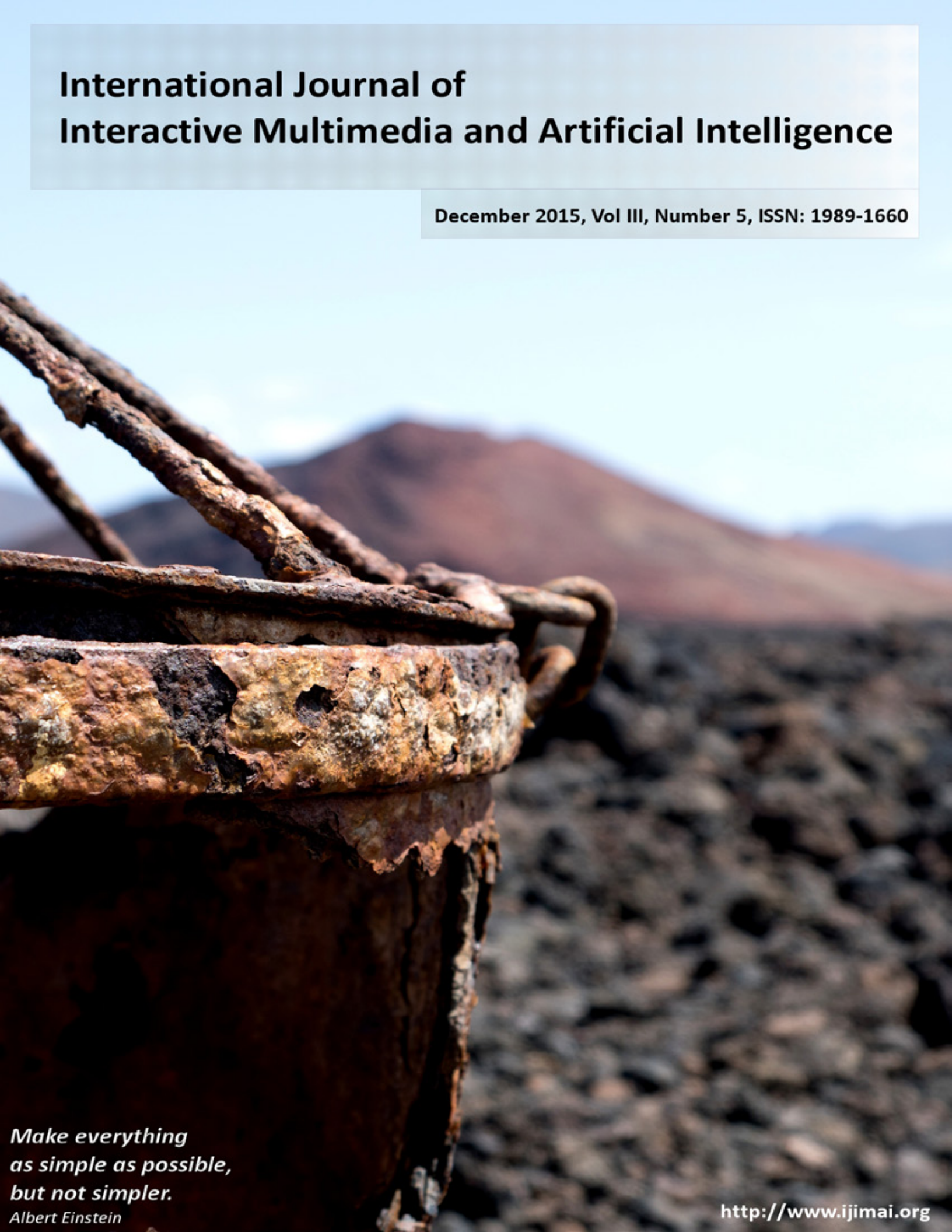
International Journal of Interactive Multimedia and Artificial Intelligence

December 2015, Vol III, Number 5, ISSN: 1989-1660



Make everything as simple as possible, but not simpler.
Albert Einstein

http://www.ijimai.org

INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE AND INTERACTIVE MULTIMEDIA

ISSN: 1989-1660–VOL. III, NUMBER 4

Editor's Note

The International Journal of Interactive Multimedia and Artificial Intelligence provides an interdisciplinary forum in which scientists and professionals can share their research results and report new advances on Artificial Intelligence and Interactive Multimedia techniques.

The research works presented in this issue are based on various topics of interest, among which are included: DSL, Machine Learning, Information hiding, Steganography, SMA, RTECTL, SMT-based bounded model checking, STS, Spatial sound, X3D, X3DOM, Web Audio API, Web3D, Real-time, Realistic 3D, 3D Audio, Apache Wave, API, Collaborative, Pedestrian Inertial, Navigation System, Indoor Location, Learning Algorithms, Information Fusion, Agile development, Scrum, Cross Functional Teams, Knowledge Transfer, Technological Innovation, Technology Transfer, Social Networks Analysis, Project Management, Links in Social Networks, Rights of Knowledge Sharing and Web 2.0.

García-Díaz, V. Et al. [1] talks about build the first step towards a language and a development environment independent of the underlying technologies, allowing developers to design solutions to solve machine learning-based problems in a simple and fast way, automatically generating code for other technologies. That can be considered a transparent bridge among current technologies. They rely on Model-Driven Engineering approach, focusing on the creation of models to abstract the definition of artifacts from the underlying technologies.

Al-asadi, T.A. Et al. [2] presents a new approach for hiding the secret image inside another image file, depending on the signature of coefficients. The proposed system consists of two general stages. The first one is the hiding stage which consist of the following steps. The second stage is extraction stage which consist of the following steps.

Mahmoud, M. A. Et al. [3] writes about an automated multi-agent negotiation framework for decision making in the construction domain. It enables software agents to conduct negotiations and autonomously make decisions. The proposed framework consists of two types of components, internal and external. Internal components are integrated into the agent architecture while the external components are blended within the environment to facilitate the negotiation process. They also discuss the decision making process flow in such system. They finally present the proposed architecture that enables software agents to conduct automated negotiation in the construction domain.

Zbrzezny, A.M. Et al. [4] talks about an SMT-based bounded model checking (BMC) method for Simply-Timed Systems (STSs) and for the existential fragment of the Real-time Computation Tree Logic. They implemented the SMT-based BMC algorithm and compared it with the SAT-based BMC method for the same systems and the same property language on several benchmarks for STSs. For the SAT-based BMC we used the PicoSAT solver and for the SMT-based BMC we used the Z3 solver. The experimental results show that the SMT-based BMC performs quite well and is, in fact, sometimes significantly faster than the tested SAT-based BMC.e agent interaction.

Stamoulias, A. Et al. [5] presents a novel method for the introduction of spatial sound components in the X3DOM framework, based on X3D specification and Web Audio API. The proposed method incorporates the introduction of enhanced sound nodes for X3DOM which are derived by the implementation of the X3D standard components, enriched with accessional features of Web Audio API. Moreover, several examples-scenarios developed for the evaluation of our approach. The implemented examples established the achievability of new registered nodes in X3DOM, for spatial sound characteristics in Web3D virtual worlds.

Ojanguren-Menendez, P. Et al. [6] talk about the real-time collaboration which is being offered by multiple libraries and APIs (Google Drive Real-time API, Microsoft Real-Time Communications API, TogetherJS, ShareJS), rapidly becoming a mainstream option for web-services developers. However, they are offered as centralized services running in a single server, regardless if they are free/open source or proprietary software. After re-engineering Apache Wave (former Google Wave), we can now provide the first decentralized and federated free/open source alternative. The new API presented allows to develop new real-time collaborative web applications in both JavaScript and Java environments

Anacleto, R. Et al. [7] shows a pedestrian inertial navigation system that is typically used to suppress the Global Navigation Satellite System limitation to track persons in indoor or in dense environments. They propose a system that uses two inertial measurement units spread in person's body, which measurements are aggregated using learning algorithms that learn the gait behaviors. In this work they present our results on using different machine learning algorithms which are used to characterize the step according to its direction and length. This characterization is then used to adapt the navigation algorithm according to the performed classifications.

Schön, E. M. Et al. [8] present a study that was carried out with 175 interdisciplinary participants from the IT industry. For the evaluation of the results, 93 participants were included who have expertise in the subject area Agile Methodologies. On one hand, it is shown that the collaborative development of product-related ideas brings benefits. On the other hand, it is investigated which effect a good understanding of the product has on decisions made during the implementation. Furthermore, the skillset of product managers, the use of pair programming, and the advantages of cross-functional teams are analyzed.

López-Cruz, O. Et al. [9] present and demonstrate in use a methodology based in complex network analysis to support research aimed at identification of sources in the process of knowledge transfer at the interorganizational level. The importance of this methodology is that it states a unified model to reveal knowledge sharing patterns and to compare results from multiple researches on data from different periods of time and different sectors of the economy. The resulting demonstrated design satisfies the objective of being a methodological model to identify sources in knowledge transfer of knowledge effectively used in innovation.

Magaña, D. Et al. [10] will carry out a literature review of papers that use Artificial Intelligence as a tool for project success estimation or critical success factor identification. Project control and monitoring tools are based on expert judgment and parametric tools. Projects are the means by which companies implement their strategies. However project success rates are still very low. This is a worrying situation that has a great economic impact so alternative tools for project success prediction must be proposed in order to estimate project success or identify critical factors of success. Some of these tools are based on Artificial Intelligence. In this paper we will carry out a literature review of those papers that use Artificial Intelligence as a tool for project success estimation or critical success factor identification.

Gil López, E. Et al. [11] explores the last developments of the legal system concerning these issues. Knowledge sharing among individuals has changed deeply with the advent of social networks in the environment of Web 2.0. Every user has the possibility of publishing

DOI: 10.9781/ijimai.2015.350

what he or she deems of interest for their audience, regardless of the origin or authorship of the piece of knowledge. It is generally accepted that as the user is sharing a link to a document or video, for example, without getting paid for it, there is no point in worrying about the rights of the original author. It seems that the concepts of authorship and originality is about to disappear as promised the structuralisms fifty years ago. Nevertheless, the legal system has not changed, nor have the economic interests concerned.

Dr. Rubén González Crespo

REFERENCES

[1] García-Díaz, V. Et al. "Towards a standard-based domain-specific platform to solve machine learning-based problems", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 6-12, 2015.

[2] Al-asadi, T.A. Et al. "A New Approach for Hiding Image Based on the Signature of Coefficients", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 13-22, 2015.

[3] Mahmoud, M. A. Et al. "An Automated Negotiation-based Framework via Multi-Agent System for the Construction Domain", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 23-27, 2015.

[4] Zbrzezny, A.M. Et al. "Checking RTECTL properties of STSs via SMT-based Bounded Model Checking", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 28-35, 2015.

[5] Stamoulias, A. Et al. "Wrapping" X3DOM around Web Audio API", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 36-46, 2015.

[6] Ojanguren-Menendez, P. Et al. "Building Real-Time Collaborative Applications with a Federated Architecture", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 47-52, 2015.

[7] Anacleto, R. Et al. "Step Characterization using Sensor Information Fusion and Machine Learning", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 53-60, 2015.

[8] Schön, E. M. Et al. "Agile Values and Their Implementation in Practice", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 61-66, 2015.

[9] López-Cruz, O. Et al. "A Network based Methodology to Reveal Patterns in Knowledge Transfer", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 67-76, 2015.

[10] Magaña, D. Et al. "Artificial Intelligence applied to Project Success: a Literature Review", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 77-82, 2015.

[11] Gil López, E. Et al. "Legal Effects of Link Sharing in Social Networks", International Journal of Artificial Intelligence and Interactive Multimedia, vol. 3, no. 5, pp. 83-86 2015.

TABLE OF CONTENTS

OPEN ACCESS JOURNAL

ISSN: 1989-1660

COPYRIGHT NOTICE

Towards a standard-based domain-specific platform to solve machine learning-based problems

Vicente García-Díaz, Jordán Pascual Espada, B. Cristina Pelayo G-Bustelo, and Juan Manuel Cueva Lovelle

Department of Computer Science, University of Oviedo, Oviedo, Spain

***Abstract* — Machine learning is one of the most important subfields of computer science and can be used to solve a variety of interesting artificial intelligence problems. There are different languages, framework and tools to define the data needed to solve machine learning-based problems. However, there is a great number of very diverse alternatives which makes it difficult the intercommunication, portability and re-usability of the definitions, designs or algorithms that any developer may create. In this paper, we take the first step towards a language and a development environment independent of the underlying technologies, allowing developers to design solutions to solve machine learning-based problems in a simple and fast way, automatically generating code for other technologies. That can be considered a transparent bridge among current technologies. We rely on Model-Driven Engineering approach, focusing on the creation of models to abstract the definition of artifacts from the underlying technologies.**

***Keywords* — Domain-Specific Language, Model-Driven Engineering, Integrated Development Environment, Machine Learning, Artificial Intelligence, Xtext**

I. Introduction

Artificial Intelligence (AI) refers to the "intelligence" provided by software included in some machines [1]. It is also a field of study commonly defined as the design of intelligent agents, which perceives their environment and takes actions that maximize their possibility of success [2]. The general problem of creating intelligence can be divided into different sub problems: 1) deduction and reasoning; 2) knowledge representation; 3) planning; 4) social intelligence; 5) natural language processing; 6) perception; 7) motion and manipulation; 8) long-term goals; or 9) machine learning.

Machine learning is one of the most important applications of artificial intelligence that evolved from the study of pattern recognition and computational learning theory. The goal is to create and study algorithms that are capable of leaning from data and make predictions on its basis [3].

Designing and implementing algorithms for machine learning is not a trivial tasks. M. Mitchell stated that a computer program is said to learn from experience E with respect to some class of tasks T and performance P, if its performance with tasks T, as measured by P, is improved with experience E [4].

In addition, machine learning is closely related to many other areas such as computational statistics for prediction-making or mathematical optimization, making the number of people involved in using related techniques very diverse with different backgrounds. Thus, there is a large number of solutions using different approaches to deal with machine learning-based problems. For example, Encog is a machine learning framework available for Java, .NET and C++ programmers [5], and Weka is a workbench that contains a group of graphical tools for data analysis and predictive modeling [6]. Moreover, there are also used General-Purpose Languages (GPL) such as Python, that is suitable for students and is one of the most popular introductory programming languages [7]. On the other hand, Domain-Specific Languages (DSL) [8] such as R, are also used for machine learning tasks [9] and their relevance continue growing.

However, although there is a great amount of solutions to deal with machine learning-based problems, all of them seem to be difficult to be used by no-experts programmers or require users to learn different technologies or applications that make the knowledge they have about a tool or platform virtually useless when they need to work with another one, when circumstances require it. This is even more problematic when the solution should be done programmatically for better control and adaptation.

Hence, different tools and software development approaches continuously appear in the software engineering field, trying to abstract the development from specific platforms or technologies (e.g., virtual machines, APIs, frameworks, etc.). It is widely considered that the Model-Driven Engineering (MDE) approach, with which the level of abstraction of developments is increased through the use of models, it is a step forward in the development of software [10], since developments are being benefited from the advantages provided by MDE (e.g., in García-Díaz et al. [11] food traceability systems for different clients are created in a quick and dynamic way).

MDE is based on the use of models, which conform to a single domain-based metamodel, which in turn are defined based on a common meta-metamodel, root of all the elements of any software development. That idea makes up the architecture of four layers defined in the Model-Driven Architecture (MDA) standard [12]. The common base allows for a wide range of supported environments and tools working together. As a result, if a metamodel for a specific knowledge domain is defined (e.g., food traceability or machine learning), it would be possible to create a DSL based on MDE tools [13], designed only to define the important specific items (e.g., food manufacturing processes or features of neural networks). Internally, the use of standard-based modeling technologies allows direct and automatic transformations to different formats or platforms defined by different software manufacturers. There are a variety of research in MDE that serve to advance in the systematic use of DSLs. For example, Cueva et al. work on bringing together the MDE approach and the Internet of Things field creating languages for automatic vehicle data capture [14][15] or García.-Díaz et al. work on improvement match algorithms for performing further operations with models [16].

The main aim of this paper is to take the first step towards the creation of a standard-based platform for defining and abstracting machine learning-based solutions in a simple and common way. Internally, definitions are automatically transformed into different

DOI: 10.9781/ijimai.2015.351

languages or platforms. Thus, the specific goals are:

1. Identify the basic elements that a representation of a language for solving machine learning-based problems must possess.

2. Create a DSL to define machine learning-based solutions. We call it AiDSL.

3. Allow automatic transformation of definitions made with AiDSL to any other platform or system.

4. Provide an Integrated Development Environment (IDE) to work with AiDSL. We call it AiIDE.

5. Study the advantages of the proposal by a comparison with other alternatives.

The remainder of this paper is structured as follows: in Section 2, we present a description of the relevant state of the art (goal [1]); in Section 3, we describe our proposal (goals [2-4]); in Section 4, we discuss a comparison of the proposal with other alternatives (goal [5]) and finally, in Section 5, we indicate our conclusions and future work to be done.

II. Background

There are a large number of approaches to deal with machine learning-based problems such as: 1) decision tree learning; 2) association rule learning; 3) artificial neural networks; 4) inductive logic programming; 5) support vector machines; 6) clustering; 7) Bayesian networks; 8) reinforcement learning; 9) representation learning; 10) similarity and metric learning; 10) sparse dictionary learning; or 11) generic algorithms.

In this work we focus on Artificial Neural Networks (ANN), that have been used to solve a great variety of problems that are difficult to solve using other techniques [17]. They can be defined as statistical learning models inspired by biological neural networks. Typically, there are presented as collections of interconnected neurons, sending messages each other. Each neuron has numeric weights that can be set using different algorithms, being them adaptive to inputs, or what it is the same, allowing them to learn.

Regarding ANNs, the Feedforward neural network was the first and the simplest type of ANN formulated. In such a type of network the information moves only in one direction. In this work, we focus on the Feedforward artificial neural network, although there are some other such as Elman Neural Network or the Jordan Neural Network, interesting depending on the type of problem to be solved.

Fig. 1 shows a small example of an artificial neural network, which can be decomposed into different layers, containing each a specific number of neurons with similar properties.

- Input layer. Typically it has one neuron for each attribute that the network will use for obtaining different kinds of solutions (e.g., classification, regression or clustering). In the example, I1 and I2 are input neurons included in the input layer.
- Output layer. It provides the output after all previous layers have processed the input. In the example, O1 is the only output neuron that is included in the output layer.
- Hidden layers. They are inserted between input and output layers and are used to better produce the expected output for the given input readjusting weights. In the example, H1 and H2 are hidden neurons contained in the only hidden layer shown.

In addition, there are also bias neurons that can be inserted in the input and hidden layers as desired (B1 and B2 in the example). They are very similar to the hidden neurons but are a special kind that allow the neural network to learn patterns more effectively, always returning the maximum value, without receiving any input.

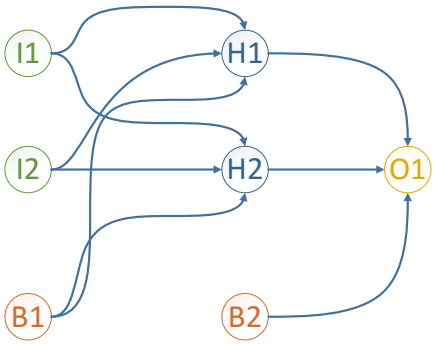


Fig 1. Artificial neural network (example)

From the point of view of classification, there are typically three broad categories into machine learning: 1) supervised learning. The algorithm is trained with example inputs and outputs. For example, Carneiro proposes a method for semantic image annotation and retrieval [18]; 2) unsupervised learning. The algorithm is not trained with examples but other techniques such as generic algorithms help to find the correct solution. For example, Kattan et al. predict the position of any particular target event in a time series [19]; and 3) reinforcement learning. The algorithm learns its behaviour based on feedback from the environment. For example, Gosavi uses reinforcement learning for control optimization [20]. There can be other definitions such as semi-supervised learning, learning to learn, developmental learning and even other classifications like for example depending on the expected kind of output.

ANNs are linked to a large amount of different type of scenarios. For example: 1) any kind of control system [21]; 2) autonomous navigation of robots [22]; 3) pattern recognition [23]; 4) forecasting [24]; or estimation of heating loads of buildings [25].

Those ANNs-based problems can be formulated using different technologies such as:

- Encog for Java, .NET or C++ [5], supporting different algorithms such as hidden Markov Models, vector machines, Bayesian networks and neural networks.
- AForge.NET for .NET [26], designed for developers and researchers in the fields of computer vision and artificial intelligence.
- The SHOGUN machine learning toolbox [27], with interfaces for MATLAB, R, Octave and Python, apart from a stand-alone command line interface.
- Apache Mahout [28], providing free implementations of scalable machine learning algorithms, using Java libraries for common operations.
- Many others such as Weka [6], Spark MLlib or ConvNetJS.

The main problem is that there are not bridges among the previous technologies, so users are highly dependent on the underlying technology. In addition, when programming is needed, software libraries that are provided does not have a high level of abstraction since they usually are designed for GPLs with no semantics in the language linked to the type of problems to be solved. That forces users to have both high programming and machine learning skills to make use of them.

III. Overview of the System

To design the prototype, we used the MDE development approach, raising the level of abstraction of software engineering. Specifically, we have used the tools built on the Eclipse Modeling Project (EMP) [29], offering one of the most accepted implementations of the standards promoted by the Object Management Group (OMG) [30].

OMG is the organization that has driven the development of MDE through the creation of the set of standards enclosed in the Model-Driven Architecture (MDA) [12] specification, for the effective and efficient work under the MDE paradigm. It is usually carried out by creating DSLs tied to a specific domain of knowledge and generating artifacts for different platforms and manufactures.

Decision in favor of a new DSL is usually not easy because it is much less expensive (both in time and cost) to adopt an existing DSL if available or even to use a GPL such as Java or C#. For that, there are mainly only two reasons why it is worth creating a new DSL [31]: 1) improved software economics, giving some authors as a reference point three developments [32] to obtain a positive return on investment; and 2) allow that people with less domain and programming expertise to develop software, even end-users with some domain, but no programming expertise [33][34].

Since our goals are compatible with both criteria, we have created a new DSL based on common elements used to describe the structural part of a feedforward neural network. To that end, we have used the Xtext framework [35], which allows the creation of both GPLs and DSLs in a relatively easy way [36]. From a grammar and some other definitions, it is possible, for example, to get a working parser and linker and also a complete Eclipse-based Integrated Development Environment [37]. Xtext also provides several mechanisms through which you can configure different aspects of languages such as validations of code, syntax highlighting, proposals to developers, code formatting or even generating artifacts through programs implemented with the programming languages defined with Xtext.

A. AiDSL

Next, there is a snippet of the context-free Xtext grammar used as the basis of the AiDSL language.

```
AI:
    neuralNetworks+=NeuralNetwork*;

NeuralNetwork:
    "neuralNetwork" name=ID "{"
        "neurons" "{"
            "input" "{"
                inputLayer = InputLayer
            "}"
            "hidden" "{"
                hiddenLayer = HiddenLayer
            "}"
            "output" "{"
                outputLayer = OutputLayer
            "}"
        "}"
        "training" "{"
            "input:" trainingInput = FloatsColection
            "output:" trainingOutput = FloatsColection
            ("type:" trainingType = TrainingType)?
("errorThreshold:" trainingErrorThreshold = Float)?
            ("result:" trainingResult = Result)?
        "}"
        "data" "{"
            "input:" dataInput = FloatsColection
            ("result:" dataResult = Result)?
        "}"
    "}"
;

InputLayer:
    "size:" size = INT
    bias ?= "bias"?
;

HiddenLayer:
    "size:" size = INT
    bias ?= "bias"?
    ("activation:" activation = Activation)?
;

OutputLayer:
    "size:" size = INT
    ("activation:" activation = Activation)?
;

enum Activation:
    BiPolar |
    Competitive |
    HyperbolicTangent |
    Linear |
    LOG |
    Sigmoid |
    SoftMax
;

enum TrainingType:
    Backpropagation |
    QuickPropagation = "QPROP"|
    LevenbergMarquardt = "LMA"|
    ManhattanUpdateRule |
    ResilentPropagation = "RPROG" |
    ScaledConjugateGradient = "SCG"
;

FloatsColection: ('[' Floats ']')+;
Floats: Float(','Float)*;
Float: INT('.'INT)?;

enum Result:
    Console |
    None
;
```

With this grammar, neural networks can be created indicating information about the input, the hidden and the output layers (e.g., number of neurons, presence of bias neurons and the activation mode). Activation functions are attached to layers and are needed to scale data output from a layer. There are different activation functions available (users can select among different functions depending on the case: bipolar, competitive, hyperbolic tangent, linear, log, sigmoid or softmax). Depending on the selection, network behavior will be different. As we focus on supervised learning, we define the way users can introduce training data with inputs and expected outputs and the algorithm used for training the system (Back propagation, Quick propagation, Levenberg Marquardt, Manhattan update rule, Resilent propagation or Scaled conjugated gradient), also depending on each particular problem. More information about the theoretical basis for designing neural networks could be found for example in Haykin [38].

The Xtext-based grammar is transformed internally into an ANTLR grammar [39] to implement the lexer (lexical analysis) and the parser (syntactic analysis) that is used when a programming language is being defined. In addition, it also generates all the necessary infrastructure to create the Abstract Syntax Tree (AST) to perform a semantic analysis on the language elements. The iteration through the tree is performed using model-based technologies, particularly the Eclipse Modeling Framework [40], which serves to ensure interoperability of the generated DSL with many other model-based existing tools such as the tools defined in the Eclipse Modeling Project [29] to help improve software development productivity.

The definition of such a grammar leads to a metamodel for the domain that is automatically generated. This metamodel makes programs that are made based on it to follow a formal definition that allows to use any tool that is compatible with all standards promoted by

the MDA such as interoperability, reusability and portability, opening a wide range of possibilities. Thus, every time a new program with AiDSL is created, a model that conforms to the proposed metamodel is instantiated, following its rules and offering a formalism that makes it very easy to perform different tasks such as validation, storing or generation of artifacts. The rules are defined in general terms based on the metamodel of the language, not for each individual case, that is, not for each model obtained during the development, which facilitates the process.

B. Transformations from AiDSL to other technologies

The code below shows a fragment of the template that is used to generate artifacts from any of the models defined using AiDSL, based on its grammar. In this example, programmed with the Xtend language, the generation is focused on the Encog machine learning framework [5] but with other templates, code for other platforms could be generated without further changes. In addition, it could be possible to directly interpret models without the need of focusing on any platform. The idea of this approach is to generate from a model, easily and automatically, the code for different architectures or platforms (it would only be necessary to add new templates). That would be a key step to benefit from all the advantages of integration and reuse offered by the MDE approach (e.g., the use of common repositories and version control systems for models).

```
@SuppressWarnings("unused")
public class «n.name.toFirstUpper» {
   public static double trainingInput[][] =
«n.trainingInput.floatsColection»;
   public static double trainingOutput[][] =
«n.trainingOutput.floatsColection»;
   public static double dataInput[][] =
«n.dataInput.floatsColection»;

   public void run() {
      BasicNetwork network = new BasicNetwork();
      network.addLayer(new BasicLayer(null, «IF
n.inputLayer.bias == true»true«ELSE»false«ENDIF»,
«n.inputLayer.size»));
      network.addLayer(new BasicLayer(new «n.hiddenLayer.
activation.toString.activation»(), «IF n.hiddenLayer.bias
== true»true«ELSE»false«ENDIF», «n.hiddenLayer.size»));
      network.addLayer(new BasicLayer(new «n.outputLayer.
activation.toString.activation»(), false, «n.outputLayer.
size»));
      network.getStructure().finalizeStructure();
      network.reset();

      MLDataSet trainingSet = new
BasicMLDataSet(trainingInput, trainingOutput);
      «trainingType(n.trainingType)»

      int epoch = 1;
      do {
         train.iteration();
         «IF n.trainingResult == Result.CONSOLE»
         System.out.println("Epoch #" + epoch + " Error:"
+ train.getError());
         «ENDIF»
         epoch++;
      } while(train.getError() >
«n.trainingErrorThreshold»);
      train.finishTraining();

      MLDataSet dataSet = new BasicMLDataSet(dataInput,
null);
      «IF n.dataResult == Result.CONSOLE»
      System.out.println("Neural Network Results:");
      «ENDIF»
      for(MLDataPair pair: dataSet ) {
         final MLData output = network.compute(pair.
getInput());
         «IF n.dataResult == Result.CONSOLE»
         System.out.println(pair.getInput() +
               " => actual=" + output);
         «ENDIF»
      }

      Encog.getInstance().shutdown();
   }
}
'''
```

C. AiIDE

Based on the Xtext architecture, some of the features included in the development environment called AiIDE are:

- Custom syntax-highlighting to distinguish the different elements of the language (e.g., keywords, comments or variables). This is done by implementing the Xtext interfaces IHighlightingConfiguration and ISemanticHighlightingCalculator.
- Content assistant to help the developer to write code faster and more efficiently through the use of the auto-complete functionality (extending the TerminalsProposalProvider class).
- Static validation of the language elements to detect syntactic and semantic issues (extending the AbstractDeclarativeValidator class).
- Suggestions for fixing errors or problems identified in the code (extending the DefaultQuickfixProvider class).
- Templates that allow developers to reduce the learning curve for typical operations.
- Formatting the code through a feature called code beautifier to distribute it properly and promote its maintenance (extending the AbstractDeclarativeFormatter class).
- Outline view fully configurable to both the elements that appear and text or icons attached to them (extending the DefaultObjectLabelProvider class).

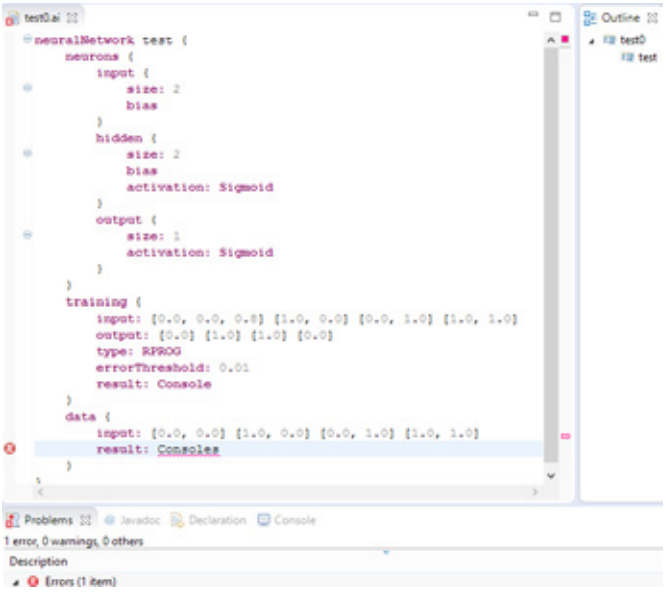


Fig. 2. AiDE working

Fig. 2 is a screenshot of the environment when a model is being created. It can be seen different features. For example, the syntax-highlighting for different elements (e.g., neurons, bias, training, etc.), the static validation marking a result type as not valid because

TABLE 1
GUIDELINES FOR A BETTER QUALITY AND A BETTER ACCEPTANCE AMONG ITS USERS

GUIDELINE	ACCOMPLISHMENT
Language purpose	
Identify language uses early	The language is used mainly for documentation of knowledge and code generation
Ask questions about uses	Any person with interest in defining neural networks will be able to model with the language
Make your language consistent [42]	It is consistent with the sole idea of defining and creating neural networks at this point
Language realization	
Decide carefully whether to use graphical or textual realization	Textual realization based on the advantages noted by Groenniger et al. [43]: 1) need of less space to display the same information; 2) more efficient creation of code; 3) easier integration with other languages; 4) more speed and quality of the formatting; 5) platform and tool independency; and 6) better version control support. Besides, graphical realizations provide a better overview and ease the understanding of models [41], but in a very close and specific domain like neural networks, we believe that the advantages of textual languages are more important
Compose existing languages where possible	Instead of starting from scratch, we have used the entire ecosystem of tools provided by the Eclipse Modeling Project, specifically Xtext, a DSL to define other DSLs, which relies heavily on the use of the Xtend language, an extension of the Java language, primarily intended to support the creation of DSLs (e.g., validations and code generation)
Reuse existing language definitions	To create AiDSL we have used the core grammar of Xtext as a basis to avoid redefining elements already defined previously
Reuse existing type systems	Related to the previous point, we have reused the core data types defined by the creators of Xtext with the aim of reusing the existing knowledge
Language content	
Reflect only the necessary domain concepts	It only contains the basic elements needed to generate code in the different target formats, so no extra domain concept is added
Keep it simple	With a small number of elements, simple syntax and reduced domain of knowledge, we think that the language is easier than other alternatives. The quantitative analysis also suggests the same idea
Avoid unnecessary generality	Due to the close domain of the language, we did not include the generalization concept, meeting with the principle of designing only what is necessary
Limit the number of language elements	The language is small, having only 13 domain-specific keywords (e.g., Java has 50 generic keywords and C# even more)
Avoid conceptual redundancy	Each fact can only be described in a unique way, avoiding redundancy
Avoid inefficient language elements	Each element is needed for clarity and used with the only purpose of allowing the generation of the final code, so there are no inefficient language elements
Concrete syntax	
Adopt existing notations domain experts use [44]	Neural networks are usually defined using a graph-based structure with inputs, outputs an intermediate nodes or states
Use descriptive notations	The language has a small number of keywords with syntax highlighting and code completion support. In addition, frequently-used symbols in other languages such as =, { or } maintain their semantics
Make elements distinguishable	Keywords, different syntax highlighting and an outline view are used to make elements distinguishable
Use syntactic sugar appropriately	We avoid syntactic sugar since we think that in a small DSL expressing the same concepts in different ways can be counterproductive, confusing users and hindering validation and code generation unnecessarily
Permit comments [45]	Support for common types of comments: single-line comments (//) and multi-line comments (/*..*/)
Provide organizational structures for models	Organizational structures such as packages are important for complex systems. However, to keep the language simple, we intend to have the definition of the set of neural networks in the same organizational structure
Balance compactness and comprehensibility	The quantitative analysis suggests that this approach may require less elements than other approaches. However, since it is a DSL with concrete semantics for the domain, it is even more comprehensible
Use the same style everywhere	All the elements of the language have the same look-and-feel and we do not embed any external language that can difficult the understanding of the language by using another syntax
Identify usage conventions	Based on an ANTLR grammar we define typical usage conventions including notation of identifiers, order of elements or type of comments
Abstract syntax	
Align abstract and concrete syntax	We took into account the three principles mentioned in Karsai et al. [41]: 1) elements that differ in the concrete syntax also have different abstract notations (e.g., input layer and type of learning are based on different metaclasses); 2) elements that have a similar meaning can be internally presented by reusing concepts of the abstract syntax (e.g., the FloatsColection rule for indicating inputs and outputs has been created using two int values along with other literals such as "[", "]" or "."); and 3) the abstract notation should not depend on the context an element is used but only on the element itself
Prefer layout which does not affect translation from concrete to abstract syntax	To simplify the usage of the DSL, the layout of the models does not affect the semantics. For example, modelers can use tabs, spaces or line breaks whenever they want. However AiIDE provides the feature called code beautifier, also provided by some environments to automatically place the language elements in a way easily understandable for most potential users
Enable modularity [46]	It is possible to decompose the code into smaller files, referencing them from other files. However, for this small language, we think that it is not necessary and it may unnecessarily increase the difficulty of use
Introduce interfaces	Interfaces are an important feature in complex systems, increasing flexibility and maintenance. However, we did not need them in our DSL because it is a simple declarative language

instead of Console, the programmer typed Consoles as the way to show the output, and the outline view showing a summary of the elements that are being used (in the example just a neural network inside the file).

In addition, it is possible to perform other customizations such as specifying the scope of the variables of the language. Thus, the AiIDE is a full-fledged development environment integrated in the Eclipse platform with the resulting advantages it provides (e.g., well-known and proven platform for developers, large amount of tools and plug-ins, open environment, etc.).

IV. Evaluation

The sections below are dedicated to a qualitative and quantitative study to show the characteristics of AiIDE and AiDSL, justifying the design and the need for its creation.

A. Qualitative analysis

To achieve a better quality of the language and the environment design and a better acceptance among its users, Karsai et al. [41] have proposed some guidelines largely based on their experience in developing languages as well as relying on existing guidelines on programming and modeling languages. Table 1 serves to verify that these guidelines are met.

B. Quantitative analysis

In this section we briefly evaluate the AiDSL language. We obtain a quantitative measurement that allows us to evaluate the main objective of our proposal; simplify and make more agile the definition of machine learning-based solutions.

In this first step of the development we are going to do a brief comparison between the definitions of two different neural networks using both AiDSL and the Encog framework. Since with AiDSL it is possible to automatically generate code for Encog and any other technology, if the syntax used by AiDSL is more compact, then it can clearly be seen as advantageous over other languages or frameworks. The measured aspects in the code and the structure are the ones below:

- Code lines: it refers to the number of lines of information needed to define the neural networks in each case.
- Words: number of words used.
- Characters: number of characters, spaces included.

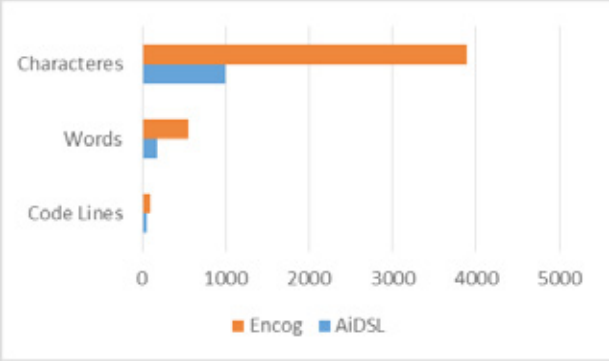


Fig. 3. Comparing AiDSL and Encog working with two neural networks

In the obtained results of the analysis (Fig. 3), we can observe that with AiDSL we require much less code lines (56 vs 102), words (180 vs 549) and characters (996 vs 3895) to define the same information than with the Encog framework. For the measurements, we defined two neural networks with a number of input, hidden and output neurons, activation method, training information and type of output expected. After we defined it with AiDSL the AiIDE automatically generated the code that should be necessary if we have worked directly with the Encog framework.

V. Conclusions and Future Work

In this paper we have presented the first version of a language for defining neural networks (AiDSL) and a development environment to facilitate working with those networks (AiIDE). This has been done by identifying basic elements that are useful to define the important aspect of any artificial neural network. In addition, we have defined mappings for transforming models made with AiDSL to the code that should be used if we worked with the Encog framework instead, and created the basis to do the same with other different popular frameworks (e.g., Weka), which favors the development and increases productivity and interoperability among systems. Finally, it the use of AiDSL through the AiIDE is easier than the manual and specific handling of other frameworks with identical purposes. Of course, both AiDSL and AiIDE are prototypes with limited scope and popular frameworks such as Encog or Weka offer many more features.

From the point of view of computer science, the focus of this paper could be set embedded in this category: Artificial Intelligence → Machine Learning → Neural Network → Feedforward Neural Network → Supervised learning. Further works will focus on other areas while they will delve into supervised learning.

Future work will be to improve and adapt both AiIDE and AiDSL with new frameworks and features to define neural networks. Finally, we will perform a usability study with real users for quantifying how simple, easy and intuitive is our proposal for them. The idea is to work with people with different profiles and ask them to define several neural networks using different techniques. That way, we will observe, among other things, the efficiency, the learning curve and the number of errors that are performed during the tasks.

References

[1] S. Russell, P. Norvig, and A. Intelligence, “A modern approach,” Artif. Intell. Prentice-Hall, Egnlewood Cliffs, vol. 25, 1995.

[2] D. Poole, A. Mackworth, and R. Goebel, Computational Intelligence: A Logical Approach. Oxford, UK: Oxford University Press, 1997.

[3] J. G. Carbonell, R. S. Michalski, and T. M. Mitchell, “An overview of machine learning,” in Machine learning, Springer, 1983, pp. 3–23.

[4] T. M. Mitchell, “Machine learning. WCB.” McGraw-Hill Boston, MA:, 1997.

[5] H. Jeff, “Programming Neural Networks with Encog3 in Java,” 2011.

[6] G. Holmes, A. Donkin, and I. H. Witten, “Weka: A machine learning workbench,” in Intelligent Information Systems, 1994. Proceedings of the 1994 Second Australian and New Zealand Conference on, 1994, pp. 357–361.

[7] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and others, “Scikit-learn: Machine learning in Python,” J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.

[8] A. Van Deursen, P. Klint, and J. Visser, “Domain-Specific Languages: An Annotated Bibliography.,” Sigplan Not., vol. 35, no. 6, pp. 26–36, 2000.

[9] B. Lantz, Machine learning with R. Packt Publishing Ltd, 2013.

[10] S. Kent, “Model driven engineering,” in Integrated Formal Methods, 2002, pp. 286–298.

[11] V. García-Díaz, J. Tolosa, B. G-Bustelo, E. Palacios-González, Ó. Sanjuan-Martínez, and R. Crespo, “TALISMAN MDE Framework: An Architecture for Intelligent Model-Driven Engineering,” in Distributed Computing Artificial Intelligence Bioinformatics Soft Computing and Ambient Assisted Living, vol. 5518, S. Omatu, M. Rocha, J. Bravo, F. Fernández, E. Corchado, A. Bustillo, and J. Corchado, Eds. Springer

Berlin / Heidelberg, 2009, pp. 299–306.
[12] S. J. Mellor, K. Scott, A. Uhl, and D. Weise, “Model-driven architecture,” in Advances in Object-Oriented Information Systems, Springer, 2002, pp. 290–297.
[13] E. P. González, H. F. Fernández, V. G. Díaz, B. C. P. G. Bustelo, J. M. C. Lovelle, and O. S. Martínez, “General purpose MDE tools,” IJIMAI, vol. 1, no. 1, pp. 72–75, 2008.
[14] G. C. Fernandez, J. P. Espada, V. G. Díaz, and M. G. Rodríguez, “Kuruma: the vehicle automatic data capture for urban computing collaborative systems,” Int. J. Interact. Multimed. Artif. Intell., vol. 2, no. 2, pp. 28–32, 2013.
[15] G. Cueva-Fernandez, J. P. Espada, V. García-Díaz, R. G. Crespo, and N. Garcia-Fernandez, “Fuzzy system to adapt web voice interfaces dynamically in a vehicle sensor tracking application definition,” Soft Comput., pp. 1–14, 2015.
[16] V. García-Díaz, B. C. P. G-Bustelo, O. Sanjuán-Martínez, E. R. N. Valdez, and J. M. C. Lovelle, “MCTest: towards an improvement of match algorithms for models,” IET Softw., vol. 6, no. 2, p. 127, Apr. 2012.
[17] B. Yegnanarayana, Artificial neural networks. PHI Learning Pvt. Ltd., 2009.
[18] G. Carneiro, A. B. Chan, P. J. Moreno, and N. Vasconcelos, “Supervised learning of semantic classes for image annotation and retrieval,” Pattern Anal. Mach. Intell. IEEE Trans., vol. 29, no. 3, pp. 394–410, 2007.
[19] A. Kattan, S. Fatima, and M. Arif, “Time-series event-based prediction: An unsupervised learning framework based on genetic programming,” Inf. Sci. (Ny)., 2015.
[20] A. Gosavi, “Control Optimization with Reinforcement Learning,” in Simulation-Based Optimization, Springer, 2015, pp. 197–268.
[21] W. T. Miller, P. J. Werbos, and R. S. Sutton, Neural networks for control. MIT press, 1995.
[22] D. A. Pomerleau, “Efficient training of artificial neural networks for autonomous navigation,” Neural Comput., vol. 3, no. 1, pp. 88–97, 1991.
[23] C. G. Looney, Pattern recognition using neural networks: theory and algorithms for engineers and scientists. Oxford University Press, Inc., 1997.
[24] G. Zhang, B. E. Patuwo, and M. Y. Hu, “Forecasting with artificial neural networks:: The state of the art,” Int. J. Forecast., vol. 14, no. 1, pp. 35–62, 1998.
[25] S. A. Kalogirou, “Artificial neural networks in renewable energy systems applications: a review,” Renew. Sustain. energy Rev., vol. 5, no. 4, pp. 373–401, 2001.
[26] [26] A. Kirillov, “AForge .NET framework,” 2010-03-02)[2010-12-20]. http://www. aforgenet. com. 2013.
[27] [27] S. Sonnenburg, G. Rätsch, S. Henschel, C. Widmer, J. Behr, A. Zien, F. de Bona, A. Binder, C. Gehl, and V. Franc, “The SHOGUN machine learning toolbox,” J. Mach. Learn. Res., vol. 11, pp. 1799–1802, 2010.
[28] [28] S. Owen, R. Anil, T. Dunning, and E. Friedman, Mahout in action. Manning, 2011.
[29] R. C. Gronback, Eclipse modeling project: a domain-specific language (DSL) toolkit. Pearson Education, 2009.
[30] O. M. G. CORBA and I. Specification, “Object Management Group.” Joint revised submission OMG document orbos/99-02-, 1999.
[31] M. Mernik, J. Heering, and A. M. Sloane, “When and How to Develop Domain-specific Languages,” ACM Comput. Surv., vol. 37, no. 4, pp. 316–344, 2005.
[32] K. Schmid and M. Verlage, “The economic impact of product line adoption and evolution,” IEEE Softw., vol. 19, no. 4, pp. 50–57, 2002.
[33] B. A. Nardi, A small matter of programming: perspectives on end user computing. MIT press, 1993.
[34] M. Voelter, “A Catalog of Patterns for Program Generation.,” in EuroPLoP, 2003, pp. 285–320.
[35] S. Efftinge and M. Völter, “oAW xText: A framework for textual DSLs,” in Workshop on Modeling Symposium at Eclipse Summit, 2006, vol. 32, p. 118.
[36] M. Eysholdt and H. Behrens, “Xtext: implement your language faster than the quick and dirty way,” in Proceedings of the ACM international conference companion on Object oriented programming systems languages and applications companion, 2010, pp. 307–309.
[37] J. desRivieres and J. Wiegand, “Eclipse: A platform for integrating development tools,” IBM Syst. J., vol. 43, no. 2, pp. 371–383, 2004.
[38] S. Haykin and N. Network, “A comprehensive foundation,” Neural Networks, vol. 2, no. 2004, 2004.
[39] T. J. Parr and R. W. Quong, “ANTLR: A predicated-LL (k) parser generator,” Softw. Pract. Exp., vol. 25, no. 7, pp. 789–810, 1995.
[40] D. Steinberg, F. Budinsky, E. Merks, and M. Paternostro, EMF: eclipse modeling framework. Pearson Education, 2008.
[41] G. Karsai, H. Krahn, C. Pinkernell, B. Rumpe, M. Schindler, and S. Völkel, “Design guidelines for domain specific languages,” arXiv Prepr. arXiv1409.2378, 2014.
[42] B. Meyer, “Eiffel-The Language Prentice Hall,” Englewood Cliffs, NJ, 1992.
[43] H. Grönninger, H. Krahn, B. Rumpe, M. Schindler, and S. Völkel, “Textbased modeling,” arXiv Prepr. arXiv1409.6623, 2014.
[44] D. Wile, “Lessons learned from real DSL experiments,” in System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on, 2003, p. 10–pp.
[45] R. S. Scowen and B. A. Wichmann, “The definition of comments in programming languages,” Softw. Pract. Exp., vol. 4, no. 2, pp. 181–188, 1974.
[46] S. Wong, Y. Cai, M. Kim, and M. Dalton, “Detecting software modularity violations,” in Proceedings of the 33rd International Conference on Software Engineering, 2011, pp. 411–420.

Vicente García-Díaz is an Associate Professor in the Computer Science Department of the University of Oviedo. He has a PhD from the University of Oviedo in computer engineering. His research interests include Domain-Specific Languages, Model-Driven Engineering, Business Process Management, Machine Learning, Internet of Things and eLearning.
Contact address is Computer Science Department, University of Oviedo. Edificio de la Facultad de Ciencias. C/ Calvo Sotelo s/n. 33007 Oviedo (Asturias, España); e-mail: garciavicente@uniovi.es

Jordán Pascual Espada is a research scientist at Computer Science Department of the University of Oviedo. Ph.D. from the University of Oviedo in Computer Engineering B.Sc. in Computer Science Engineering and a M.Sc. in Web. He has published several articles in international journals and conferences, he has worked in several national research projects. His research interests include the Internet of Things, exploration of new applications and associated human computer interaction issues in ubiquitous computing and emerging technologies, particularly mobile and Web applications.
Contact address is Computer Science Department, University of Oviedo. Edificio de la Facultad de Ciencias. C/ Calvo Sotelo s/n. 33007 Oviedo (Asturias, España); e-mail: pascualjordan@uniovi.es

Cristina Pelayo García-Bustelo is a lecturer in the Computer Science Department of the University of Oviedo. She has a PhD from the University of Oviedo in computer engineering. Her research interests include object-oriented technology, Web engineering, eGovernment, modeling software with BPM, DSL and MDA.
Contact address is Computer Science Department, University of Oviedo Edificio de la Facultad de Ciencias. C/ Calvo Sotelo s/n. 33007 Oviedo (Asturias, España); e-mail: crispelayo@uniovi.es

Juan Manuel Cueva Lovelle became a mining engineer from Oviedo Mining Engineers Technical School in 1983 (Oviedo University, Spain). He has a PhD from Madrid Polytechnic University, Spain (1990). From 1985 he has been a professor at the languages and computers systems area in Oviedo University (Spain), and is an ACM and IEEE voting member. His research interests include object-oriented technology, language processors, human-computer interface, Web engineering, modeling software with BPM, DSL and MDA.
Contact address is Computer Science Department, University of Oviedo. Edificio de la Facultad de Ciencias. C/ Calvo Sotelo s/n. 33007 Oviedo (Asturias, España); e-mail: cueva(at)uniovi.es

A New Approach for Hiding Image Based on the Signature of Coefficients

Tawfiq A. Al-asadi, Israa Hadi Ali, Abdul kadhem Abdul kareem Abdul kadhem

Information Technology College, University of Babylon, Iraq

***Abstract* — This paper presents a new approach for hiding the secret image inside another image file, depending on the signature of coefficients. The proposed system consists of two general stages. The first one is the hiding stage which consist of the following steps (Read the cover image and message image, Block collections using the chain code and similarity measure, Apply DCT Transform, Signature of coefficients, Hiding algorithm , Save information of block in boundary, Reconstruct block to stego image and checking process). The second stage is extraction stage which consist of the following steps (read the stego image, Extract information of block from boundary, Block collection, Apply DCT transform, Extract bits of message and save it to buffer, Extracting message).**

***Keywords* — Information Hiding, Steganography, Chain Code, DCT, Signature of Coefficients.**

I. Introduction

Communication in our world are today based on the use of wide world website (the Internet) where anyone can send anything to any place on earth, every day trillions of messages and information are sent and received in seconds. Encrypted data are always suspicious, and considered illegal in some places in the world. Information hiding has become very important as it works on converting information to be hidden and hard to discover [1].

Information hiding is a general term that contains many other topics (steganography, watermarking, copyrighting). This paper concentrate on Steganography, which is a field of science that works on hidden communication that hides a message in a way that no one knows about it only the sender and the intended recipient, where there is no attention for the hidden information and cannot be attacked easily . It is different from cryptography where information is unreadable but not invisible [2].

Steganography technique hides important information that should be secret in normal media (audio, digital image, video, and so on). Any attempt to extract hidden information from stego is called Steganalysis. The Steganographic algorithm is considered to be broken if a steganalytic algorithm detect a given media to be a carrier for a secret message [3].

The secret message can be hidden inside the cover image in some locations according to a particular algorithm. This paper will introduce a new technique for information hiding based on chain code and DCT. The chain code is used to determine some locations of cover image for hiding a message by developing the traditional freeman chain code with 8-connectivity to quad chain code , the quad chain code is found depending on one of the measures which is the similarity between two adjacent vectors (quad pixels). While the DCT is used to determine the locations of a block for hiding operation and use it as a signature to support the hide operation.

II. Related Works

Zuheir in [4] proposed a steganography method to hide text in the cover image by using traditional chain code, first generate a chain code and store it in the cover image then store the embedded text in the cover image according to the generated chain code. The system uses the first pixel of the cover image to specify the location of the starting point to begin with it. The second pixel contains the length of the secret message where each character needs 8 bits for representation. The system divided the image into two sections the first section contains the chain code which represent the map of the secret message and the second section include the secret message which the sender pass.

Our proposed system implements a new hiding technique by using similarity measure (cosine similarity) to generate chain code, each value of chain code represent the movement between two vectors (two neighbor quad pixels) in the cover image. The chain code was used to collect blocks, each block with size 8*8 for hiding operation.

Ajit and et al [5] proposed a novel image steganography technique by embedding a bit that is randomized. First operation is obtaining the DCT of the cover image, and then constructing the stego image by hiding the secret message that was given in least significant bit of the cover image in random locations depending on the threshold. The locations of the randomized pixel are determined by DCT coefficients for hiding.

Authors in [6] proposed an LSB and DCT steganographic technique for data hiding that applies spatial domain with frequency domain of steganography methods and asymmetric key cryptography. It's done by utilizing a significant bit of low frequency DCT coefficients of the cover image blocks hiding encrypted message bits. 2D-DCT converts the image block from the spatial domain to the frequency domain, and then bits of data are embedded by changing LSB of DCT coefficients.

Our proposed system implements the information hiding technique in the spatial domain. The hiding operation is not sequential but randomly based on the chain code and the nature of data that are dealing with it. The chain code technique is used to equipment pixels or quads (each quad = 4 pixels) of the cover image that are more similar with each other as blocks. The system utilizes the DCT to find the coefficients values that are zero or near to zero (between 1,-1) that were used to determine the signature of coefficients. That means, if the block elements are more similar with each other, is that will give high zero values after applying the 2D-DCT. According to the simulated results (MSE and PSNR), it is so difficult differing the genuine image from the stego image of our proposed algorithm, as they seem to be identical.

III. Background

A. Steganography.

Fundamentally information hiding, popularly known as steganography, is the method of concealing a message that is secret

DOI: 10.9781/ijimai.2015.352

in another message or carrier called “cover media”. The use of cover media facilitates, secure information transfer over commonly available public domain open channels that are insecure networks like the internet and wireless mobile networks. Although the evolution of these broadband networks has permitted cost-effective and high speed multimedia information transfer, but it has posed increasing threats to the information security due to greater possibilities of unauthorized information access. The issue of securing information becomesmore challenging in internet environment because it is open across the globe and being extensively used for information access and dissemination [7].

Steganography is a field of science for hiding information by embedding a message in other information. It is derived from the Greek words “STEGOS” which means ‘ cover ’ and “GRAFIA” which means ‘ writing ’ to be ‘ covered writing ’. Steganography is done by replacing bits of unused data in regular computer files (like sounds, graphics, HTML, texts, or image) with bits of invisible information. The information that is hidden can be any kind of text (plain text, cipher text), sound or images [8].

Information hiding in digital steganography is achieved by hiding the secret data with other seemingly innocuous cover object or carrier where message data is embedded under the cover using a key to generate the stego object. The object that is covered refers to the used object as the carrier to hide the messages and the stego object is one which contains the secret message [9]. Fig (1) explains steganography principle.

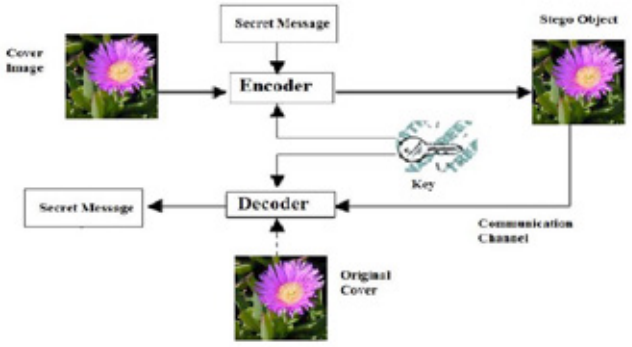


Fig. 1. Steganography principles.

General model to hide data in another data can be described as follows [10]:

1. Cover file (carrier): is used for hiding information and the size of the file is carefully selected to be enough for embedding the information that have to be secret.
2. Message file (secret file): contains secret information that has to be hidden and must be kept save during transformation.
3. Steganography algorithm: refers to the deterministic sequence for using a cover file to hide a secret message.
4. Stego file: is what follows embedding that embeds the secret message in the cover file using steganography algorithm.
5. Retrieve algorithm: is used for extracting the message that is secret from the stego file, the retrieved algorithm runs in a reversing way than the embedding algorithm.
6. The cover file and the message file may take many different files like audio files, text, image, and video files (e.g. .doc, .html, .wma, .jpeg, .tiff, .bmp, .gif, .wav, .mp3, and mp4, etc.).

B. Chain code.

Freeman Chain Code (FCC) was the first technique to represent an image that uses chain code; it was introduced by Freeman in 1961. Straight-line segments that are connected in sequence with particular length and direction are represented as a boundary by using this chain code. This representation is based on 4- or 8-connectivity of the segments. A numbering scheme is used to code the direction of each segment. 4-connected Freeman Chain Code is shown in fig (2-a) while fig (2-b) shows 8-connected Freeman Chain Code of 8-directional (FCCE) [11].

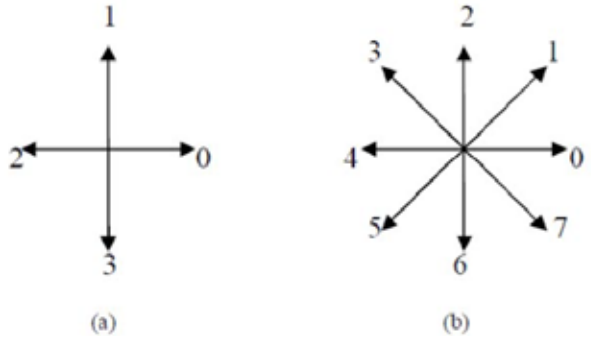


Fig. 2. Neighbor directions of freeman chain code.

In the Freeman Chain Code 8-connected (FCCE), 8 directions from one pixel to a neighbor pixel are possible, every code is considered as an angular direction, multiplied by 450 moving from one contour pixel to the next. Direction 0 means move “to the right of”, 2 means “immediately above”, and 1 is at 45 degrees, bisecting 0 and 2, and so on. Fig (3) shows the examples using the 8-connected path and 4-connected path of Freeman Chain Code.

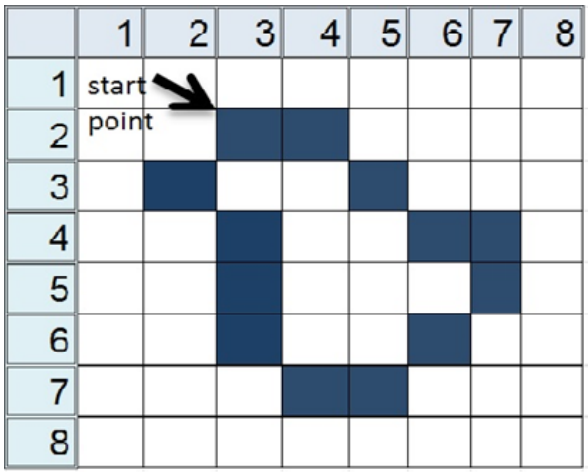


Fig. 3. Example of chain code

A. 4-connectivity chain code = 00303032323221112110

B. 8-connectivity chain code = 0770655432231

C. Discreet Cosine Transform (DCT)

The DCT is used widely in transformation for data compression (loosy). It’s an orthogonal transform, with a fixed size of (image independent) basis functions, properties of an excellent energy compaction and correlation reduction, and an algorithm that is efficient for computation. Ahmed et al found that the Karhunen Lòeve Transform (KLT) basis function of a first order Markov image resembles those of the DCT closely. As the correlation between the adjacent pixels approaches to one, they become identical [12].

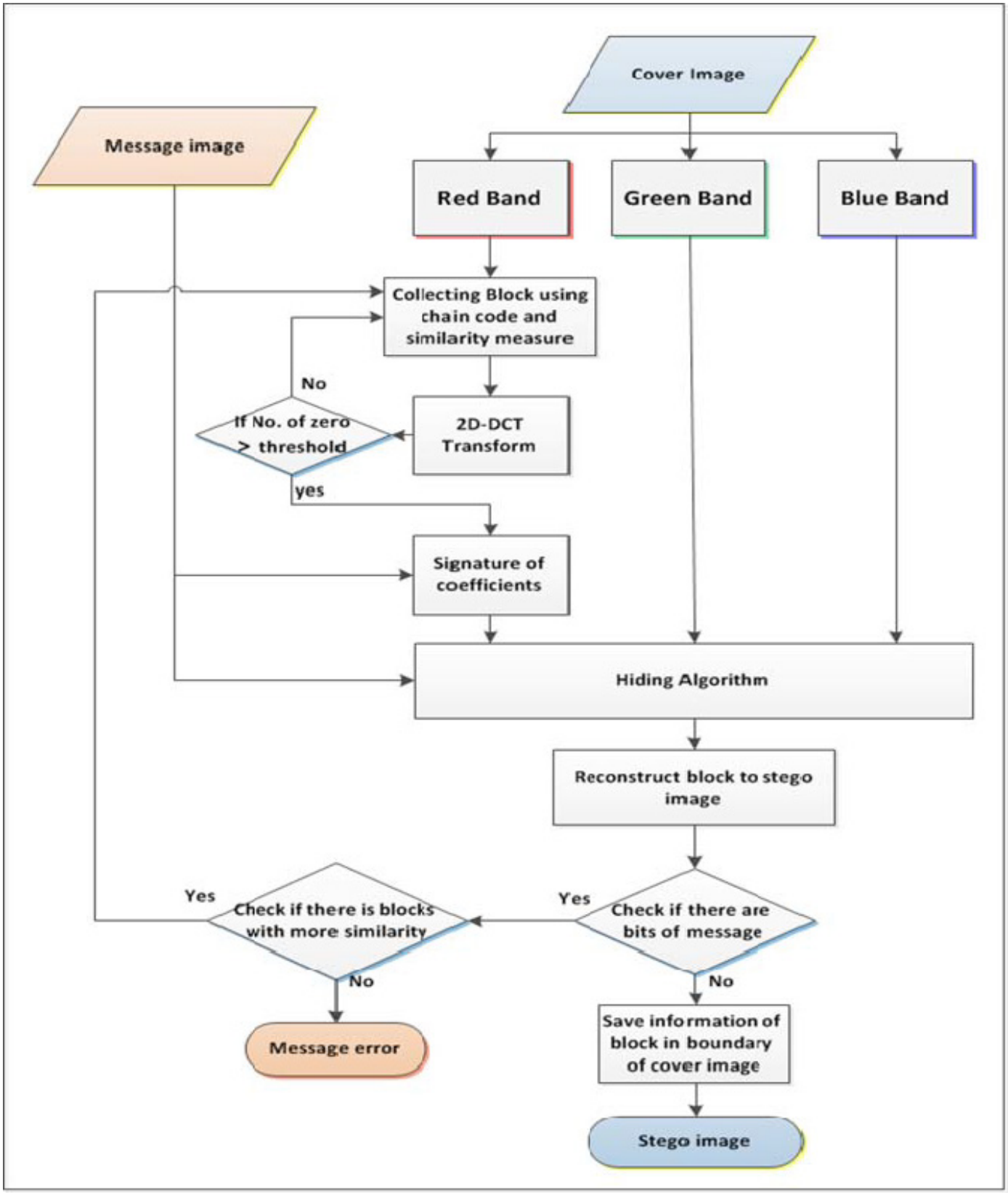


Fig. 4. Flow chart of the proposed system

The input data points are represented by DCT as the sum of cosine functions that are oscillating at different dimension and frequencies. There are generally two kinds of DCT: one dimensional (1-D) DCT and two dimensional (2-D) DCT. 2-D DCT is considered for this study work. For an input sequence, the 2-D DCT can be described as follows: [13].

$$F(u,v)=\frac{1}{4}C(u)C(v)\sum_{x=0}^{7}\sum_{y=0}^{7}f(x,y)\cos\left[\frac{\pi(2x+1)u}{16}\right]\cos\left[\frac{\pi(2y+1)v}{16}\right]$$

for $u=0,\ldots,7$ and $v=0,\ldots,7$

where $C(k)=\begin{cases}1/\sqrt{2} & \text{for } k=0\\ 1 & \text{otherwise}\end{cases}$

As effectively as possible, the DCT is protected against blocking artifact with no blocks that are interconnected since all DCT basis functions have a zero gradient at the edges of their blocks. In another word, only the DC level affects the blocking artifact and then can be targeted. In DCT operation ringing is a main problem. To make the image shaper DCT depends on the high frequency components, when edges happen in an image.

Though the components with the high frequency that continued firmly across the whole block are effective at improving the edge quality, they 'ring' in the flat areas of the block [13].

IV. The Proposed System

The proposed method consists of two general stages (the message hiding stage, the extracting message stage). Ech one consists of many steps.

A. The message hiding stage

In this stage there are many steps as illustrated in Fig (4) of hiding a gray scale image (8 bit per pixel) in the cover image (color image – 24 bit per pixel) depending on the signature of coefficients , it consists of the following steps :

1) Read the cover image and message image

Images are simply 2D arrays of colors where each color is represented using one of the color formats (8, 16, 24, or 32 bits).This step will choose an image with size M*N to be a cover which is a color BMP image (24 bit per pixel). This image will be converted to three bands (Red, green, blue) RGB each band is a 2D array with size M*N too. While the secret message is a BMP gray scale image (8 bit per pixel) that we need to hide inside the cover image and convert it to a stream of bits.

2) Collecting block based on chain code and similarity measure

In this step, we will collect a block with size 8*8 which pixels more similar with each other. The proposed system will we use the chain code to choose the similar pixels by developing the freeman chain code to quad chain code using quadruple pixel (each quad as a vector of 4 pixels), this step is find the chain code for the red band of the cover image depending on one of the similarity measures that were used to find the two adjacent vectors (four pixels or quad) that are almost similar.

The similarity measures can be applied in the system to find vectors (quad of pixels) that are more alike (cosine similarity) as illustrated in the following equation [14]:

$$Cosine(A,B)=\frac{A\cdot B}{\|A\|\cdot\|B\|}$$

For example, if A and B as the vector, when A vector =(1,3,7,4) and B vector =(5,3,1,6). Where:

$\|A\| = (1*1 + 3*3 + 7*7 + 4*4) = 75.$

$\|B\| = (5*5 + 3*3 + 1*1 + 6*6) = 71.$

$A\cdot B = (1*5 + 3*3 + 7*1 + 4*6) = 45.$

Cosine (A,B) = 0.61669.

As long as similarity measure return values in the range [0,1],when the similarity value is closer to 1 that mean blocks are more alike and vice versa.

Quad chain code that was found by scanning the pixels of the red band of cover image checks if there are (16) quads similar with each other to collect a block with size 8*8 (16 quad pixels equal 64 pixel).

Fig (5) explains how to collect a block from the cover image depending on chain code and similarity measure.

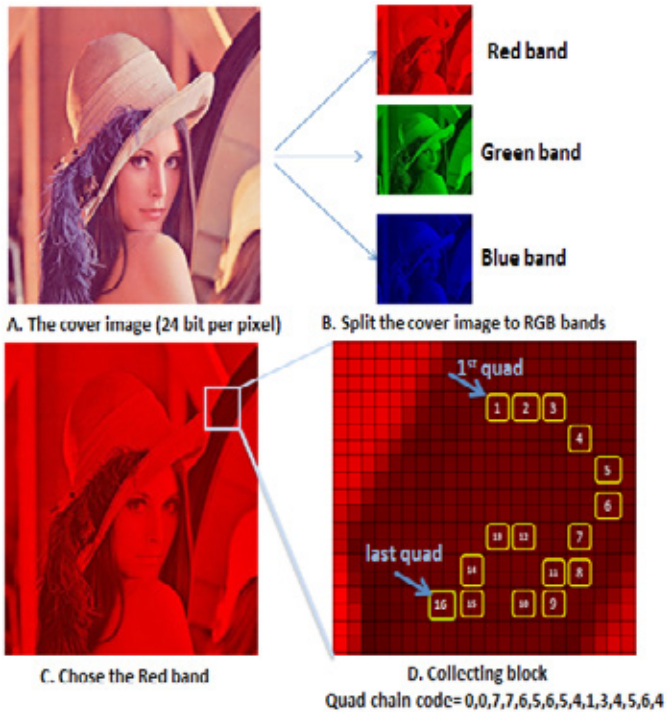


Fig. 5. Collecting block with size 8*8 pixels

3) Apply DCT Transform for a block

This step will apply the 2D-DCT transform for the block that was equipmented from the above step to convert values of pixels from the spatial domain to frequency domain. As shown in fig (6) which explain this step.

118	118	118	119	119	119	118	115
118	119	118	118	118	118	120	118
119	118	120	119	119	118	118	117
119	120	120	119	118	118	118	117
117	118	118	118	119	120	119	115
118	118	118	118	118	120	118	118
118	117	119	119	118	117	119	118
119	117	120	118	119	116	118	118

947	-0.4	-1.2	0.5	-0.5	-0.9	0.2	-1
1	0	0	-0.7	-0.6	0.6	-0.7	1.2
-0.8	0.3	0.5	0.8	1.6	-0.6	-0.3	-1.2
-1	3	0.1	-1.3	-0.3	1	0.4	2.2
0.5	-0.1	-0.1	0.3	0.5	1	0.3	-0.1
0.2	-1.7	-0.6	0.8	0.5	-0.7	0.3	0.3
0	-0.7	-0.7	-0.6	0.7	0.5	1.6	0
0	1.5	-0.8	0.1	0.8	-0.2	0.4	-0.5

A. Collect a block B. After apply 2D-DCT transforms

Fig. 6. Apply 2D DCT transform for a block

The next operation of this step is computing the number of zero or zeros values (between 1,-1) in coefficients of DCT array and compare it with threshold, if the number of zero is larger than threshold then go to the next step, else must go to the above step and collect another block for hiding operation.

4) Signature of coefficients

This step is checking the number of bits (message) that can be embedded inside the block that was collected from above. The first operation is extracting the mask matrix with size 8*8 by determining which locations of DCT coefficients have a zero value or values between (1,-1). That means if each element of DCT coefficients is between (1,-1), then the corresponding value in mask matrix will be (1), otherwise the remaining values will be (0). Except the values of the first row and the first column of mask matrix that will be the (0) value.

The next operation is computing the number of (1) values in the mask matrix to determine three parameters (low, mid, high). The values of these parameters are computed as follow:

Low =0.

High= number of (1) values in the mask matrix.

Mid= (Low + High) / 2.

The mid value is used to determine how many hiding operations are there for the block. The bits of the message must be hidden in locations of corresponding spatial domain block that mask matrix is (1) by using least significant bit (LSB) algorithm .the strategy that is going to be applied is bottom- up. For each hide operation to a number of the block positions, we will find the 2D-DCT to the block after hiding. Then we create the signature coefficients matrix according to the values of the DCT block (which means each element of DCT coefficients between (1,-1) the corresponding value in signature coefficients matrix will be (1), otherwise the remaining values will be (0). Except the values of the first row and the first column, the signature coefficients matrix will be (0) value).

The next operation is comparing the signature coefficients matrix with the mask matrix that has been extracted in the beginning to update the values of (low, mid, high). This operation will be benefited from the idea of binary search algorithm (divide and conquer) to update the (low, mid, high) if the signature coefficients matrix are equal to the mask matrix, that means increase bits of the message to hide in the block, otherwise must decrease bits of message that are hidden in the block. The new values of (low, mid, high) are explained in the following paragraph:

If signature coefficients matrix equal to the mask matrix then

Low = Mid.

High=High.

Mid = (Low + High) /2.

If signature coefficients matrix not equal to the mask matrix then

Low= Low

High= Mid

Mid= (Low+ High) /2.

In the next level, We will process a new hide by embedding bits of message in the block based on the value of (mid) which were extracted from the first level, also we apply 2D-DCT for the block after hiding, and extract the new signature of coefficients matrix from the DCT block, comparing between the signature of coefficients matrix and the mask matrix and the values (low, mid, high) that were updated based on the comparison result.

The successive levels will continue until all values (low, mid, high) are equal with each other. The output of this step is (mid value) that represent the number of bits (message) that can be embedded in the block without changing the signature of coefficients matrix compared with the mask matrix. The aim of this step is to ensure the integrity of the message without any loss in the extraction message from the stego-image stage.

To explain this step, consider the following message "**011010000 110010101101100011011000101101111**" that was needed to hide in the following cover block with size 8*8 as illustrated in figure (7).

119	118	119	119	119	118	118	118
118	119	118	118	118	118	119	118
119	118	118	119	118	118	118	119
119	118	118	118	119	118	119	119
119	119	118	119	118	118	118	119
118	118	118	118	118	118	118	118
118	119	118	118	119	119	118	118
118	118	118	119	118	119	119	118

Fig. 7. An example of the cover block.

The first operation is finding the mask matrix that was extracted depending on the DCT block (applying the 2-D DCT transform for the cover block and find the mask matrix from DCT transform block), and compute the parameters (low, mid and high) according to the mask matrix, figure (8) shows this operation.

946.9	0.11	0.2	0.22	0.52	0.21	-0.49	-0.29
0.32	0.58	0.39	-0.47	0.57	0.33	-0.02	0.66
-0.03	0.07	-1.22	0.29	-0.9	0.28	0.04	0
-0.34	0.81	-0.39	-0.17	-0.04	-0.52	0.17	0.27
0.88	0.01	-0.07	-0.06	0.63	0.63	0.16	0.2
0.38	-0.02	-0.42	0.51	0.57	-0.51	0.51	0.61
-0.39	0.06	-0.21	-0.39	0.78	0.88	0.72	-0.15
0.57	0.69	-0.06	-0.33	0.64	-0.22	0.33	-0.9

A. DCT transform block

0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1
0	1	0	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1

B. mask matrix

Fig. 8. Find the mask matrix.

The low value is (0) in the beginning, the high value is (48) that represent number of (1) values in the mask matrix while the mid value is (24).

Depending on the mid value, the first level is hide (24) bit of message "**011010000110010101101100**" in the cover block from down to up by using least significant bit (LSB), the next operation is applying 2D-DCT transform for a block after hiding, the next operation is finding the signature of coefficients matrix after the first hiding operation. Fig (9) explains the first level for the signature of coefficients.

119	118	119	119	119	118	118	118
118	119	118	118	118	118	119	118
119	118	118	119	118	118	118	119
119	118	118	118	119	118	119	119
119	119	118	119	118	118	118	119
118	119	118	119	119	118	119	118
118	119	118	118	119	119	118	118
118	118	118	119	118	119	119	118

A. The Cover block after level (1)

947.37	0.11	0.01	0.22	0.63	0.21	-0.95	-0.29
-0.07	0.58	0.54	-0.47	0.67	0.33	0.34	0.66
-0.3	0.07	-1.12	0.29	-0.9	0.28	0.29	0
0.35	0.81	-0.65	-0.17	-0.04	-0.52	-0.47	0.27
0.38	0.01	0.12	-0.06	0.63	0.63	0.63	0.2
0.24	-0.02	-0.37	0.51	0.57	-0.51	0.63	0.81
0.26	0.06	-0.46	-0.39	0.78	0.88	0.12	-0.15
-0.01	0.69	0.16	-0.33	0.64	-0.22	0.87	-0.9

0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1
0	1	0	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1

B. Apply DCT transform for the block C. signature matrix

Fig. 9. An example of the first level for the signature of coefficients.

The last operation of the first level is finding the comparison between the mask matrix and the signature matrix to extract the new values of (low, mid, high). We note that the mask matrix is equal to the signature matrix that means increase the bits of message hidden in the cover block. The new low value is (24), the high value is (48) without any change, while the mid value is ((24 + 48) / 2= 36).

The second level is hide (36) bit of message "**011010000110010101101100011011000110**" in the cover block from down to up by using least significant bit (LSB), the next operation is applying 2D-DCT transform for a block after the second hiding, the next operation is finding the signature of coefficients matrix for the cover block. We notice that the mask matrix is equal to the signature matrix, which means, decreasing the bits of message hidden in the next level, Fig (10) explains the second level for the signature of coefficients.

119	118	119	119	119	118	118	118
118	119	118	118	118	118	119	118
119	118	118	119	118	118	118	118
119	119	119	118	118	118	119	119
119	118	119	119	118	118	118	119
118	119	118	119	119	118	119	118
118	119	118	118	119	119	118	118
118	118	118	119	118	119	119	118

A. Cover block after level (2)

947.4	0.51	-0.12	-0.08	0.12	0.52	-0.63	-0.13
-0.14	0.81	0.49	-0.4	0.47	0.36	0.21	0.58
-0.39	-0.33	-1.07	0.79	-0.34	-0.06	-0.19	-0.17
0.43	0.31	-0.66	-0.24	0.42	-0.5	-0.17	0.51
0.63	0.07	0.32	-0.65	0.38	0.74	1.09	0.29
0.35	0.42	-0.06	0.35	0.1	-0.72	0.31	0.38
0.03	0.16	-0.69	0.04	0.82	0.89	-0.18	-0.17
-0.33	0.44	-0.44	0.07	1.02	0.17	1.15	-0.35

0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1
0	1	0	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	0	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	0	1	0	1

B. Apply DCT transform for the block C. signature matrix

Fig. 10. An example of the second level for the signature of coefficients.

The new low value is (24) without any change, the high value is (36), while the mid value is ((24 + 36) / 2= 30).

The third level is hiding (30) bits of message "**011010000110010101101100011011**" in the cover block from down to up by using least significant bit (LSB), the next operation is applying 2D-DCT transform for a block after the third hiding, next operation is finding the signature of coefficients matrix for the cover block. We notice that the mask matrix is not equal to the signature matrix, that means, decrease the bits of message hidden in the next level. Fig (11) explains the third level for the signature of coefficients.

119	118	119	119	119	118	118	118
118	119	118	118	118	118	119	118
119	118	118	119	118	118	118	119
119	118	118	118	119	118	119	119
119	118	119	119	118	118	118	119
118	119	118	119	199	118	119	118
118	119	118	118	119	119	118	118
118	118	118	119	118	119	119	118

A. The Cover block after level (3)

947.4	0.06	-0.12	0.08	0.63	0.42	-0.63	-0.04
-0.07	0.59	0.57	-0.43	0.67	0.28	0.25	0.59
-0.3	0.13	-0.94	0.47	-0.9	0	-0.14	-0.32
0.35	0.77	-0.76	-0.28	-0.04	-0.36	-0.21	0.46
0.38	-0.04	-0.01	-0.19	0.63	0.83	0.95	0.45
0.24	0.04	-0.21	0.67	0.57	-0.75	0.25	0.52
0.26	0.09	-0.39	-0.32	0.78	0.77	-0.06	-0.29
-0.01	0.62	-0.02	-0.53	0.64	0.07	1.32	-0.56

0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	0	1

B. Apply DCT transform for the block C. signature mask

Fig. 11. An example of the third level for the signature of coefficients.

The new low value is (24) without any change, the high value is (30) , while the mid value is ((24 + 30) / 2= 27).

The fourth level is hiding (27) bits of message "**011010000110010101101100011**" in the cover block from down to up by using least significant bit (LSB), the next operation is applying 2D-DCT transform for a block after the fourth hiding, next operation is finding the signature of coefficients matrix for the cover block. We notice that the mask matrix is not equal to the signature matrix, that means, decrease the bits of the message hidden in the next level. Fig (12) explains the fourth level for the signature of coefficients.

119	118	119	119	119	118	118	118
118	119	118	118	118	118	119	118
119	118	118	119	118	118	118	119
119	118	118	118	119	118	119	119
119	119	119	119	118	118	118	119
118	119	118	119	119	118	119	118
118	119	118	118	119	119	118	118
118	118	118	119	118	119	119	118

A. The Cover block after level (4)

947.5	0.21	-0.06	0.05	0.5	0.25	-0.79	-0.11
-0.11	0.55	0.55	-0.42	0.7	0.33	0.29	0.62
-0.46	-0.06	-1.03	0.52	-0.73	0.23	0.07	-0.19
0.45	0.89	-0.71	-0.3	-0.14	-0.5	-0.34	0.38
0.5	0.11	0.06	-0.23	0.5	0.66	0.79	0.35
0.09	-0.13	-0.29	0.71	0.72	-0.55	0.44	0.64
0.19	0.01	-0.43	-0.3	0.84	0.86	0.03	-0.23
0.16	0.82	0.07	-0.58	0.47	-0.17	1.1	-0.7

0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1
0	1	0	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	0	1

B. Apply DCT transform for the block C. signature mask
Fig . 12. An example of the fourth level for the signature of coefficients.

The new low value is (24) without any change, the high value is (27), while the mid value is ((24 + 27) / 2= 25).

The fifth level is hiding (25) bits of message "01101000011001010110110001" in the cover block from down to up by using least significant bit (LSB), the next operation is applying 2D-DCT transform for a block after the fifth hiding, next operation is finding the signature of coefficients matrix for the cover block. We notice that the mask matrix is equal to the signature matrix, which means, increase the bits of message hidden in the next level. Fig (13) explains the fourth level for the signature of coefficients.

119	118	119	119	119	118	118	118
118	119	118	118	118	118	119	118
119	118	118	119	118	118	118	119
119	118	118	118	119	118	119	119
119	119	118	119	118	118	118	119
118	119	118	119	119	118	119	118
118	119	118	118	119	119	118	118
118	118	118	119	118	119	119	118

A. The Cover block after level (5)

947.37	0.11	0.01	0.22	0.63	0.21	-0.95	-0.29
-0.07	0.58	0.54	-0.47	0.67	0.33	0.34	0.66
-0.3	0.07	-1.12	0.29	-0.9	0.28	0.29	0
0.35	0.81	-0.65	-0.17	-0.04	-0.52	-0.47	0.27
0.38	0.01	0.12	-0.06	0.63	0.63	0.63	0.2
0.24	-0.02	-0.37	0.51	0.57	-0.51	0.63	0.81
0.26	0.06	-0.46	-0.39	0.78	0.88	0.12	-0.15
-0.01	0.69	0.16	-0.33	0.64	-0.22	0.87	-0.9

0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1
0	1	0	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1

B. Apply DCT transform for the block C. signature mask Fig. 13. An example of the fifth level for the signature of coefficients.

The new low value is (25), the high value is (27) without change, while the mid value is ((25 + 27) / 2= 26).

The sixth level is hiding (25) bits of message "**01101000011001010110110000**" in the cover block from down to up by using least significant bit (LSB), the next operation is applying 2D-DCT transform for a block after sixth hiding, next operation is finding the signature of coefficients matrix for the cover block. . Also fig (14) explains the sixth level for the signature of coefficients.

119	118	119	119	119	118	118	118
118	119	118	118	118	118	119	118
119	118	118	119	118	118	118	119
119	118	118	118	119	118	119	119
119	119	118	119	118	118	118	119
118	119	118	119	119	118	119	118
118	119	118	118	119	119	118	118
118	118	118	119	118	119	119	118

A. The Cover block after level (6)

947.37	0.11	0.01	0.22	0.63	0.21	-0.95	-0.29
-0.07	0.58	0.54	-0.47	0.67	0.33	0.34	0.66
-0.3	0.07	-1.12	0.29	-0.9	0.28	0.29	0
0.35	0.81	-0.65	-0.17	-0.04	-0.52	-0.47	0.27
0.38	0.01	0.12	-0.06	0.63	0.63	0.63	0.2
0.24	-0.02	-0.37	0.51	0.57	-0.51	0.63	0.81
0.26	0.06	-0.46	-0.39	0.78	0.88	0.12	-0.15
-0.01	0.69	0.16	-0.33	0.64	-0.22	0.87	-0.9

0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1
0	1	0	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1

B. Apply DCT transform for the block C. signature mask Fig. 14. An example the sixth level for the signature of coefficients

The new low value is (26), the high value is (27) without change, while the mid value is ((26 + 27) / 2= 26).

We notice the mid value is equal to the low value .the algorithm will be finished in this level. The number of bits of the message that can be embedded in the cover block are (26) bits.

5) Hiding algorithm

From the above steps, the system determines the number of bits that can be embedded in a block of the red band of cover image, and collecting the two corresponding blocks of the green and the blue bands depending on the quad chain code which was used to collect a block of red band. This step will hide stream of bits of the message in the least significant bit of pixels of the three blocks (the first block that was extracted in step 2 from the red band, the second block of green band and the third block of blue band that were extracted based on the quad chain code of the red band) from bottom to up (that means the initial hiding of each block will be in location (7,7) and we continue until we reach the location (1,1)).

6) Reconstruct block of stego image

This step will reconstruct the block from the above step to the stego image by assigning new values after hiding bits of a message image.

7) Checking step

There are two checking levels:

- **The first one**: check if there are bits of the message still not hidden in the stego image, if yes there will be another checking, otherwise must go to the next step (save information of the red block in the boundary of the stego image.
- **The second one**: check if there is a block in the red band of cover

image that pixels are more similarity with each other. If yes then go to step (2). Otherwise must end the algorithm and upload another cover image because the current cover image cannot embed all bits of the secret image (message).

8) Save information of blocks in boundary

After the hiding operation for each block, the system will save the stream of chain code for the block of red band, start point of each block (red band) and number of hidden bits for blocks of red band in the first two boundaries of the red band of the stego image. The number of blocks that contain hidden bits will be saved in the 4 corner locations of stego image (embedded two bits for each corner location by using LSB algorithm, these locations are [0,0],[0,199],[199,199],[199,0]) . The stream of chain code(with size 15 for each block) will be embedded in the first boundary of the stego image except the corners, we note the values of chain code between (0-7) need 3 bits to be represented, then we need to embed 3 bits for each pixel in the first boundary of the stego image. After this step, the message hiding process is completed and the stego image is ready to be sent to the destination.

B. The extract message stage

In this stage there are several steps to extract a secret message from the stego image as follow:

1. Read the stego image and divide it into RGB (red band, green band, blue band).
2. Extract information of blocks that gets the hiding operation from the boundary of the red band.
3. Collect blocks with size (8*8) depending on the information hidden in the boundary of the red band, and find the two corresponding blocks of green and blue bands based on the same information that is used to collect the block of the red band.
4. Apply 2D DCT transform for each block and extract the signature of coefficients to determine locations of the block that gets hiding operation in it.
5. Extract bits of the message from blocks and save these bits in a buffer.
6. Convert bits in the buffer to the message image.

V. Experimental Results

Experiments of proposed method carried out to prove the efficiency. The proposed method has been simulated using the visual basic 6.0 program on Windows 7 platform on Intel core i5 2.5 GHz with 4 GB of main memory. The quality of the stego image is measured through the mean square error (MSE) that returns cumulative squared error between the cover image and the stego image , and the Peak Signal to Noise Ratio (PSNR) that returns the ratio of the maximum signal to noise between two images(cover, stego), in decibels. The best values of error measures are when the MSE is low and the PSNR is large. The mathematical equation for this error measures are:

$$PSNR(cover, stego) = 10 \times \log_{10}\left(\frac{255^2}{MSE}\right) \quad (3)$$

$$MSE(cover, stego) = \frac{\sum_{i=0}^{N-1}\sum_{j}^{M-1}(cover(i,j) - stego(i,j))}{N \times M} \quad (4)$$

Where N, M are the dimensions of the cover image and stego image. The proposed system was hiding gray image in color image. the hiding operation is done by embedding the pixels of message in a color image depending on the red band that it used to determine regions of cover image which will be hiding operation (each pixel of the message image (one byte) will be hidden in (three byte) of cover image, hide 2 bits in the red band, and hide 3 bits in the green band, and hide 3 bits in the blue band for each pixel in the cover pixels). There are two cases.

Case (1):

The cover image is a bitmap color image (24 bit for each pixel) with size (486*486), while the message is a bitmap gray scale image with size (50*50) which was used to test the proposed system as illustrated in fig (15). The similarity ratio between quads is 95%.

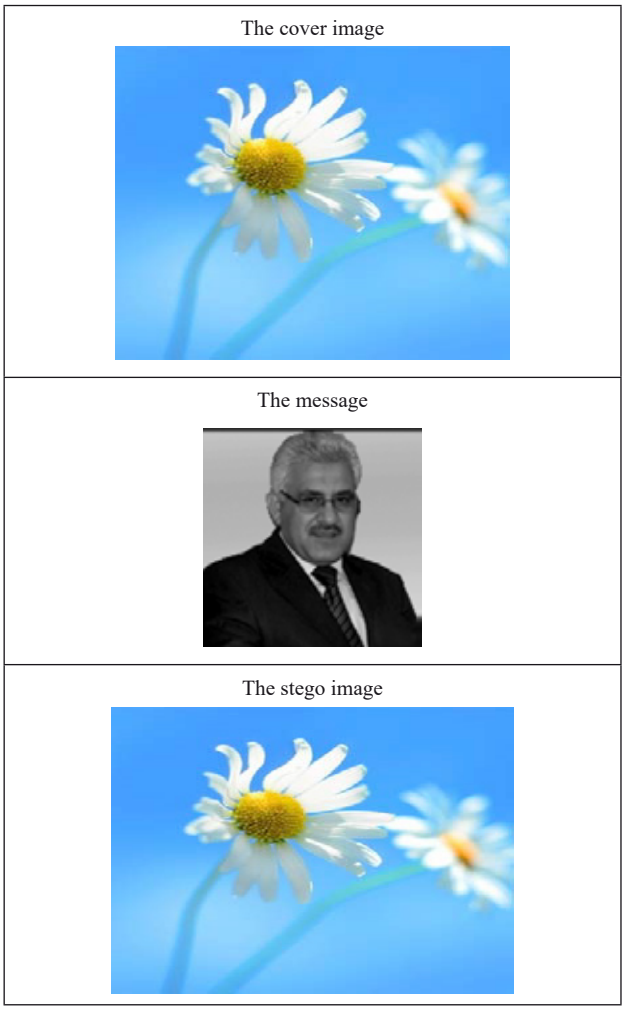


Fig. 15. The cover image, the message, the stego image for case (1)

We note the size of message image are (2500) pixels, then converting these pixels to the binary representation for a given stream of bits with size (20000) bits. Figure (16) explains simulation for hiding operation as follows.

1. Read the cover image.
2. Apply quad chain code by scanning the cover image to find locations of the pixels (quads) that are more similar with each other (as illustrate in green color).
3. Determine the start point for each block of the cover image(as illustrated in red color, the bold red color represent the start point of blocks that got hiding operation, while the remaining red color represent the start point of blocks that did not get hiding operation .
4. Determine the blocks (locations of the cover image) which got hiding process.

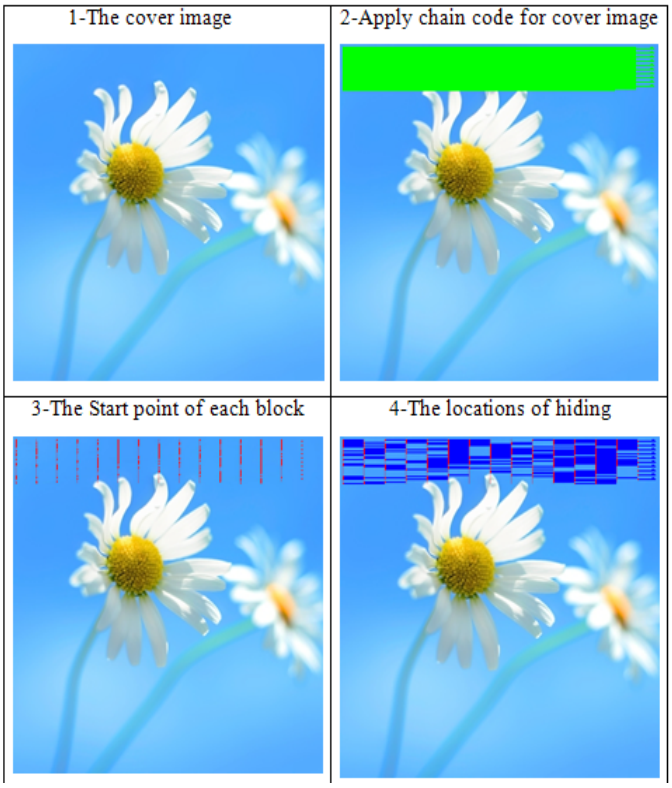


Fig. 16. Simulation of hiding operation for case (1).

Case (2):

The cover image is a bitmap color image (24 bit for each pixel) with size (450*450), while the message is a bitmap gray scale image with size (51*51) which was used to test the proposed system as illustrated in figure (17). The similarity ratio between quads is 96%.

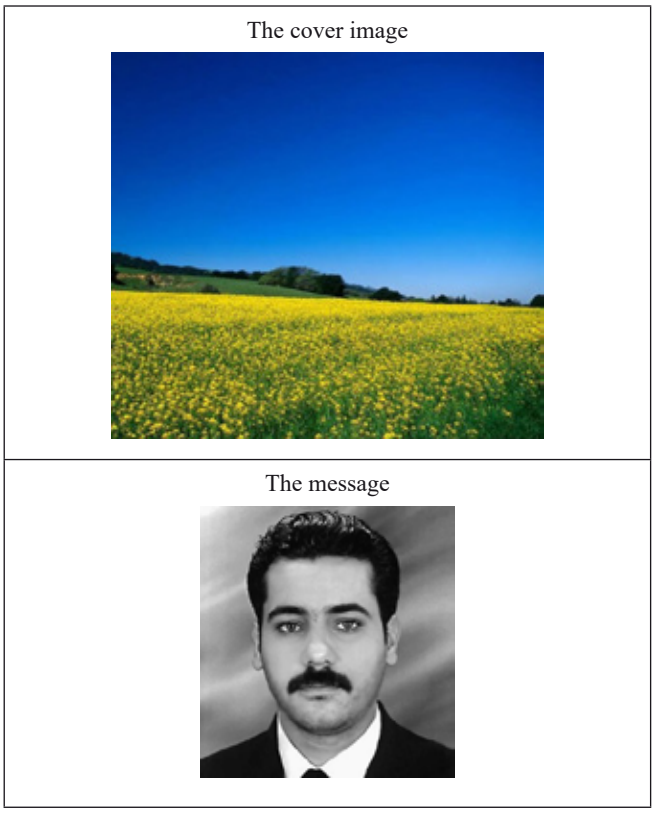


Fig.17.The cover image, the message, the stego image for case (17).

We note the size of message image are (2601) pixels, then converting these pixels to the binary representation for a given stream of bits with a size (20808) bit. Figure (18) explains simulation for hiding operation as follows.

1. Read the cover image.
2. Apply quad chain code by scanning the cover image to find locations of pixels (quads) that are more similar with each other (as illustrated in green color).
3. Determine the start point for each block of the cover image(as illustrate in red color, the bold red color represent the start point of blocks that got hiding operation, while the remaining red color represent start point of blocks that did not get hiding operation .
4. Determine the blocks (locations of the cover image) which got hiding process.

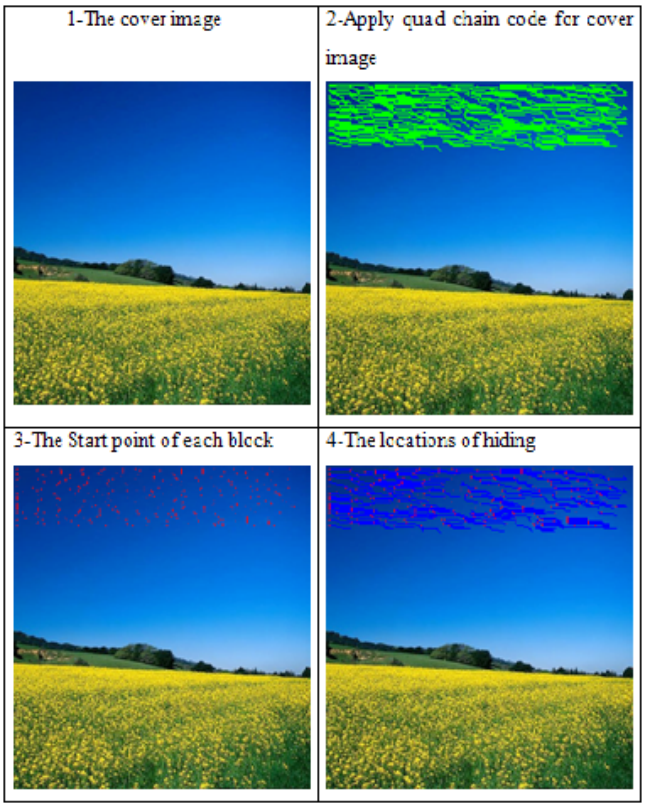


Fig. 18. Simulation of hiding operation for case (2).

After applying the proposed system on two cases, we have access the following information as illustrated in table (1):

TABLE 1
ILLUSTRATES THE ANALYSIS OF ALL CASES THAT ARE USED

Cover images	Case(1)	Case(2)
Size of cover image	486*486	450*450
Number of symbol (message)	2500 pixels	2601 pixels
Number of all blocks	458	306
Number of blocks which got in it operation hiding	250	225
Input Similarity Ratio	95%	96%
MSE with Boundary	0.6520	10.34
PSNR with Boundary	49.9881	37.9833
MSE without boundary	0.2542	0.4829
PSNR without boundary	54.0783	51.2916
Time for hiding	11 sec.	12 sec.
Time for extracting	3 sec.	3 sec.

VI. Conclusions

The proposed method is experimented and efficiency of the approach is demonstrated. The randomization that we apply makes this scheme stronger and secured. The randomization comes from three directions as follow:

1. The first one is applying chain code that select quad pixels that are not sequential but random (may take the zigzag shape).
2. The second one is applying the 2D-DCT transform to find what location gives zeros values (between 1,-1) which represent the signature of coefficients.
3. The third one is applying signature coefficients algorithm that is used to determine number of bits of the message that can be embedded in a block.

The proposed scheme can resist blind steganalysis schemes effectively. In the future, the security of the proposed scheme can be further improved by employing compression and encryption techniques.

References

[1] Joshua Michael Buchanan, "creating a robust form of steganography", department of computer science, Wake Forest University, North Carolina, USA, 2004..

[2] Peter Hanzlik,"Steganography in Reed-Solomon Codes", Lulea University of Technology, Sweden, 2011.

[3] Hardik Patel, Preeti Dave, "Steganography Technique Based on DCT Coefficients", International Journal of Engineering Research and Applications (IJERA), vol. 2, no. 1, pp.713-717, 2012.

[4] Zuheir H.Ali, "information hiding Using Chain code technique ", Al-Turath University College Magazine, vol. 1, no. 9, pp. 44-55, 2010.

[5] Ajit Danti and Preethi Acharya, "Randomized Embedding Scheme Based on DCT Coefficients for Image Steganography", IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition" RTIPPR, 2010.

[6] Deepak Singla and Rupali Syal, "Data Security Using LSB & DCT Steganography In Images", International Journal Of Computational Engineering Research, vol. 2, no. 2, 2012.

[7] Harsh Vikram Singh, information hiding technique for image cover, LAMBERT Academic Publishing, 2010.

[8] Vidyabharati Mahavidyalaya, "INFORMATION HIDING TECHNOLOGY- A WATERMARKING", Advances in Computational Research, vol. 3, no. 1, PP-37-41, 2011.

[9] Mehdi Kharrazi, Husrev T. Sencar, and Nasir Memon, "Image Steganography: Concepts and Practice", WSPC/Lecture Notes Series, Polytechnic University, Brooklyn, NY 11201, USA, 2004.

[10] Manish Mahajan , Dr. Navdeep Kaur, "Adaptive Steganography: A survey of Recent Statistical Aware Steganography Techniques", Computer Network and Information Security, 2012.

[11] R. C. Gonzalez and R. E. Woods, "Digital Image Processing", Second Edition, Prentice-Hall Inc, 2002.

[12] Swastik Das and Rasmi Ranjan Sethy, "Digital Image Compression Using Discrete Cosine Transform & Discrete Wavelet Transform", National Institute of Technology, Rourkela, India, 2009.

[13] Suchitra Shrestha, "Hybrid Dwt-Dct Algorithm For Image And Video Compression Applications", Master thesis, University of Saskatchewan, Saskatoon, Canada, 2010.

[14] Vikas Thada, Dr Vivek Jaglan , "Comparison of Jaccard, Dice, Cosine Similarity Coefficient To Find Best Fitness Value for Web Retrieved Documents Using Genetic Algorithm" ,International Journal of Innovations in Engineering and Technology (IJIET), vol. 2 no. 4, pp. 202-205, 2013.

Prof. Dr. Tawfiq A. Al-asadi is a professor at the college of Information technology at the University of Babylon, Iraq. His research interests are primarily in image processing, computer graphics and data compression.

Prof. Dr. Israa Hadi Ali is a professor at the Department of Software in college of IT at the University of Babylon, Iraq. Her research interests are primarily multimedia and video tracking.

M.Sc. Abdul kadhem Abdul kareem Abdul kadhem is a graduate student at the college of information technology, Department of software, university of Babylon, Iraq. He is currently pursuing his MSC degree. Abdul Kadhem holds a BS degree from the Department of computer science from university of Babylon

An Automated Negotiation-based Framework via Multi-Agent System for the Construction Domain

Moamin A. Mahmoud, Mohd Sharifuddin Ahmad, Mohd Zaliman M. Yusoff, and Arazi Idrus

Dept. of Computer Science and Information Technology, Universiti Tenaga Nasional (UNITEN), Malaysia

***Abstract* — In this paper, we propose an automated multi-agent negotiation framework for decision making in the construction domain. It enables software agents to conduct negotiations and autonomously make decisions. The proposed framework consists of two types of components, internal and external. Internal components are integrated into the agent architecture while the external components are blended within the environment to facilitate the negotiation process. The internal components are negotiation algorithm, negotiation style, negotiation protocol, and solution generators. The external components are the negotiation base and the conflict resolution algorithm. We also discuss the decision making process flow in such system. There are three main processes in decision making for specific projects, which are propose solutions, negotiate solutions and handling conflict outcomes (conflict resolution). We finally present the proposed architecture that enables software agents to conduct automated negotiation in the construction domain.**

***Keywords* — Intelligent Software Agent, Multi-agent Systems, Agent and Negotiation, Automated Negotiation, Value Management, Construction Domain.**

I. Introduction

In the construction domain, deciding on a new project is dependent upon a company's strategy. If the strategy is based on a decision made by a stakeholder, then it takes a very short time to decide. However, such decision has no value in terms of value management, because the decision-making process does not include other experienced stakeholders that hold different backgrounds.

Figure 1 considers a value management approach that emphasizes on involving various stakeholders in the decision-making process to arrive at a single valued solution. In other words, the various stakeholders with different backgrounds that have a stake in the project must contribute to the decision. In fact, these stakeholders often belong to different departments and possess different perspectives about the solutions according to their background and positions they hold.

For example, a project manager usually cares more about the cost of a project than the function while a design manager is more concerned about the function than the cost. Thus, for any decision to be made regarding a new project, stakeholders must propose a single optimal solution. However, a problem may arise when stakeholders propose more than one solution. In such situation, stakeholders need to negotiate on the proposed solutions and agree on a single solution. But the negotiation may not be easy and smooth because when stakeholders possess different backgrounds, often their views about the optimal solution for a particular project are different. Such differences cause conflicts in arriving at a decision. In addition, stakeholders may work at different branches throughout the country or other parts of the world which make a meeting for decision more difficult and costly. While applying Value Management on decision making in the construction domain is useful, it faces communication difficulties between stockholders and conflicting issues that require negotiation.

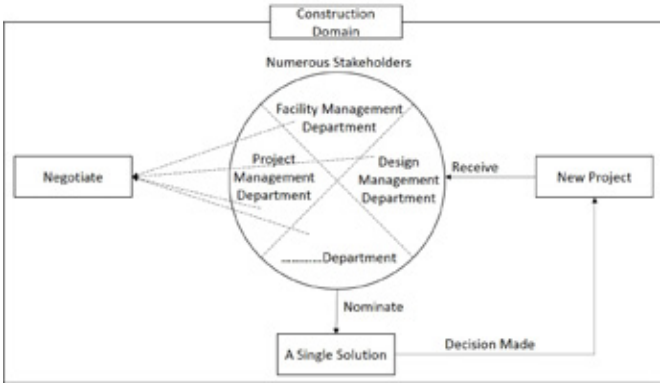


Figure 1 A Decision Making based on Value Management in Construction Domain

In this paper, we attempt to overcome these difficulties by proposing a Value-based Automated Negotiation Model utilizing the multi-agent system's approach. It enables software agents to conduct negotiations and autonomously arrive at a decision.

While this work is inspired by the work of Utomo [1], his study is only in conceptual level and lacks an intelligent agent architecture that aids an agent to interact with other agents and respond to its environment and eventually influences its autonomy level in decision making. Automated Negotiation as a very complicated system could not be efficiently used if agents have trivial architecture. Moreover, he does not incorporate the different negotiation styles to the agent architecture which could help the agents in mimicking humans' styles in negotiation. Briefly, the major development that we intend to do is to develop concrete agent architecture such as the Belief-Desir-Intention architecture and explore the potential components that an agent could employ to conduct useful and efficient negotiations. Consequently, we consult the various resources that are presented by Utomo [1] to come up with our framework.

The next section dwells upon the related work on automated negotiation. Section 3 presents the proposed framework. In Section 4, we discuss the decision making process flow. Section 5 presents the agent architecture and Section 6 concludes the paper.

II. Related Work

In this section, we discuss two prominent topics to this research which are, value management, and applications of negotiation in multi-agent systems.

Value Management (VM) is defined as "a structured, organized team approach to identify the functions of a project, product, or service

DOI: 10.9781/ijimai.2015.353

that will recognize techniques and provide the necessary functions to meet the required performance at the lowest overall cost" [17]. VM works on identifying and eliminating unnecessary cost [18] but without affecting a quality parameter [19]. VM is based on data collection method from reliable resources and functional requirements to fulfill the needs, wants and desires of the customers [1]. According to Kelly and Male [20], VM is a multidisciplinary, team-oriented approach to problem solving.

The application of VM in decision making has been reported by many researchers [1, 21, 22]. One of the techniques that is relevant to VM is weighting and scoring in which a decision needs to be made in selecting an option from a number of competing options, and the best option is not immediately identifiable [1, 23, 24].

Intelligent software agents have been widely used in distributed artificial intelligence and due to their autonomous, self-interested and rational abilities, agents are well-suited for automated negotiation on behalf of humans [2, 25, 26, 27, 28, 29, 30]. According to Kexing [2], automated negotiation is a system that applies artificial intelligence and information and communication technology to negotiation strategies, utilizing agent and decision theories.

Numerous research have discussed the negotiation on multi-agents systems in various domains [3, 4, 5, 6, 7]. Few of them study the issues of conflict resolution and negotiation in construction domain [1, 8, 9]. Anumba et al. [9] presented two main negotiation theories; mechanical and behavior theories. The mechanical theory is inspired by game theories which are mathematical models relied on rational behavior assumption, while the behavior theory studies human behavior in negotiation.

Coutinho et al. [10] proposed a negotiation framework to serve collaboration in enterprise networks to improve the sustainability of interoperability within enterprise information systems. Utomo [1] presented a conceptual model of automated negotiation that consists of methodology of negotiation and agent based negotiation. Dzeng and Lin [11] presented an agent-based system to support negotiation between construction and suppliers via the Internet. Anumba et al. [12] proposed a collaborative design of light industrial buildings based on multi-agent systems to automate the interaction and negotiation between the design members. Ren et al. [4] developed a multi-agent system representing participants, who negotiate with each other to resolve construction claims.

III. A Conceptual Automated Negotiation Framework

From our initial investigation of the literature [1, 2, 3, 4, 8, 9], we observe that agents need to be integrated with six main components to conduct negotiations, which are classified, in this research, into internal and external components. The internal components are negotiation algorithm, negotiation style, negotiation protocol, and solution generator. The components are integrated with a BDI agent architecture (as discussed in Section 5). The external components are the negotiation base and the conflict resolution algorithm.

As shown in Figure 2, the negotiation algorithm presents a formal and intelligent procedure that maintains negotiations with other agents. Each agent is endowed with a negotiation style that represents the agent's approach to negotiation. Each agent possesses one style, either competing style or collaborating style.

For agents to conduct negotiations systematically, they must have a negotiation protocol, which controls the negotiations process between agents. For example, agents could possibly negotiate individually or form groups (coalitions) before negotiating. They could also negotiate directly by sending messages to each other or deploy another method to share their inputs. Finally, agents must be able to generate solutions that conform to their interests and reap the benefits of negotiation.

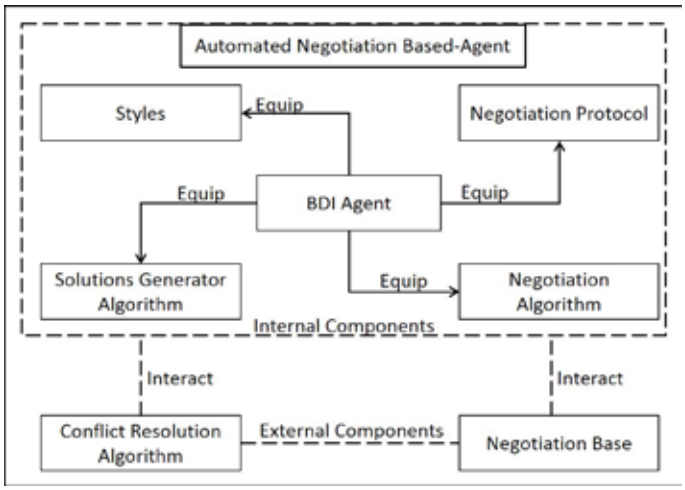


Figure 2. Automated Negotiation Model based on Multi-agent System (AN-MAS)

The Negotiation Base represents the negotiation hub that contains suggested solutions of negotiations used by agents in sharing their solutions and form coalitions. The base reduces direct interactions between agents that would increase the network load.

The conflict resolution algorithm handles negotiations outcomes. If agents have not agreed on a single solution, the conflict resolution algorithm works on solving that conflict. Consequently, the proposed framework with the internal and external components manifests the Automated Negotiation Model based on the Multi-agent System (AN-MAS).

A. A Negotiation Styles

According to Utomo [1], there are five main negotiation styles that constitute two types of outcomes, which are Competing, Avoiding, Collaborating, Accommodating, and Compromising, and the two outcomes are Cooperative and Assertiveness.

Accordingly, Each agent possess one style, and this style forms one negotiation outcome whether it is Cooperative or Assertiveness, e.g. agent a is Cooperative type and possess Accommodating style.

B. Negotiation Protocol

Agents conduct negotiation according to a predefined protocol. Such protocol ensures that the negotiation progresses smoothly.

C. Solution Generator Algorithm

For an agent to conduct negotiation, it should be able to propose solutions and rank them from 1st to nth solution for the next stage of the negotiation operation.

In real situations, various stakeholders have different level of interest about the cost and function parameters based on their positions and values they uphold. Thus, those stakeholders appraise their solutions based on their interest level on these parameters. For example, in the construction domain, a Design Manager cares more about the function in contrast with a Project Manager who cares more about the cost, while a Facility Manager's interest is in between the Design and Project Managers' interests. Therefore, the Design Manager normally attempts to find a solution that provides high function, whereas the Project Manager normally attempts to find a solution that provides low cost. The Facility Manager attempts to find a moderate solution that provides acceptable cost and function.

Consequently, The Solutions Generator Algorithm will be inspired by the two main parameters of Value Management which are Cost and Function to deliver value solutions.

D. Negotiation base

The Negotiation Base represents the negotiation hub that is used by agents to form negotiations by sharing their solutions and form coalitions. The base helps in reducing direct interactions between agents that increase the network load. All negotiations are processed via this base which is accessible by all agents.

E. Negotiation Algorithm

The algorithm implements the negotiation process between agents. The process starts when each agent submits its solutions to the negotiation base. Each agent then reviews each solution's and accordingly sets a plan to conduct negotiation.

F. Conflict Resolution Algorithm

The need for this algorithm is based on the negotiation algorithm outcomes. Since any project needs a single solution, then when the negotiation algorithm outcome is a single solution, agents skip this algorithm. But when the outcome is several solutions, then another process is needed to resolve this conflict. Such situation represents a conflict between agents about the solution of that project.

IV. Decision Making Process Flow

A decision made by agents goes through several processes. These processes work by gradually reducing candidate solutions of a project until a single solution is reached. Consequently, in this work, the process of nominating a single solution from a set of solutions is called decision making.

There are three main processes in decision making for a specific project, which are propose solutions, negotiate solutions and handling conflict outcomes (conflict resolution).

- Propose solutions: In this process, each agent proposes solutions and ranks them from 1st to nth solution where n is any natural number.
- Negotiate solutions: When ranked solutions are ready, agents negotiate by submitting their ranked solutions to each other. Since each agent's target is to maximize its utility by selecting a solution that has a better order, each agent prepares a plan. Using these plans, agents form coalitions among them based on similar plans. These coalitions continuously compare plans with each other until a single or more solutions converge after exhausting all attempts.
- Resolve conflict: If agent coalitions agree upon a single solution, then this process is forfeited, but if there are two or more conflicting solutions, then the conflicts need to be resolved. This process resolves conflicts based on each coalition's strength and its solutions' risks. From these two parameters, this process drops solutions until a single solution is reached.

Figure 3 shows the decision making flowchart as described above. The process starts when agents receive a new project. The agents first propose solutions in ranked order. They then negotiate these solutions. If they agree upon a single solution, then the decision is made, otherwise, the conflict resolution process takes over to drop the weak and risky solutions. If the outcome of the conflict resolution process is a single solution then the decision is made. Otherwise, the agents negotiate the outcome of the conflict resolution process. Ultimately, one coalition's solution is accepted.

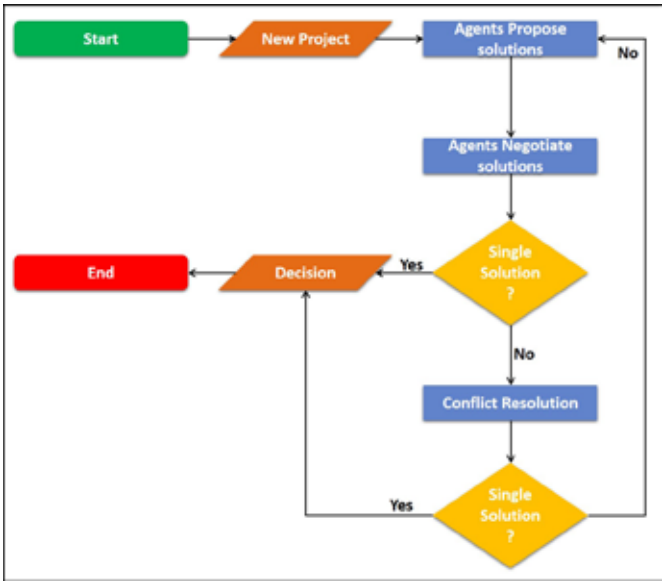


Figure 3 Decision Making Flowchart

V. The Belief-Desire-Intention (BDI) Agent Architecture

This section presents an architecture that enables software agents to mimic human behaviors and styles in building an automated negotiation system in the construction domain. In this work, we develop BDI agents that are widely used by researchers to build intelligent agents.

The BDI agent consists of three main components that are affected by the environment; Belief, Desire, and Intention. Agents usually perform tasks within an environment and they exploit the environment to update their goals. The agents' beliefs are influenced by the environmental changes. The belief in turn updates their desires and the intentions.

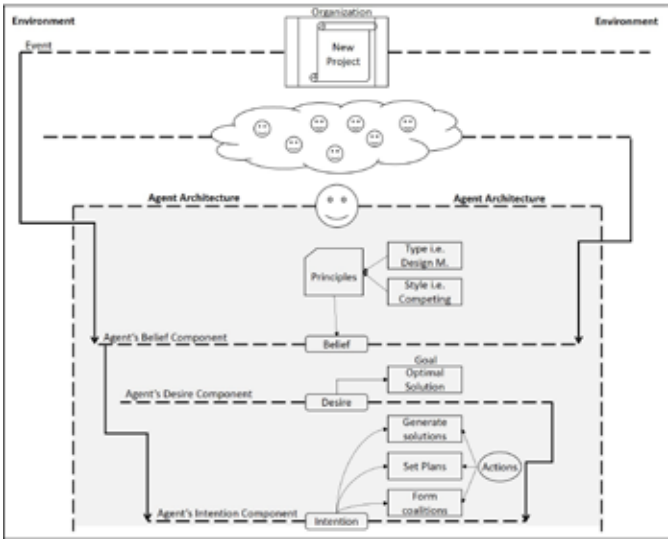


Figure 4. The Proposed Agent Architecture

As shown in Figure 4, the proposed agent architecture consists of the outer area that represents the environment and the inner area that represents the decision making process. The environment constitutes the variables of new project information and agent's activities information, e.g. interactions, decisions, negotiations, coalitions. The belief component within the architecture is influenced by the environment and agent attributes, which include the agent type, e.g.

Design Manager, and the agent style e.g. competing. The desire component represents an agent's goal.

In the construction domain, each agent attempts to ensure its first rank solution wins. If it is not possible, it works on the second rank and so on. This scenario represents its desire or goal. From the agent's belief and desire, it performs actions which represent the intention components. The intention component represents a bridge between the belief and the desire, in other words, it represents the practical steps to achieve the desire according to its belief about the environment and the attributes.

VI. Conclusion and Future Work

To create a multi-agent automated negotiation model, agents need to be integrated with several components. In this paper, we identify four internal components (negotiation algorithm, negotiation style, negotiation protocol, and solution generators) integrated with the agent design and two external components (the negotiation base and the conflict resolution algorithm) within the environment. These components constitute the proposed framework.

We also discuss the decision making process flow in such system, consisting of three main processes which are propose solutions, negotiate solutions and handling conflict outcomes (conflict resolution). We finally reveal our proposed agent's architecture to conduct automated negotiation in the construction domain.

Since this work is in its theoretical stage, it only presents the conceptual underpinnings of pertinent issues in negotiation and does not present the experimental results. Such outcome will be presented in our future work.

In addition, for our future work, we shall study and propose mechanisms for the three methods needed by the decision making process which are Agent Proposes Solutions, Agent Negotiate Solutions and Conflict Resolution.

References

[1] Utomo C., Development of a negotiation Support Model for Value Management in Construction, PhD Thesis, University Teknologi PETRONAS, December 2009.

[2] Kexing L.. A survey of agent based automated negotiation. In Network Computing and Information Security (NCIS), 2011 International Conference on, vol. 2, pp. 24–27 (IEEE, 2011).

[3] Beer M., d'Inverno M., Jennings R.N., Luck M., Preist C., Schroeder M., Negotiation in multi-agent systems Knowledge Engineering Review, 14 (3) (1999), pp. 285–289

[4] Z. Ren, C.J. Anumba, Multi-agent systems in construction—state of the art and prospects, Automation in Construction, 13 (2004), pp. 421–434

[5] M. Wang, H. Wang, D. Vogel, K. Kumar, D.K.W. Chiu Agent-based negotiation and decision making for dynamic supply chain formation Engineering Applications of Artificial Intelligence, 22 (7) (2009), pp. 1046–1055

[6] Utomo C., Idrus A., A Concept toward Negotiation Support for Value Management on Sustainable Construction, Journal of Sustainable Development Vol 4, No 6 (2011).

[7] Victor Sanchez-Anguix , Vicente Julian , Vicente Botti , Ana García-Fornes, Tasks for agent-based negotiation teams: Analysis, review, and challenges, Engineering Applications of Artificial Intelligence, v.26 n.10, p.2480-2494, November, 2013.

[8] Anderson, R. M., Hobbs, B. F., & Bell, M. L. (2002). Multi-objective decision-making in negotiation and conflict resolution. Encyclopedia of Life Support Systems. New York: Eolss Publishers.,

[9] C. J. Anumba , O. O. Ugwu , Z. Ren, Agents And Multi-agent Systems In Construction, Spon Press, 2005.

[10] Coutinho, C., Cretant, A., Ferreira da Silva, C., Ghodous, P., & Jardim-Goncalves, R. (2014). Service-based negotiation for advanced collaboration in enterprise networks. Journal of Intelligent Manufacturing. doi:10.1007/s10845-013-0857-4.

[11] Dzeng, R. J., & Lin, Y. C. (2004). Intelligent agents for supporting construction procurement negotiation. Expert Systems with Applications, 27(1), 107-119.

[12] C.J. Anumba, Z. Ren, A. Thorpe, O.O. Ugwu, L. Newnham, Negotiation within a multi-agent system for the collaborative design of light industrial buildings Adv Eng Software, 34 (7) (2003), pp. 389–401

[13] Johnson, H., Johnson, P., Task Knowledge Structures: Psychological basis and integration into system design, Acta Psychologica, Vol. 78, 1991, pp. 3-26.

[14] Friend, M. and Cook, L. The new Mainstreeem. Instructor 30-36.

[15] Matthews , J. Implications for collaborative educators preparation and development: A sample instructional approach in Diana G Pounder (Editor) Restructuring School of Collaboration: Promise and Pitfalls, State University of NewYwork Press. 1998

[16] Holley, W. H. and Jennings K.M. The labour relations process. 8th ed., Chapter 6. Mason, OH: Thomsan/South-Westeren, 2005.

[17] SAVE International, value methodology standards, 2001.

[18] Kelly J. and Male S., Value Management. Chapter 5 in Kelly J., Morledge R. and Wilkinson S (eds.) Best Value in Construction Oxford UK, Blackwell, pp 77-99, 2002.

[19] Mukhopadhyaya A.K., Value Engineering Concept, Techniques and Applications, Response Books, New Delhi, 2003.

[20] Kelly J. and Male S., Value Management in Decision and Construction, The Economic Management of Projects. Spon Press, London.

[21] Jaapar, A., Endut, I.R., Bari, N.A.A. and Takim, R. The impact of value management implementation in Malaysia. Journal of Sustainable Development 2 (2), 2009.

[22] Shen, Q., Chung, J.K.H., Li, H. and Shen, L.. A Group Support System for improving value management studies in construction. Automation in Construction, 13 (2004): 209–224, 2004.

[23] Cariaga, I, El-Diraby, T and Osman, H., Integrating Value Analysis and Quality Function Deployment for Evaluating Design Alternatives. Construction Engineering and Management, 133(10), 761-770, 2007.

[24] Qing Y. and Wanhua Q., 2007, Value Engineering Analysis and Evaluation. For the Second Beijing Capital Airport. Value World, Spring, SAVE International.

[25] Ahmed M., Ahmad M S, Yusoff M Z M, Modeling Agent-based Collaborative Process , The 2nd International Conference on Computational Collective Intelligence Technology and Applications (ICCCI 2010), pp. 296-305, ISBN:3-642-16692-X 978-3-642-16692-1, 10-12 November, 2010 Taiwan.

[26] Moamin A. Mahmoud, Mohd Sharifuddin Ahmad, Mohd Zaliman M. Yusoff, and Arazi Idrus. "Automated Multi-agent Negotiation Framework for the Construction Domain." Distributed Computing and Artificial Intelligence, (DCAI' 15), Spain, 3th-5th June, 2015. In Springer's Advances in Intelligent Systems and Computing, Springer International Publishing, Volume 373, 2015, pp 203-210.

[27] Itaiwi A. K., Ahmad M. S., Hamid N. H. A., Jaafar N. H., Mahmoud M. A., A Framework for Resolving Task Overload Problems Using Intelligent Software Agents, 2011 IEEE International Conference on Control System, Computing and Engineering, ICCSCE11,2011.

[28] Ahmed M., Ahmad M. S., and Yusoff M. Z. M., "A Collaborative Framework for Multiagent Systems." International Journal of Agent Technologies and Systems (IJATS), 3(4):1-18, 2011.

[29] Ahmed M., Ahmad M. S., and Yusoff M. Z. M.,, Mitigating Human-Human Collaboration Problems using Software Agents, The 4th International KES Symposium on Agents and Multi-Agent Systems – Technologies and Application (AMSTA 2010), pp. 203-212, ISBN:3-642-13479-3 978-3-642-13479-1, Gdynia, Poland, 23 – 25 June 2010.

[30] Al-Mutazbellah Khamees Itaiwi, Mohd Sharifuddin Ahmad, Nurzeatul Hamimah Abd Hamid, Nur Huda Jaafar, Moamin A. Mahmoud, A Multi-agent Framework for Dynamic Task Assignment and Delegation in Workload Distribution, International Conference on Computer & Information Sciences, 12 June 2012.

Moamin A. Mahmoud obtained his Bachelor in Mathematics from the College of Mathematics and Computer Science, University of Mosul, Iraq in 2007. He obtained his Master of Information Technology at the College of Graduate Studies, Universiti Tenaga Nasional (UNITEN), Malaysia in 2010, and PhD of Information and Communication Technology from Universiti Tenaga Nasional (UNITEN), Malaysia in 2013. His research interests are in the area of software agents, agent behavior in open societies, and social sciences simulation.

Mohd Sharifuddin. Ahmad received his B.Sc. in Electrical and Electronic Engineering from Brighton Polytechnic, UK in 1980. He started his career as a power plant engineer specialising in Process Instrumentation and Control in 1980. After completing his MSc in Artificial Intelligence from Cranfield University, UK in 1995, he joined UNITEN as a Principal Lecturer and Head of Dept. of Computer Science and Information Technology. He obtained his PhD from Imperial College, London, UK in 2005. He has been an associate professor at UNITEN since 2006. His research interests include applying constraints to develop collaborative frameworks in multi-agent systems, collaborative interactions in multi-agent systems and tacit knowledge management using AI techniques.

Mohd Zaliman Mohd Yusoff obtained his Msc and Phd in Computer Science from Universiti Kebangsaan Malaysia in 1998 and 2013 respectively. He started his career as a Lecturer at UNITEN in 1998 and has been appointed as a Principle Lecturer at UNITEN since 2008. His has produced and presented about 86 papers for local and international conferences. His research interest includes modeling and applying emotions in various domains including educational systems, norms based system including detection and assimilation strategies using software agents achitecture. He is also interested in modeling agent's trust and reputation in computer forensic and knowledge discovery systems. He is an active member of IEEE (91279801) and serves as the exco committee of Computer Society Malaysia Chapter since 2013.

AraziIdrusB.Eng.(Hons)inCivilandStructuralEngineering (Sheffield University, UK), M.Sc. (Cranfield University, UK), Ph.D (Imperial College, University of London, UK). Field of Specialization: Project Management, Construction It, Construction Productivity, Value Management; PPP/PFI, IBS Construction, Building Maintenance Management, Risk-Based Assessment, Concrete Repairs,Blast Design, and Load Resistance Factor Design of Offshore Structures

Checking RTECTL properties of STSs via SMT-based Bounded Model Checking

Agnieszka M. Zbrzezny, Andrzej Zbrzezny

MCS, Jan Długosz University in Częstochowa, Poland

***Abstract* —We present an SMT-based bounded model checking (BMC) method for Simply-Timed Systems (STSs) and for the existential fragment of the Real-time Computation Tree Logic. We implemented the SMT-based BMC algorithm and compared it with the SAT-based BMC method for the same systems and the same property language on several benchmarks for STSs. For the SAT-based BMC we used the PicoSAT solver and for the SMT-based BMC we used the Z3 solver. The experimental results show that the SMT-based BMC performs quite well and is, in fact, sometimes significantly faster than the tested SAT-based BMC.**

***Keywords* — RTECTL, SMT-Based Bounded Model Checking, STS**

I. Introduction

Verification of soft real-time systems is an actively developing field of research [2,9,10]. Popular models of such systems include, among others, timed automata [1], and simply-timed systems (STSs) [5], i.e., Kripke models where each transition holds a duration, which can be any integer value (including zero).

The fundamental thought behind bounded model checking (BMC) is, given a system, a property, and an integer bound $k \geq 0$, to define a formula such that the formula is satisfiable if and only if the system has a counterexample (of the length at most k) violating the property. The bound is incremented until a satisfiable formula is discovered or a completeness threshold is reached without discovering any satisfiable formulae. The SMT problem [3] is a generalisation of the SAT problem, where Boolean variables are replaced by predicates from various background theories, such as linear, real, and integer arithmetic. SMT generalises SAT by adding equality reasoning, arithmetic, fixed-size bit-vectors, arrays, quantifiers, and other useful first-order theories.

There are three main reasons why it is interesting to consider STSs instead of standard Kripke models. First, STSs allow for transitions that take a long time, e.g. 100 time units. Such transitions could be simulated in standard Kripke models by inserting 99 intermediate states. But this increases the size of the model, and so it makes the model checking process more difficult. Second, STSs allow transitions to have zero duration. This is very convenient in models where some steps are described indirectly, as a short succession of micro-steps. Third, the transitions with the zero duration allow for counting specific events only and thus omitting the irrelevant ones from the model checking point of view.

The original contribution of this paper consists in defining a SMT-based BMC method for the existential fragment of RTCTL (RTECTL) interpreted over simply-timed systems (STSs) generated by simply-timed automata with discrete data (STADDs). We implemented our SMT-based BMC algorithm and we compared it with the SAT-based BMC method for RTCTL and STSs. For a constructive evaluation of our SMT-based BMC method we have used two scalable benchmarks: a modified *bridge-crossing problem* [8] and a modified *generic pipeline paradigm* [7].

The rest of the paper is organised as follows. We begin in Section II by introducing simply-timed automata with discrete data, simply-timed systems, and we present the syntax and semantics of RTECTL over simply-timed systems. In Section III we present our SMT-based BMC method for RTECTL and simply-timed systems. In Section IV we discuss our experimental results. In the last section we conclude the paper.

II. Preliminaries

In this section we first define simply-timed automata with discrete data and simply-timed systems, and next we introduce syntax and semantics of RTECTL. The formalism of STADD was introduced in [12] and formalism of STS in [10].

A. Simply-timed automata with discrete data and simply-timed systems

Let $\mathbb{Z}$ be the set of integer numbers, $\bar{Z}$ a finite set of integer variables $\{z_1, \ldots, z_2\}$, $c \in Z$, $z \in \bar{Z}$, and $\oplus \in \{+, -, *, mod, \div\}$. The set $Expr(\bar{Z})$ of all the *arithmetic expressions* over $\bar{Z}$ is defined by the following grammar: $ae ::= c \mid z \mid ae \oplus ae \mid -ae \mid (ae)$. Next, for $ae \in Expr(\bar{Z})$ and $\sim \in \{=, \neq, <, \leq, \geq, >\}$, the set $BoE(\bar{Z})$ of all the Boolean expressions over $\bar{Z}$ is defined by the following grammar:

$$\beta ::= true \mid ae \sim ae \mid \beta \wedge \beta \mid \beta \vee \beta \mid \neg\beta \mid (\beta).$$

For $z \in \bar{Z}$, $ae \in Expr(\bar{Z})$, ϵ denoting the empty sequence, the set $SimAss(\bar{Z})$ of all the simultanoues assignments over $\bar{Z}$ is defined as $\alpha ::= \epsilon \mid z_{i_1}, \ldots, z_{i_m} := ae_{i_1}, \ldots, ae_{i_m}$, where $i_j \in \{1, \ldots, n\}$ and any $z \in \bar{Z}$ appears on the left-hand side of $:=$ at most once.

A *variables valuation* is a total mapping $v : \bar{Z} \to \mathbb{Z}$. We extend this mapping to expressions of $Expr(\bar{Z})$ in the usual way. Moreover, we assume that a domain of values for each variable is finite. Satisfiability of a Boolean expression $\beta \in BoE(\bar{Z})$ by a valuation v, denoted $v \models \beta$, is defined inductively as follows: $v \models true$, $v \models ae_1 \sim ae_2$ iff $v(ae_1) \sim v(ae_2)$, $v \models \beta_1 \wedge \beta_2$ iff $v \models \beta_1$ and $v \models \beta_2$, $v \models \beta_1 \vee \beta_2$ iff $v \models \beta_1$ or $v \models \beta_2$, $v \models \neg\beta$ iff $v \not\models \beta$, $v \models (\beta)$ iff $(v \models \beta)$. Given a variables valuation v and an instruction $\alpha \in Ins(\mathbb{Z})$, we denote by $v(\alpha)$ a valuation v' such that: if $\alpha = \epsilon$, then $v' = v$; if $\alpha = (z := expr)$, then for all $z' \in \mathbb{Z}$ it holds $v'(z') = v(expr)$ if $z' = z$, and $v'(z') = v(z')$ otherwise; if $\alpha = \alpha_1; \alpha_2$, then $v' = (v(\alpha_1))(\alpha_2)$.

DOI: 10.9781/ijimai.2015.354

Definition 1. *Let* PV *be a set of atomic propositions and* $\mathbb{N}=\{0,1,2,\ldots\}$. A simply-timed automaton with discrete data (STADD) *is a tuple* $A=(\Sigma,L,l^0,Z,E,d,V_A)$, *where* Σ *is a finite set of actions,* L *is a finite set of locations,* l^0 *is an initial location,* $\bar{Z}$ *is a finite set of integer variables,* $E\subseteq L\times\Sigma\times BoE(\bar{Z})\times SimAss(\bar{Z})\times L$ *is a transition relation,* $d:\Sigma\to\mathbb{N}$ *is a duration function, and* $V_A:L\to 2^{PV}$ *is a valuation function that assigns to each location a set of propositional variables that are assumed to be true at that location.*

The semantics of the STADD is defined by associating to it a simply-timed system as defined below.

Definition 2. *Let* PV *be a set of atomic propositions,* $v^0:\bar{Z}\to\mathbb{Z}$ *an initial variables valuation, and* $A=(\Sigma,L,l^0,\bar{Z},E,d,V_A)$ *a simply-timed automaton with discrete data.* A simply-timed system *(or* a model*) for* A *is a tuple* $M=(\Sigma,S,\iota,T,d,V)$, *where* Σ *is a finite set of actions of* A, $S=L\times\mathbb{Z}^{|\bar{Z}|}$ *is a set of states,* $d:\Sigma\to N$ *is the duration function of* A, $V:S\to 2^{PV}$ *is a valuation function defined as* $V((l,v))=V_A(l)$, *and* $\iota=(l^0,v^0)\in S$ *is the initial state,* $T\subseteq S\times\Sigma\times S$ *is the smallest simply-timed transition relation defined in the following way: for* $\sigma\in\Sigma$, $(l,v)\xrightarrow{\sigma}(l',v')$ *iff there exists a transition* $(l,\sigma,\beta,\alpha,l')\in E$ *such that* $v\models\beta$, $v'=v(\alpha)$. *We assume that the relation* T *is total, i.e., for any* $s\in S$ *there exists* $s'\in S$ *and an action* $\sigma\in\Sigma$ *such that* $(s,\sigma,s')\in T$ *(or* $s\xrightarrow{\sigma}s'$*).*

A *path* in M is an infinite sequence $\pi=s_0\xrightarrow{\sigma_1}s_1\xrightarrow{\sigma_2}s_2\xrightarrow{\sigma_2}\ldots$ of transitions. For such a path, and for $m\in\mathbb{N}$, by $\pi(m)$ we denote the m-th state s_m. For $j\le m\in\mathbb{N}$, $\pi[j..m]$ denotes the finite sequence $s_j\xrightarrow{\sigma_{j+1}}s_{j+1}\xrightarrow{\sigma_{j+2}}\ldots s_m$ with $m-j$ transitions and $m-j+1$ states. The (cumulative) duration $D\pi[j..m]$ of such a finite sequence is $d(\sigma_{j+1})+\ldots+d(\sigma_m)$ (hence 0 when $j=m$). By $\Pi(s)$ we denote the set of all paths starting at $s\in S$.

B. RTECTL: an existential fragment of a soft real-time temporal logic.

In the syntax of RTECTL we assume the following: $p\in PV$ is an atomic proposition, and I is an interval in $\mathbb{N}=\{0,1,2,\ldots\}$ of the form: $[a,b)$ or $[a,\infty)$, for $a,b\in N$ and $a\neq b$. The RTECTL formulae are defined by the following grammar:

$$\varphi::=true\,|\,false\,|\,p\,|\,\neg p\,|\,\varphi\wedge\varphi\,|\,\varphi\vee\varphi\,|\,EX\varphi\,|\,E(\varphi U_I\varphi)\,|\,E(\varphi R_I\varphi).$$

Intuitively, we have an existential path quantifier E, and the symbols X, U_I, and R_I that are the temporal operators for "neXt time", "bounded until", and "bounded release", respectively. The formula $E(\alpha U_I\beta)$ means that it is possible to reach a state satisfying β via a finite path whose cumulative duration is in I, and always earlier α holds. The formula $E(\alpha R_I\beta)$ means that either it is possible to reach a state satisfying α and β via a finite path whose cumulative duration is in I, and always earlier β holds, or there is a path along which β holds at all states with cumulative duration being in I. The formulae for the "bounded eventually", and "bounded always" are defined as standard: $EF_I\varphi\stackrel{\text{def}}{=}E(true\,U_I\varphi)$, $EG_I\varphi\stackrel{\text{def}}{=}E(false\,R_I\varphi)$.

An RTECTL formula φ is true in the model M (in symbols $M\models\varphi$) iff $M,\iota\models\varphi$ (i.e., φ is true at the initial state of the model M). For every $s\in S$ the relation $\models$ is defined inductively as follows:

- $M,s\models\alpha\wedge\beta$ iff $M,s\models\alpha$ and $M,s\models\beta$
- $M,s\models\alpha\vee\beta$ iff $M,s\models\alpha$ or $M,s\models\beta$
- $M,s\models EX\alpha$ iff $(\exists\pi\in\Pi(s))(M,\pi(1)\models\alpha)$
- $M,s\models E(\alpha U_I\beta)$ iff $(\exists\pi\in\Pi(s))(\exists m\ge 0)$ $(D\pi[0..m]\in I$ and $M,\pi(m)\models\beta$ and $(\forall\, j<m)M,\pi(j)\models\alpha)$
- $M,s\models E(\alpha R_I\beta)$ iff $(\exists\pi\in\Pi(s))(\exists m\ge 0)$ $(D\pi[0..m]\in I$ and $M,\pi(m)\models\alpha$ and $(\forall\, j\le m)M,\pi(j)\models\beta)$ or $(\forall\, m\ge 0)D\pi[0..m]\in I$ implies $M,\pi(m)\models\beta$.

III. SMT-based bounded model checking

In this section we define the SMT-based BMC method for the existential fragment of RTCTL (RTECTL) [9]. Similarly to SAT-based BMC, the SMT-based BMC is based on the notion of the bounded semantics for RTECTL ([10]) in which one inductively defines for every $s\in S$ the relation $\models_k$. Let M be a model, $k\ge 0$ a bound, φ an RTECTL formula, and let $M,s\models_k\varphi$ denotes that φ is k-true at the state s of M. The formula φ is k-true in M (in symbols $M\models_k\varphi$) iff $M,\iota\models_k\varphi$ (i.e., φ is k-true at the initial state of the model M).

A. Bounded Semantics for RTECTL

Let M be a model, $k\ge 0$ a bound, φ an RTECTL formula, and $M,s\models_k\varphi$ denote that φ is k-true at the state s of M.

The formula φ is k-true in M (in symbols $M\models_k\varphi$) iff $M,\iota\models_k\varphi$ (i.e., φ is k-true at the initial state of the model M).

For every $s\in S$, the relation $\models_k$ (the bounded semantics) is defined inductively as follows:

- $M,s\models_k EX\alpha$ iff $k>0$ and $(\exists\pi\in\Pi_k(s))(M,\pi(1)\models_k\alpha)$
- $M,s\models_k E(\alpha U_I\beta)$ iff $(\exists\pi\in\Pi_k(s))(\exists 0\le m\le k)$ $(D\pi[0..m]\in I$ and $M,\pi(m)\models_k\beta$ and $(\forall\, 0\le j<m)M,\pi(j)\models_k\alpha)$
- $M,s\models_k E(\alpha R_I\beta)$ iff $(\exists\pi\in\Pi_k(s))(\exists 0\le m\le k)$ $(D\pi[0..m]\in I$ and $M,\pi(m)\models\alpha$ and $(\forall\, 0\le j\le m)M,\pi(j)\models_k\beta)$ or $(D\pi[0..k]\ge\text{right}(I)$ and $(\forall\, 0\le j\le k)$ $D\pi[0..j]\in I$ implies$)$ or $(\forall\, m\ge 0)D\pi[0..m]\in I$ implies $M,\pi(m)\models\beta$

Theorem 1 ([10]). *Let* M *be a model and an RTECTL formula. Then, the following equivalence holds:* $M \models \varphi$ *iff there exists* $k \geq 0$ *such that* $M \models_k \varphi$.

The bounded model checking problem asks whether there exists $k \in N$ such that $M \models_k \varphi$. The following theorem states that for a given model and an RTECTL formula there exists a bound k such that the model checking problem ($M \models_k \varphi$) can be reduced to the BMC problem ($M \models_k \varphi$). The theorem can be proven by induction on the length of the formula φ.

B. Translation to SMT

The translation to SMT is based on the bounded semantics. Let M M be a simply-timed model, φ an RTECTL formula, and k a bound. The presented SMT encoding of the BMC problem for RTECTL and for STS is based on the SAT encoding of the same problem [10,11], and it relies on defining the quantifier-free first-order formula

$$[M,\varphi]_k := [M^{\varphi,\iota}]_k \wedge [\varphi]_{M,k}$$

that is satisfiable if and only if $M \models_k \varphi$ holds.

The definition of the formula $[M^{\varphi,\iota}]_k$assumes that states of the model M are encoded in a symbolic way. Such a symbolic encoding is possible, since the set of states of M is finite. In particular, each state s can be represented by a vector w (called a *symbolic state*) of different individual variables ranging over the natural numbers (called *individual state variables*).

The formula $[M^{\varphi,\iota}]_k$ encodes a rooted tree of k -paths of the model M. The number of branches of the tree depends on the value of the auxiliary function $f_\iota : RTECTL \to \mathbb{N}$ defined in [10].

Given the above, the j -th symbolic k -path is defined as the following sequence $((d_{0,j}, w_{0,j}), \ldots, (d_{k,j}, w_{k,j}))$, where $w_{i,j}$ are symbolic states and $d_{i,j}$ are *symbolic durations*, for $0 \leq i \leq k$ and $0 \leq j \leq f_k(\varphi)$. The *symbolic duration* $d_{i,j}$ is a individual variable ranging over natural numbers.

Let w and w' (resp., d and d') be two different symbolic states (resp., durations). We assume definitions of the following auxiliary quantifier-free first-order formulae:

- $I_\iota(w)$ - encodes the initial state of the model M,
- $T((d,w),(d',w'))$ - encodes the transition relation of M,
- $p(w)$ - encodes the set of states of M in which $p \in PV$ holds,
- $B_k^I(\pi_n)$ encodes that the duration time represented by the sequence $d_{1,n}, \ldots, d_{k,n}$ of symbolic durations is less than right(I),
- $D_j^I(\pi_n)$ - encodes that the duration time represented by the sequence $d_{1,n}, \ldots, d_{j,n}$ of symbolic durations belongs to the interval I,
- $D_{k,l,m}^I(\pi_n)$ for $l \leq m$ - encodes that the duration time represented by the sequences $d_{1,n}, \ldots, d_{k,n}$ and $d_{l+1,n}, \ldots, d_{m,n}$ of symbolic durations belongs to the interval I.

The formula $[M^{\varphi,\iota}]_k$ encoding the unfolding of the transition relation of the model M $f_k(\varphi)$-times to the depth k is defined as follows:

$$[M^{\varphi,\iota}]_k = I_i(w_{0,0}) \wedge \bigwedge_{j=0}^{f_k(\varphi)-1} \bigwedge_{i=0}^{k-1} T((d_{i,j}, w_{i,j}), (d_{i+1,j}, w_{i+1,j}))$$

For every RTECTL formula φ the function f_k determines how many symbolic k -paths are needed for translating the formula φ. Given a formula φ and a set A of k -paths such that $|A| = f_k(\varphi)$, we divide the set A into subsets needed for translating the subformulae of φ. To accomplish this goal we need the auxiliary functions $g_s(A)$, $h_n^U(A,e)$ and $h_n^R(A,e)$ that were defined in [11].

Let φ be an RTECTL formula, M a model, and $k \in \mathbb{N}$ a bound. The quantifier-free first-order formula $[\varphi]_{M,k} := [\varphi]_k^{[0,0,F_k(\varphi)]}$, where $F_k(\varphi) = \{j \in N \mid 0 \leq j < f_k(\varphi)\}$, encodes the bounded semantics for RTECTL, and It is defined inductively as shown below. Namely, let $0 \leq n < f_k(\varphi)$, $m \leq k$, $n' = min(A)$, $h_k^U = h_k^U(A, f_k(\beta))$, $h_k^R = h_k^R(A, f_k(\alpha))$, then:

- $[EX\,\alpha]_k^{m,n,A} := w_{m,n} = w_{0,n'} \wedge [\alpha]_k^{[1,n',g_s(A)]}$ if $k > 0$; *false*, *otherwise*,
- $[E(\alpha U_I \beta)]_k^{m,n,A} := w_{m,n} = w_{0,n'} \wedge \bigvee_{i=0}^{k} \left([\beta]_k^{[i,n',h_k^U(k)]} \wedge D_i^I(\pi_{n'}) \wedge \bigwedge_{j=0}^{i-1} [\alpha]_k^{[j,n',h_k^R(j)]} \right)$,
- $[E(\alpha R_I \beta)]_k^{m,n,A} := w_{m,n} = w_{0,n'} \wedge \bigvee_{i=0}^{k} \left([\alpha]_k^{[i,n',h_k^R(k+1)]} \wedge D_i^I(\pi_{n'}) \wedge \bigwedge_{j=0}^{i} [\beta]_k^{[j,n',h_k^R(j)]} \right) \vee \left(\neg B_k^I(\pi_{n'}) \wedge \bigwedge_{j=0}^{k} \left(D_j^I(\pi_{n'}) \rightarrow [\beta]_k^{[j,n',h_k^R(j)]} \right) \right) \vee \left(B_k^I(\pi_{n'}) \wedge \bigwedge_{j=0}^{k} \left(D_j^I(\pi_{n'}) \rightarrow [\beta]_k^{[j,n',h_k^R(j)]} \right) \wedge \bigvee_{l=0}^{k-1} \left[w_{k,n'} = w_{l,n'} \right] \wedge \bigwedge_{l=0}^{k-1} \left(D_{k;l,j+1}^I(\pi_{n'}) \rightarrow [\beta]_k^{[j,n',h_k^R(j)]} \right) \right) \right)$.

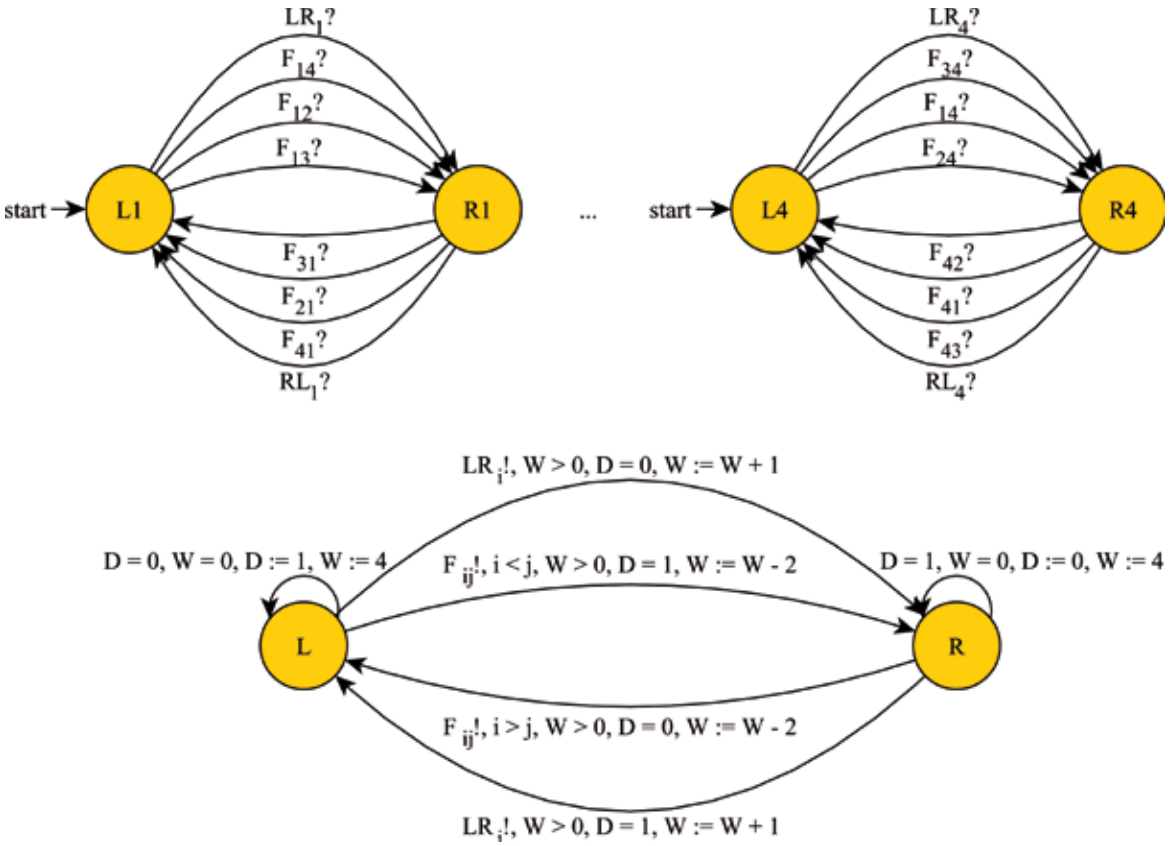


Figure 1: A network of STADD automata that models BCP for 4 persons. The variableD indicates the crossing direction: D = 1 (D = 0) means that all the persons crossthe bridge from its left side to its right side, (from its right side to its left side). Thevariable W denotes the number of persons waiting on the left (right) side of the bridge,if D = 1 (D = 0).

IV. Experimental results

Our SAT-based and SMT-based BMC algorithms were implemented as standalone programs written in the programming language C++. For the SAT-based BMC module we used the state of the art SAT-solver PicoSAT [4], and for our SMT-based BMC module we used the state of the art SMT-solver Z3 [6].

In this section we experimentally evaluate the performance of our SMT-based BMC encoding for RTECTL over the STS se- mantics. We compare our experimental results with the SAT-based BMC [10], the only existing method that is suitable with respect to the input formalism and checked properties.

We have conducted the experiments using two benchmarks: the generic simply-time pipeline paradigm (GSPP) STS model [10] and the bridge crossing problem (BCP) STS model [10]. We would like to point out that both benchmarks are very useful and scalable examples. Further, we specify each property for the considered benchmarks in the existential form, and for every specification given, there exists a witness.

We have computed our experimental results on a computer equipped with I7-3770 processor, 32 GB of RAM, and the operating system Arch Linux with the kernel 3.15.3. We set the CPU time limit to 3600 seconds. Moreover, in order to compare our SMT-based BMC with the SAT-based BMC, we have asked the authors of [10] to provide us the binary version of their implementation of the SAT-based BMC method. We have obtained the requested binaries. Furthermore, our SMT-based BMC algorithm is implemented as standalone program written in the programming language C++.

For the SAT-based BMC module we used the state of the art SAT-solver PicoSAT (http://fmv.jku.at/picosat/) [4], and for our SMT-based BMC module we used the state of the art SMT- solver Z3 [6] (http://z3.codeplex.com/).

A. The bridge-crossing problem

The bridge-crossing problem (BCP) [8] is a famous mathematical puzzle. To generate experimental results we have tested BCP system defined in [10]. We have five automata that run in parallel and synchronised on actions LR_i, RL_i, and F_{ij} for $i \neq j$ and $i, j \in \{1,...4\}$. The action LR_i (respectively, RL_i) means that the i-th person goes from the left side of the bridge to its right side (respectively, from the right side of the bridge to its left side) bringing back the lamp. The F_{ij} action with $i<j$ (respectively, F_{ij} with $i>j$) means that the persons i and j cross the bridge together from its left side to its right side (respectively, from its right side to its left side). Four automata (those with states named as Li and Ri, for $1 \leq i \leq 4$) represent persons, and one represents a lamp that keeps track of the position of the lamp, and ensures that at most two persons cross in one move. Let Min denote the minimum time required to cross the bridge, $N_p \geq 2$ be the number of persons, and $\text{right} = (2 \cdot N_p - 3) \cdot (t_1 + (N_p - 1) \cdot 3)$. We have tested BCP for $N_p \geq 4$ persons, $t_i - t_1 + (i-1) \cdot 3$ with $1 \leq i \leq N - p$ and $t_1 \geq 10$, on the following RTECTL formulae:

- $\varphi_{1\,BCP} = EF_{|Min, Min+1|}\left(\wedge_{i=1}^{N_p} Ri\right)$,
- $\varphi_{2\,BCP} = EG_{|0,\text{right}|}\left(\vee_{i=1}^{N_p} \neg Ri\right)$;

the formulae are true in the model for BCP.

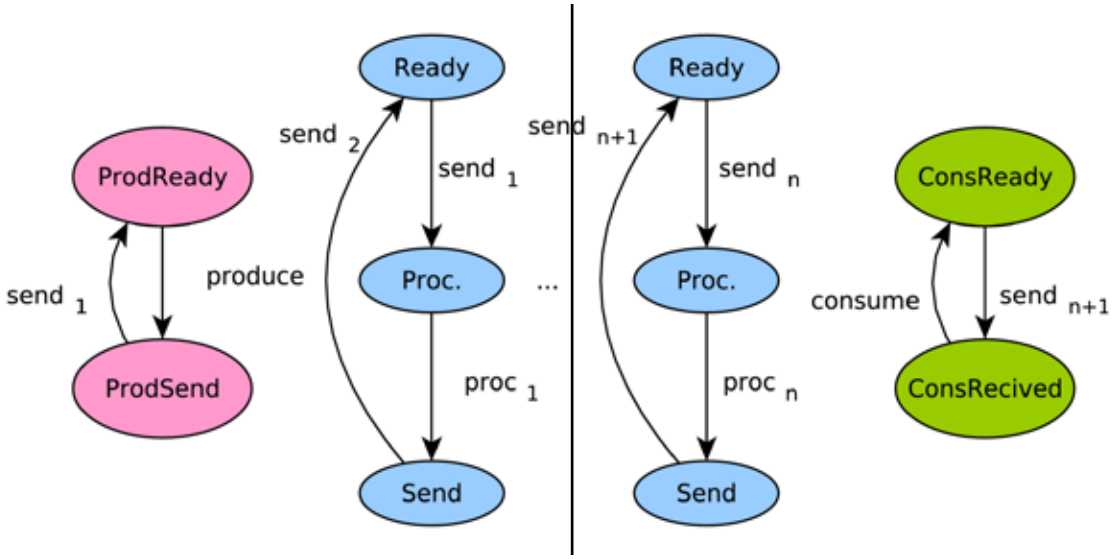


Figure 2: A network of STADD automata that models GSPP.

B. Generic Simply-timed Pipeline Paradigm

We adapted the benchmark scenario of *a generic pipeline paradigm* [7], and we called it the *generic simply-timed pipeline paradigm* (GSPP). The model of GSPP involves Producer producing data, Consumer receiving data, and a chain of n intermediate Nodes that transmit data produced by Producer to Consumer. Producer, Nodes, and Consumer have different producing, sending, processing, and consuming times. A STADD automata model of GSPP is shown in Fig 2. We have $n+2$ automata (n automata representing Nodes, one automaton for Producer, and one automaton for Consumer) that run in parallel and synchronise on actions $Send_i$ ($1 \le i \le n+1$). Action $Send_i$ $1 \le i \le n$ means that i -th Node has received data produced by Producer. Action $Proc_i$ $1 \le i \le n$ means that Consumer has received data produced by Producer.

Action $Proc_i$ $1 \le i \le n$ means that i -th Node processes data. Action $Produce$ means that Producer generates data. Action $Consume$ means that Consumer consumes data produced by Producer.

Let $1 \le i \le n$. We have tested the GSPP problem with the following basic durations: $d(Produce)=2$, $d(send_i)=2$, $d(Proc_i)=4$, $d(Consume)=2$ and their multiplications by 50, 100, 150, etc., on the following RTECTL formulae:

- $\varphi_{1GSPP} = EF_{[Min, Min+1]} ConsReceived$,
- $\varphi_{2GSPP} = EG_{[0,\infty)}(\neg ProdSend \vee EF_{[0, Min-d(Produce)+1]} ConsReceived)$,
- $\varphi_{3GSPP} = EG_{[0,\infty)}(\neg ProdSend \vee EG_{[0, Min-d(Produce)]} ConsReady)$,

where Min denotes the minimum time required to receive by Consumer the data produced by Producer.

Note that the φ_{2GSPP} and φ_{3GSPP} are properties, respectively, of the type the existential *bounded-response* and existential *bounded-invariance*. All the above formulae are true in the model for GSPP.

C. Performance evaluation

The evaluation of both the BMC algorithms is given by means of the running time and the memory used. In most cases, the experimental results show that the SMT-based BMC method is significantly faster than the SAT-based BMC method.

1) GSPP

From Fig. 3-8 and Tables 1-3 we can notice that for the GSPP system and all considered formulae the SMT-based BMC is faster than the SAT-based BMC, however, the SAT-based BMC consumes less memory. Moreover, the SMT-based method is able to verify more nodes for all the tested formulae. In particular, in the time limit set for the benchmarks, the SMT-based BMC is able to verify the formula φ_{1GSPP} for 54 nodes while the SAT-based BMC can handle 40 nodes, for the formula φ_{2GSPP} respectively 25 nodes and 21 nodes. For φ_{3GSPP} the SMT-based BMC is still more efficient - it is able to verify 20 nodes, whereas the SAT-based BMC verifies only 17 nodes for $t=1$ and 19 nodes for $t=1000000$.

TABLE 1: SMT-BMC: EXPERIMENTAL RESULTS FOR GSPP AND φ_{1GSPP} SCALING UP THE NUMBER OF NODES AND BASIC DURATION

n	Sec.	MB	Whitness length
1	0.1	12.5	5
5	0.1	13.5	13
10	0.7	16.6	23
15	3.0	21.8	33
20	8.4	29.7	43
25	21.4	40.3	53
30	56.3	50.7	63
35	139.0	70.6	73
40	323.0	86.3	83
45	577.9	109.8	93
50	1615.1	145.0	103
53	3220.1	153.5	109
54	3957.4	180.8	111

TABLE 2: SMT-BMC: EXPERIMENTAL RESULTS FOR GSPP AND φ_{2GSPP} SCALING UP THE NUMBER OF NODES AND BASIC DURATION

n	Sec.	MB	Whitness length
1	0.6	4.4	6
2	1.9	7.2	8
3	3.4	11.2	10
4	6.1	15.9	12
5	10.9	22.2	14
6	24.6	29.9	16
7	31.8	38.9	18
8	50.0	50.4	20
9	79.1	63.4	22
10	98.7	78.7	24
12	170.1	120.3	28
15	615.5	198.2	34
20	2304.1	423.5	44
22	2462.1	537.3	48
25	7203.7	777.5	54

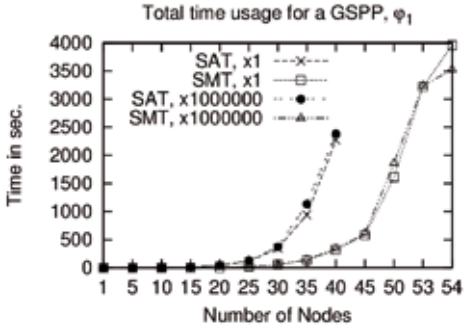
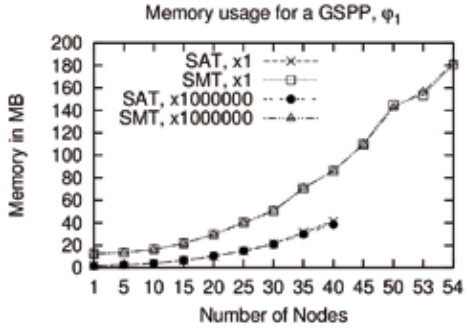


Figure 3: φ_{1GSPP} SAT/SMT-BMC: GSPP scaling up both the number of nodes and durations



4: φ_{1GSPP} SAT/SMT-BMC: GSPP scaling up both the number of nodes and durations

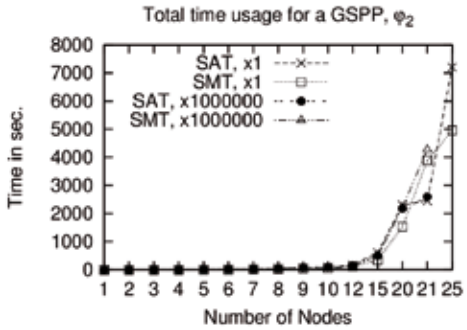


Figure 5:SAT/SMT-BMC: GSPP scaling up both the number of nodes and durations

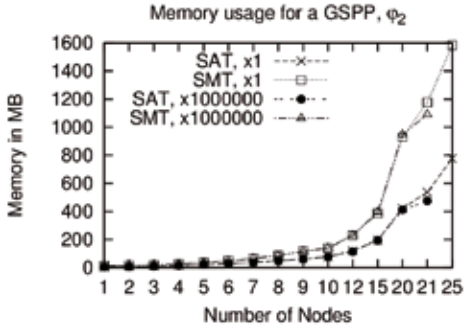


Figure 6: φ_{2GSPP} SAT/SMT-BMC: GSPP scaling up both the number of nodes and durations

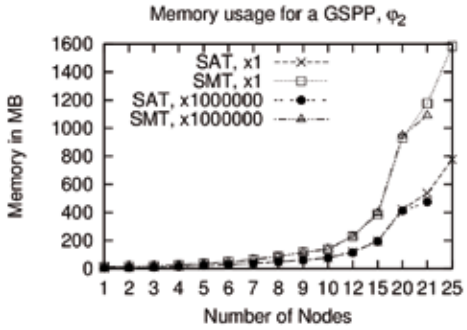


Figure 7: φ_{3GSPP} SAT/SMT-BMC: GSPP scaling up both the number of nodes and durations

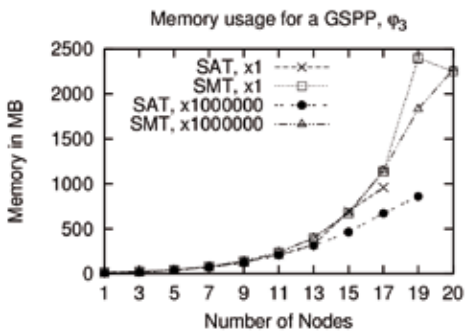


Figure 8: φ_{3GSPP} SAT/SMT-BMC: GSPP scaling up both the number of nodes and durations

TABLE 3: SMT-BMC: EXPERIMENTAL RESULTS FOR GSPP AND φ_{3GSPP} SCALING UP THE NUMBER OF NODES AND BASIC DURATION

n	Sec.	MB	Whitness length
1	0.2	14.4	5
2	0.7	17.7	7
3	1.4	23.6	9
4	2.7	31.8	11
5	4.8	43.6	13
6	8.4	62.6	15
7	16.9	81.6	17
8	25.8	103.3	19
9	37.7	138.1	21
10	78.0	200.0	23
11	90.8	232.3	25
12	143.0	301.1	27
13	203.3	393.9	29
14	407.9	503.8	31
15	402.0	670.1	33
16	675.1	1014.5	35
17	762.1	1136.4	37
18	1663.9	1901.3	39
19	2235.8	2397.2	41
20	3641.5	2252.4	43

2) BCP

As one can see from the line charts for the BCP system (Figures 9-12, Tables 4-5), in the case of this benchmark the SMT-based BMC and SAT-based BMC are complementary. In the case of the formula SMT-based BMC is able to verify system with 10 persons while the SAT-based BMC can handle 11 persons. For the SMT-based BMC is more efficient - it is able to verify 31 persons, whereas the SAT-based BMC verifies only 27 nodes for and 29 nodes for , but the SAT-based BMC consumes less memory.

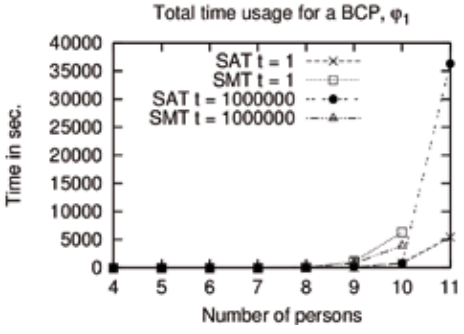


Figure 9: φ_{1BCP} SAT/SMT BMC: BCP scaling up both the number of persons and durations

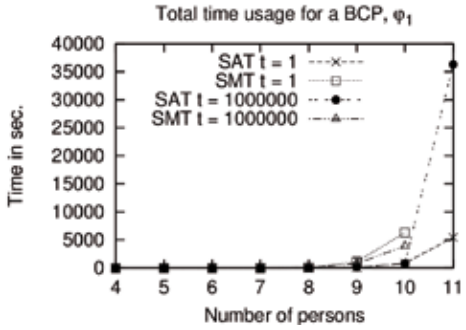


Figure 10: φ_{1BCP} SAT/SMT BMC: BCP scaling up both the number of persons and durations

TABLE 4: SMT-BMC: EXPERIMENTAL RESULTS FOR BCP AND φ_{1BCP} SCALING UP THE NUMBER OF NODES AND BASIC DURATION

n	Sec.	MB	Whitness length
4	0.0	13.6	6
5	0.1	14.4	8
6	0.8	16.4	10
7	4.8	21.3	12
8	67.5	44.3	14
9	1094.3	158.8	16
10	6298.2	163.6	18

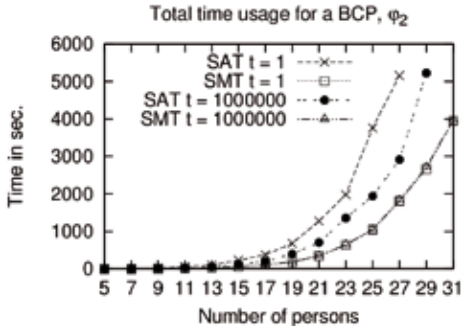


Figure 11: φ_{2BCP} SAT/SMT BMC: BCP scaling up both the number of persons and durations

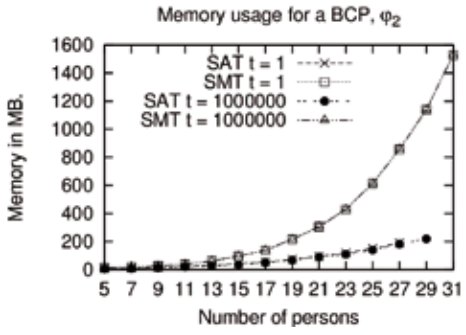


Figure 12: φ_{2BCP} SAT/SMT BMC: BCP scaling up both the number of persons and durations

TABLE 5: SMT-BMC: EXPERIMENTAL RESULTS FOR BCP AND φ_{2BCP} SCALING UP THE NUMBER OF NODES AND BASIC DURATION

n	Sec.	MB	Whitness length
4	0.1	13.7	6
5	0.2	14.6	8
6	0.5	16.4	10
7	1.2	19.3	12
8	2.1	22.4	14
9	3.7	27.3	16
10	5.9	33.1	18
11	9.9	39.7	20
12	14.9	49.0	22
13	25.9	61.7	24
14	34.1	73.8	26
15	54.4	96.5	28
16	71.7	114.3	30
17	105.3	138.4	32
18	141.7	172.2	34
19	182.1	216.8	36
20	265.5	251.8	38
21	340.1	305.9	40
22	436.0	367.3	42
23	630.3	428.9	44
24	816.6	521.3	46
25	1037.9	614.6	48
26	1309.7	734.7	50
27	1804.1	857.4	52
28	2385.1	986.7	54
29	2666.8	1139.2	56
30	3527.2	1326.0	58
31	3946.9	1528.0	60

V. Conclusion

We have proposed an SMT-based BMC verification method for model checking RTECTL properties interpreted over the simply-time systems that are generated for simply-timed automata with discrete data. We have provided a preliminary experimental results showing that our method is worth interest.

References

[1] R. Alur. Timed Automata. In *Proceedings of CAV'99*, volume 1633 of *LNCS*, pages 8-22. Springer-Verlag, 1999.

[2] R. Alur, C. Courcoubetis, and D. Dill. Model checking in dense real-time. *Information and Computation*, 104(1):2-34, 1993.

[3] Clark Barrett, Roberto Sebastiani, Sanjit Seshia, and Cesare Tinelli. Satisfiability modulo theories. In Armin Biere, Marijn J. H. Heule, Hans van Maaren, and Toby Walsh, editors, *Handbook of Satisfiability*, volume 185 of *Frontiers in Artificial Intelligence and Applications*, chapter 26, pages 825-885. IOS Press, 2009.

[4] A. Biere. PicoSAT essentials. *Journal on Satisfiability, Boolean Modeling and Computation (JSAT)*, 4:75-97, 2008.

[5] N. Markey and Ph. Schnoebelen. Symbolic model checking of simply-timed systems. In *Proceedings of FORMATS'04 and FTRTFT'04*, volume 3253 of *LNCS*, pages 102-117. Springer, 2004.

[6] L. De Moura and N. Bjørner. Z3: an efficient SMT solver. In *Proceedings of TACAS'2008*, volume 4963 of *LNCS*, pages 337-340. Springer-Verlag, 2008.

[7] D. Peled. All from one, one for all: On model checking using representatives. In *Proceedings of CAV'93*, volume 697 of *LNCS*, pages 409-423. Springer-Verlag, 1993.

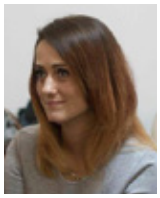
[8] Elizabeth Early Cook Saul X. Levmore. *Super Strategies for Puzzles and Games*. Doubleday, Garden City, N.Y., 1981.

[9] B. Woźna-Szcześniak, A. M. Zbrzezny, and A. Zbrzezny. The BMC method for the existential part of RTCTLK and interleaved interpreted systems. In *Proceedings of EPIA'2011*, volume 7026 of *LNAI*, pages 551-565. Springer-Verlag, 2011.

[10] B. Woźna-Szcześniak, A. M. Zbrzezny, and A. Zbrzezny. SAT-based bounded model checking for RTECTL and simply-timed systems. In *Proceedings of EPEW 2013*, volume 8168 of *LNCS*, pages 337-349. Springer-Verlag, 2013.

[11] A. Zbrzezny. Improving the translation from ECTL to SAT. *Fundamenta Informaticae*, 85(1-4):513-531, 2008.

[12] A. Zbrzezny and A. Pólrola. SAT-based reachability checking for timed automata with discrete data. *Fundamenta Informaticae*, 79(3-4):579-593, 2007.



Agnieszka M. Zbrzezny received the M.S. in Computer Science from Jan Długosz University in Częstocjowa, now she is finishing the doctorate in the Institute of Computer Science Polish Acadamy of Sciences. Her research interest include model checking of multi-agent system and real-time systems.

Andrzej Zbrzezny received M.S. deegree in computer science from Jagielonian University in Cracow in 1981, PhD deegre in logic from Wrocław University in 1991, and habilitation in computer science from Polish Academy of Sciences in Warsaw in 2014. Now he is an associate professor at Jan Dlugosz University in Częstochowa. His research interest include model checking of concurrent systems, multi-agent system and real-time systems.

"Wrapping" X3DOM around Web Audio API

Andreas Stamoulias[1], Eftychia Lakka[1,2], Athanasios G. Malamos[1]

[1]Department of Informatics Engineering Technological Educational Institute of Crete Heraklion, Greece, GR 71004,
[2]Faculty of Computing, Engineering and Science, University of South Wales, Treforest, UK

***Abstract* — Spatial sound has a conceptual role in the Web3D environments, due to highly realism scenes that can provide. Lately the efforts are concentrated on the extension of the X3D/ X3DOM through spatial sound attributes. This paper presents a novel method for the introduction of spatial sound components in the X3DOM framework, based on X3D specification and Web Audio API. The proposed method incorporates the introduction of enhanced sound nodes for X3DOM which are derived by the implementation of the X3D standard components, enriched with accessional features of Web Audio API. Moreover, several examples-scenarios developed for the evaluation of our approach. The implemented examples established the achievability of new registered nodes in X3DOM, for spatial sound characteristics in Web3D virtual worlds.**

***Keywords* — Spatial sound, X3D, X3DOM, Web Audio API, Web3D, Real-time, Realistic 3D, 3D Audio.**

I. Introduction

THE three Dimensional (3D) visualization plays a vital role in Computer Graphics. In the last few years, there is a growing interest in the technologies which have been designed to present Web 3D scenes. This effort has started since 1995, when a text based meta-language, Virtual Reality Modeling Language (VRML), was design with the influence of HTML [1]. Following this, the Web Graphics Library (WebGL) introduced in 2011 by the Kronos Group [2] and attracted considerable attention by Eicke et al. [3], due to the effective suggested method for Web 3D representation, without any particular hardware and software requirements. The so-called extensible 3D (X3D), is one of the latest architectures that has come up as an extension of VRML [4]. The use of the Extensible Markup Language (XML) in X3D had as a consequence increased flexibility and maturity in comparison to VRML [5]. However, the drawback of plugins installation was remaining. This problem was solved by the proposed X3DOM framework, as reported by Behr et al. [6]. For this reason, the last five years almost 70% of the X3D specification has been realized into the X3DOM framework, but the spatial sound attributes are remaining an open issue and not implemented yet, despite the fact that it is a critical issue for the 3D scene.

Moreover, the literature on 3D sound shows that it has attracted much attention from research teams. The group of Garbe [7] and Ding et al. [8], through the enrichment of their X3D audio nodes, succeed in better Web 3D representation. In the meantime, the most interesting approach to this issue has been proposed by the Web Audio API. This describes a high-level JavaScript API [9] that provides a group of new and reinforced audio properties in Web applications [10]. Wyse et al [11] and Pettersson et al. [12] have developed a novel implementation by using Web Audio API in order to achieve their goals for spatial sound. Most of the 3D applications support spatial sound. For example, the position and the intensity of a sound source could be managed by the user and accordingly, could be changed depending on the movement of the user. All the above are important for realistic scene. Consequently, spatial sound is an equally decisive issue for both, the 3D and the Web 3D applications.

However, while X3D ISO [13] includes the specification of the spatial sound components, it has not been implemented in X3DOM yet. Moreover, there is no efficient implementation of spatial sound in X3D in general. Besides that, few publications can be available in the literature that address the issue of the X3D sound extension [7] [14] but they work only with a very limited number of sophisticated sound properties and methods. Therefore, we demonstrate an innovative method to improve the spatial auralization in X3DOM, through the X3D specification and the structure of Web Audio API. In detail, we make an effort to combine the benefits of X3DOM, such as the open technology and the lack of plugins use, with the flexible HTML5 and the efficient Web Audio API. Accordingly, our work intends to include 3D sound in a declarative web 3D scene, with the ability to run natively. Likewise, the paper does not only focus on the implementation but also on the evaluation through a variety of scenarios.

The rest of the paper is organized in 4 sections: Section II describes the technologies that are used, Section III analyses the implementation. The experimental results of the evaluation are presented in Section IV and Section V concludes the paper.

II. Background

This section presents background information for the X3D/X3DOM and for the 3D Audio (Web 3D Audio - Web Audio API - X3D/X3DOM Audio) as well.

A. Web 3D - X3D/X3DOM

The X3D has been proposed by Web3D consortium, a nonprofit organization which is a combination of business companies, government agencies, academic institutions, and individual professionals. It has been formally approved by the International Standards Organization (ISO) as ISO/IEC 19775, since 2004 [15]. X3D is modern descendant of VRML, it not only includes the capabilities of VRML, but also predominates in several aspects. Particularly, it uses XML in order to express the geometry and integrates plainly with other applications [16]. Moreover, it is organized into logical groupings of functionality (components) [17], which could be expanded and enriched with new ones. Likewise, X3D applications are reliable and predictable; similarly X3D binary format provides encryption and compression [18]. For the aforementioned reasons, the X3D has become more attractive and effective than the VRML.

On one hand, X3D has been used already in HTML5 in order to declare the web 3D worlds. On the other hand, it does not address the problem of connection between the web-browser frontends and the X3D backends. X3DOM comes to overcome this issue [19].

DOI: 10.9781/ijimai.2015.355

TABLE I
OUR IMPLEMENTATION FOR THE REGISTRATION OF X3D SOUND NODES (ATTRIBUTES) INTO X3DOM.

Node	X3D	X3DOM	OUR IMPLEMENTATION	
	Attributes		Web Audio API	Enhanced X3D Nodes
X3DSoundSourceNode: AudioClip	SFString description			AudioSource
	SFBool loop	✔	AudioBufferSourceNode	AudioSource
	SFNode metadata	✔		AudioSource
	SFTime pauseTime		AudioBuffer AudioContext	AudioSource
	SFFloat pitch		AudioBufferSourceNode	AudioSource
	SFTime resumeTime		AudioBuffer AudioContext	AudioSource
	SFTime startTime		AudioBufferSourceNode	AudioSource
	SFTime stopTime		AudioBufferSourceNode	AudioSource
	MFString url	✔		AudioSource
	SFTime duration_changed			
	SFTime elapsedTime			
	SFBool isActive			
	SFBool isPaused			
	✘	SFBool enabled		
X3DSoundNode: Sound	SFVec3f location		PannerNode	PannerNode
	SFVec3f direction		PannerNode	PannerNode
	SFFloat maxBack		PannerNode	PannerNode
	SFFloat maxFront		PannerNode	PannerNode
	SFFloat minBack		PannerNode	PannerNode
	SFFloat minFront		PannerNode	PannerNode
	SFNode metadata	✔		
	SFFloat intensity		GainNode	
	SFFloat priority		DynamicsCompressorNode	
	SFNode source	✔		AudioSource
	SFBool spatialize		TRUE	TRUE

It is a framework, which incorporates a set of technologies such as X3D, WebGL, HTML, CSS and JavaScript. The main objective is the development of an X3D scene by the use of the HTML DOM and the 3D content management through the DOM elements and without extra plugins [20]. Equally important is the capability of X3DOM, as a JavaScript framework, to allow three interactions; the first is an event in the scene that causes a behavior in the scene. The second is an event in the scene that causes a behavior in the HTML5. The last is an event in HTML5 that causes a behavior in scene [21].

The present research efforts are focusing on the development of Web 3D interactive virtual worlds, based on the X3DOM. However, some studies draw attention to the X3DOM node extension. Stamoulias et al. [22] suggest the implementation of the rigid body physics component in the X3DOM environment. A concept for the improvement of the shadow representation for X3DOM is provided by the Kuijper's group [3]. Additionally, Kapetanakis et. al [23] present an approach to extend the adaptation methods of X3DOM by adding a mechanism to perform dynamic adaptation and achieve HD video delivery in 3D Virtual Reality (VR) worlds. All the above have the same aim; to increase the level of realism in the Web 3D scene, through the enhancement of the X3DOM structure. The aim of the current work, in accordance to the previous examples, is to open up the field of the spatial sound in a web 3D scene, beyond the limits of the X3D-X3DOM technology.

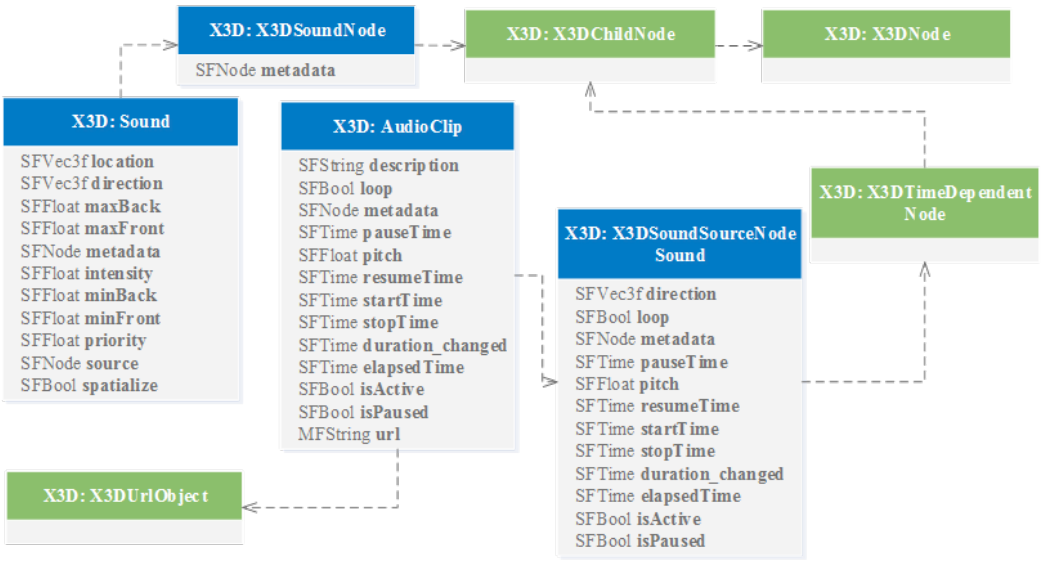


Fig. 1. Inheritance diagram of Audio nodes in X3D.

B. 3D Audio

The sound comprises a fundamental part of real life applications and can be produced from a plenty of audio sources. Every sound has a specific direction and can be easily identified, due to its distinctive characterizations and the familiarity that we have developed with sounds [24]. Additionally, a characteristic of the human hearing system is the ability to perceive the 3D sound. General, the meaning of 3D sound lies in the way that the listener receives the incoming sounds from all the directions. The concept of 3D sound is the same when it is simulated by a computer [25]. In other words, a listener can recognize meaningful spatial cues from a sound source, for example the direction, the distance and the spaciousness [26].

According to the above, the 3D sound is important to be included in virtual worlds, in order to improve the realism and the sense of the immersivity in the scene. In this case, the user can understand any sound source which is included in the virtual 3D scene through the combination of visual and auditory sense. Adding natural spatial sound to interactive 3D immersive applications, more sophisticated information is conveyed, in conjunction with using other modalities (e.g., vision) [27].

1) Web 3D Audio

Many approaches for Web 3D applications aimed at realistic visualization of the scene. The most often overlooked point is the sound, even though if it can offer further details to a 3D graphic world. Specifically, a high immersion level is accomplished and the natural interaction is increased, on the ground of that the graphic scene simulates the real world in the best possible way [23]. Equally important is the spatial audio and the rendering of spatial attributes of each auditory objects, in a Web 3D world. These attributes involve perceived directions, distances and spatial extends of the auditory objects; furthermore these attributes should be conceived in the same way as they are recognized in the real world [28].

For all the above reasons, considerable attention has been paid to introduce sound in web application. The first attempt took place via the

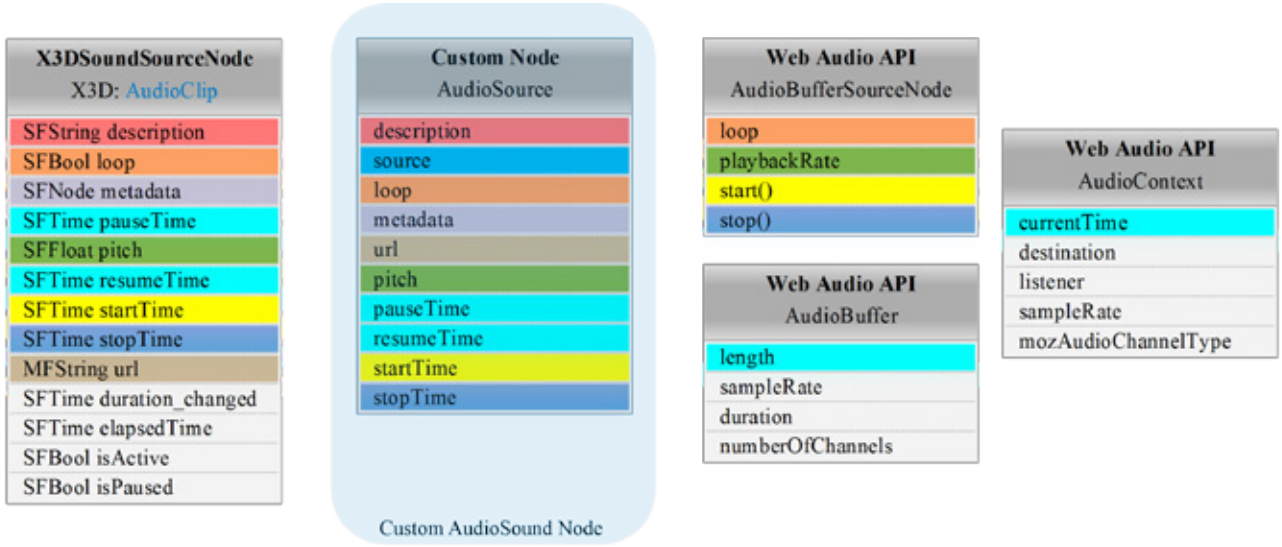


Fig. 2a. This component diagram presents our proposal for spatial sound in X3DOM. The attributes of AudioClip (X3D node) have been registered in X3DOM with the use of new enhanced X3D nodes and Web Audio API nodes. Each color expresses which fields combine in order to create the respective components in X3DOM.

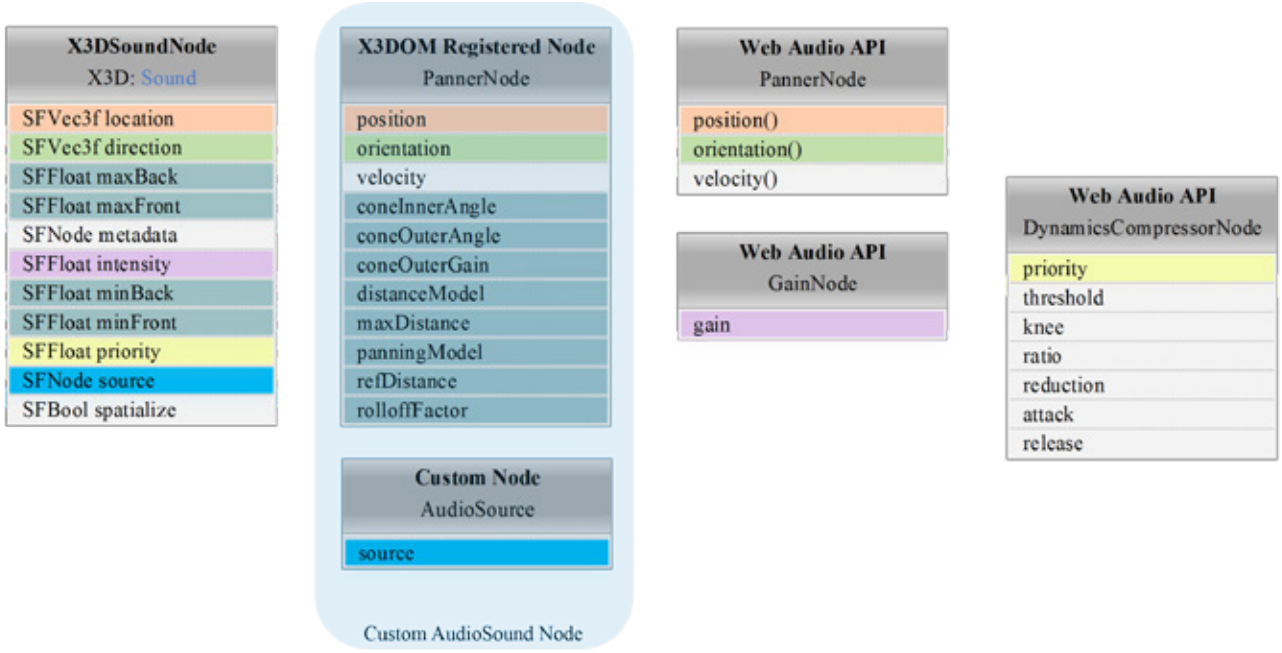


Fig. 2b. This component diagram presents our proposal for spatial sound in X3DOM. The attributes of Sound (X3D node) have been registered in X3DOM with the use of new enhanced X3D nodes and Web Audio API nodes. Each color expresses which fields combine in order to create the respective components in X3DOM.

<bgsound> tag, in which only background music could be contained in a web page and was available for specific browsers. After that, flash was the first cross-browser way of audio on the Web, but plugins were required. The focus of recent research was concentrated on the element <audio> in HTML5, which could avoid the plugins, but was not designed for sophisticated and ambitious developments [29], [30]. Specifically, the element <audio> is inferior to apply filters to the sound signal and access the raw PCM data. Moreover, it does not include the concept of position and direction of sources and listeners. Lastly, it does not afford low-latency precise-timing model, which is very important for interactive applications for the need of fast auditory response to user actions [31]. Thus, it is not adequate for a 3D interactive web environment with demanding sound design.

Under these circumstances, several alternatives have been proposed, in order to establish an effective API, which attends to overcome the most of these limitations. One of the most interesting approach to this issue is Web Audio API, which has been proposed by Mozilla Foundation. Predecessor of Web Audio API was the Audio Data API. This API provided a distinct structure for writing audio callbacks in JavaScript, but it did not provide specifications for lower level, native, pre-compiled agent to be included in browsers [32]. As a result, Mozilla Foundation resulted in the adoption of the Web Audio API, which was endorsed by the other browsers.

2) Web Audio API

Web Audio API is a high-level JavaScript API, which can be used to synthesize audio in web applications. Additionally, it is characterized as an extremely powerful tool for controlling audio in the browser and tends to become a de facto standard in modern browsers [33].

Until now, 3D audio rendering engines were utilized, such as OpenAL (Open Audio Library) [34], FMOD library, in order to develop 3D auralization. Even thought that the existing solutions afford friendly interfaces, they have a number of limitations. Namely, they confined to the rendering of static audio sources, in contrast with the Web Audio API [35].

Moreover, the significance of Web Audio API can be obvious from the fact that a number of web audio libraries and APIs have been developed, in order to use/handle it. Specifically, Three.js, uses the Web Audio API to play the sound and determine the correct volume. It is a JavaScript library, which offers a simple way programming WebGL directly from JavaScript, with the view to create and animate 3D scenes [36], [37]. Moreover, the webaudiox.js library disposes a set of helpers for using the Web Audio API and the howler.js library supports automatic caching for Web Audio API [38], [39]. Furthermore, the pedalboard.js is an open-source JavaScript framework which develops audio effects through the Web Audio API [40]. Another one is the Wad library that helps to simplify manipulating audio using the Web Audio API [41]. Lastly, the Fifer Javascript library is a lightweight conductor for the Web Audio API with Flash Fallback. [42].

As mentioned above, the Web Audio API becomes attractive for the spatial audio in web environments. Except the previous reasons, it also distinguished for extra benefits. In detail, it is open source and be supported from the most browsers. Furthermore, multi-channel audio is available and is integrated with Web Real-Time Communications (WebRTC). Also, high-level sound abilities as filters, delay lines, amplifiers, spatial effects (such as panning) are offered. At the same time, audio channels can have 3D distribution according to the position, speed or direction of the viewer and the sound source [43]. Additionally, the Web Audio API is characterized by compositionality, using audio node structure, which can be linked together in order to form an audio routing graph. Besides that, it is much faster since it is written in C++, rather than it has been written in JavaScript. However, Web Audio API also provides a node (ScriptProcessorNode) that allows the web developers to manage audio using JavaScript. All the above assets make the Web Audio API to break new ground for the web sound synthesis [10].

Indeed, it was not a coincidence that Web Audio API has drawn much attention from research teams in the last two years. In particular, the literature demonstrates a variety of approaches which utilize Web Audio API, in order to accomplish the sound in web environment [44]-[47]. Furthermore, many researchers have proposed various methods of adjusting the Web Audio API in their application [48], [49], [50].

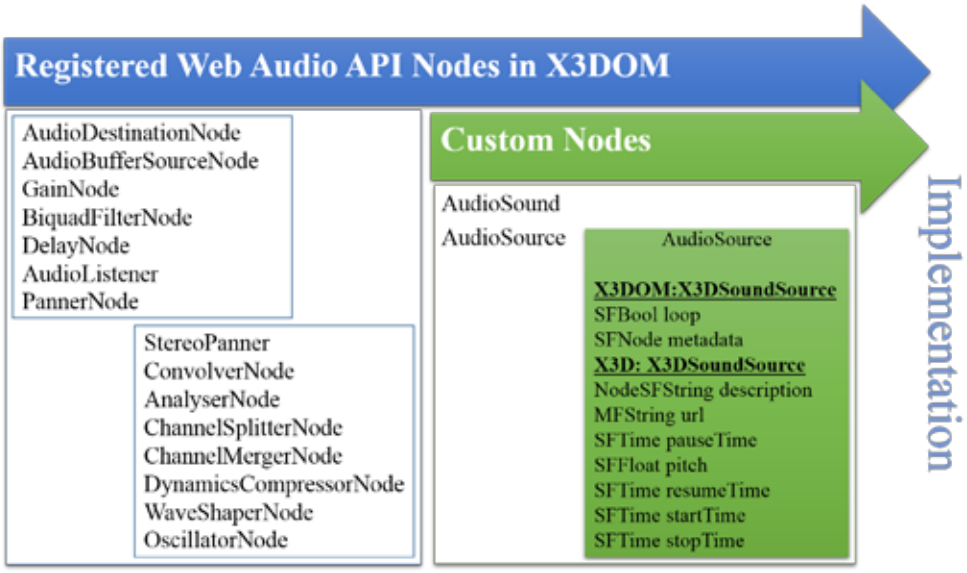


Fig. 3. Registered nodes for the paper implementation

In either case, the approach of Web Audio API is based on the concept of audio context, which presents the direction of audio stream flows, between sound nodes. In every node, the properties of the sound can be adapted and changed, depending on the application requirements.

The included nodes are possible to be sorted by type: a) Source nodes (audio buffers, live audio inputs, oscillators and JS processors), b) Modification nodes (filters, convolvers and panners), c) Analysis nodes (analyzers and JS processors), d) Destination nodes (audio outputs and offline processing buffers) [29].

3) X3D/X3DOM Audio

The current literature shows that insufficient efforts have been done in order to incorporate the spatial sound in X3D. Particularly, two types of nodes are included in the X3D, the first is about the sound description and the other one for the sound source. Specifically, the first node is the X3DSoundNode, which is an abstract node for all sound nodes. It is minimalist with only one attribute, metadata, which expresses important information for the significance, appearance and the proposed role of the model [51]. The second node is the X3DSoundSourceNode, which is the abstract node for each node that is used to emit sound and it has a number of common fields with the TimeSensor, for example the loop, the startTime, the stopTime, the pauseTime and resumeTime (Figure 1).

The third node is the Sound, which is derived from the X3DSoundNode. It is designed for the description of the X3D scene sounds. Specifically, it determines both the location and the behavior of the sound. Additionally, the geometry describes that the sound can be directed and be emitted in an elliptical pattern. Two ellipsoids constitute the pattern, which specifies the borders for level of loudness of the sound. Also, ellipsoids can be reshaped in order to provide more or less directional focus from the location of the sound [52]. Consequently, the sound node is intended to recognize the source and is related to the direction, the location, the priority and general, the spatial features of the sound source (Figure 1) [53].

The forth node is the AudioClip, which is derived from the X3DSoundSourceNode. It specifies audio data that can be referenced by Sound nodes. Basically, it loads an external audio file with a view to handle playing, stopping and starting. As regard the attributes, AudioClip has a number of fields in common with TimeSensor, because it is an X3DSoundSourceNode and implements the X3DTimeDependentNode abstract type. Basically, the fields of the sound nodes and their interrelation are presented in Figure 1 by an interpretive diagram.

Besides that the X3DOM is a descendant of X3D, the only thing that it has been implemented for sound is the registration of X3DSoundNode, X3DSoundSourceNode, AudioClip and Sound nodes, but without the most of the properties of the respective X3D nodes. Specifically, an X3DOM scene could include an audio file for playing, without any spatial characteristics. Only the attribute “SFBool enabled” has been extra added to X3DOM in comparison with X3D, which specifies whether the clip is enabled or not. The first two columns of Table 1 illustrate the attributes of sound nodes that are transferred to X3DOM from X3D.

C. Motivation

The X3DOM framework comprises a fundamental part in the 3D web development, because it provides an approach for the integration of declarative 3D in HTML5 [19]. However, the implementation of spatial sound in X3DOM is still lacking, even thought that the spatial sound should be an integral part of an immersive 3D application. In order to overcome this drawback, we present an innovative solution of the spatial sound in X3DOM framework that based on a combinational methodology. Specifically, we suggested the enrichment of X3DOM with spatial sound features, using both the X3D sound nodes and the structure of Web Audio API. We selected this combination for the reason that both the X3D and the Web Audio API has been influenced by OpenAL. Particularly, the Web Audio API has borrowed many concepts from OpenAL (position and orientation of sources and listeners, parameters associated with the source audio cones, relative velocities of sources and listeners) [29]. In the same time, the structure of OpenAL has been used for X3D sound nodes extension, by the group of Garbe [7] and several authors [54], [55], [18] have proposed the implementation of 3D virtual environment with the cooperation of X3D (for 3D scene) and OpenAL (for 3D sound). The compatibility which stems from the common ground that the API and X3D has, is the main asset that we are taking advantage in order to incorporate the spatial sound to X3DOM framework.

A measurement of the impact of our contribution is the number of web applications in X3DOM platform that may benefit from the

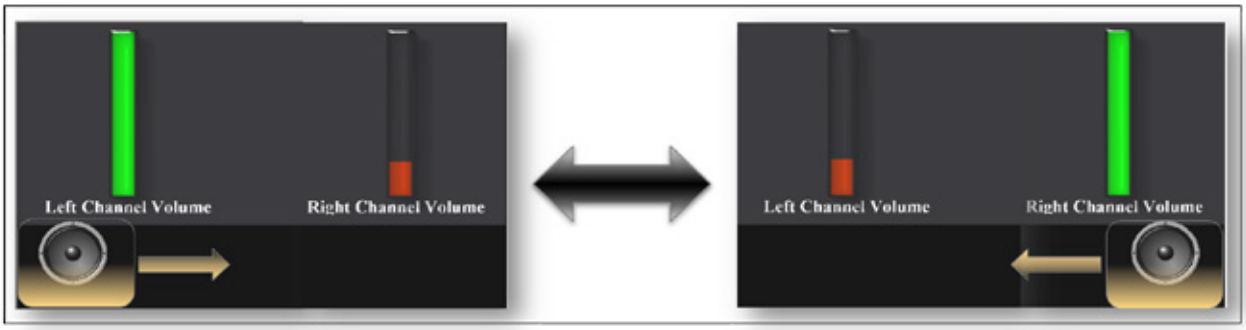


Fig. 4. The second example: spatial sound with a sound source which moves in the scene (right-left).

spatial sound components we introduce. Teleconferencing, gaming, immersive internet and entertainment are some of the areas that achieve a truly immersive listening experience, through 3D sound of X3DOM.

Although Web Audio API has been coupled with WebGL for spatial sound in the past [36], these efforts are custom implementations that use API's methods natively and are still away of becoming a platform or even more a language. Our effort was not just to customize an environment but instead to fully integrate Web Audio API into X3D language, by introducing new sound nodes and incorporate them in X3DOM API. In other words, with this approach the complex sound design and implementation is becoming transparent to the programmer and moreover the applications are independent of the sound libraries are employed.

Fig 3 presents the new nodes which have been registered in X3DOM according to our proposal (a combination of Web Audio API nodes and Custom Nodes). Table 1, Fig. 2a and Fig. 2b depict the matching of X3D and Web Audio API attributes, which have been integrated in order to create the respective components in X3DOM.

III. Implementation

This Section is devoted to the implementation of the proposed design. In the first place is the registration of Web Audio API and custom components into X3DOM framework. All these nodes are used by HTML DOM, which is essentially an X3D scene, directly from the X3DOM and no through JavaScript structure. On the other hand, JavaScript code, which is indicated as JavaScript Controller, accords all the indispensable instruction of nodes design and role. It interacts with the HTML file, for the purpose of parsing the 3D scene and being updated on any potential change in the scene. In the following sections, a more detailed description is presented of which nodes are entered on the X3DOM core in order to be recognized and be used as X3DOM element, as well as the process of JavaScript Controller.

A. Web Audio API/Enhanced X3D nodes registration into X3DOM

The attributes of AudioClip (X3D node) and Sound (X3D node) have been re-introduced in X3DOM with the use of:

- Two enhanced X3D nodes
  1. AudioSound
  2. AudioSource
- An added value set of Web Audio API nodes.

The first stage of our work is the registration of new components in X3DOM core. Besides that, it mutually exchanges information with the AudioBufferSourceNode component of Web Audio API. For visual representation of the implementation structure, the reader is referred to Fig 3.

1) Registered Web Audio API nodes in X3DOM

A set of Web Audio API nodes were registered in the X3DOM, in order to implement the proposed work. To enumerate, we utilized the AudioDestinationNode and AudioBufferSourceNode, which represent the final audio destination and the audio source consisting of in-memory audio data. Likewise, the GainNode affects the loudness of a sound and the BiquadFilterNode supports all of the commonly used second-order filter types. Furthermore, the DelayNode causes a delay between the arrival of an input data and its propagation to the output. The AudioListener represents the position of a person listening to an audio source in 3D space and each source can be passed through a PannerNode, which spatializes the input audio. Based on the relative position of the sources and the listener, the correct gain modifications can be computed by this method. Also, we registered the StereoPannerNode in order to pan an audio stream left or right. A further register node is the ConvolverNode which is effectively a very complex filter (like the BiquadFilterNode), but rather than selecting from a set of effect types, it can be configured with an arbitrary filter response. The AnalyserNode is intended to provide real-time frequency and time-domain analysis information. The ChannelSplitterNode was utilized to separate the different channels of an audio source and it often used in conjunction with its opposite the ChannelMergerNode. Lastly, we recommended to insert the DynamicsCompressorNode which was used for the compression of effects, the WaveShaperNode for the representation of a non-linear distorter and the OscillatorNode for the rendition of a periodic waveform [9].

It is noteworthy that the nodes StereoPannerNode, ConvolverNode, WaveShaperNode and OscillatorNode have been registered as extra nodes, in order to provide further sound features in the web 3D scene and succeed a higher dimension of realism in the 3D environment.

2) Enhanced X3D nodes

The enhanced X3D nodes, which were developed for the needs of the proposed work, are the AudioSound and AudioSource. The first one is the "parent" element of any other new component and the second is a combination of the development node in X3DOM, X3DSoundSourceNode and set of attributes of the corresponding X3D node. Specifically, the registration of AudioSound has been developed with the follow structure:

Registration of AudioSound Node in X3DOM

```
x3dom.registerNodeType("AudioSound"){
this.addField_SFNode("transform",x3dom.nodeTypes.Transform),
this.addField_SFNode("source", x3dom.nodeTypes.AudioSource),
this.addField_SFNode("panner", x3dom.nodeTypes.PannerNode),
this.addField_SFNode("filter",x3dom.nodeTypes.BiquadFilterNode),
```

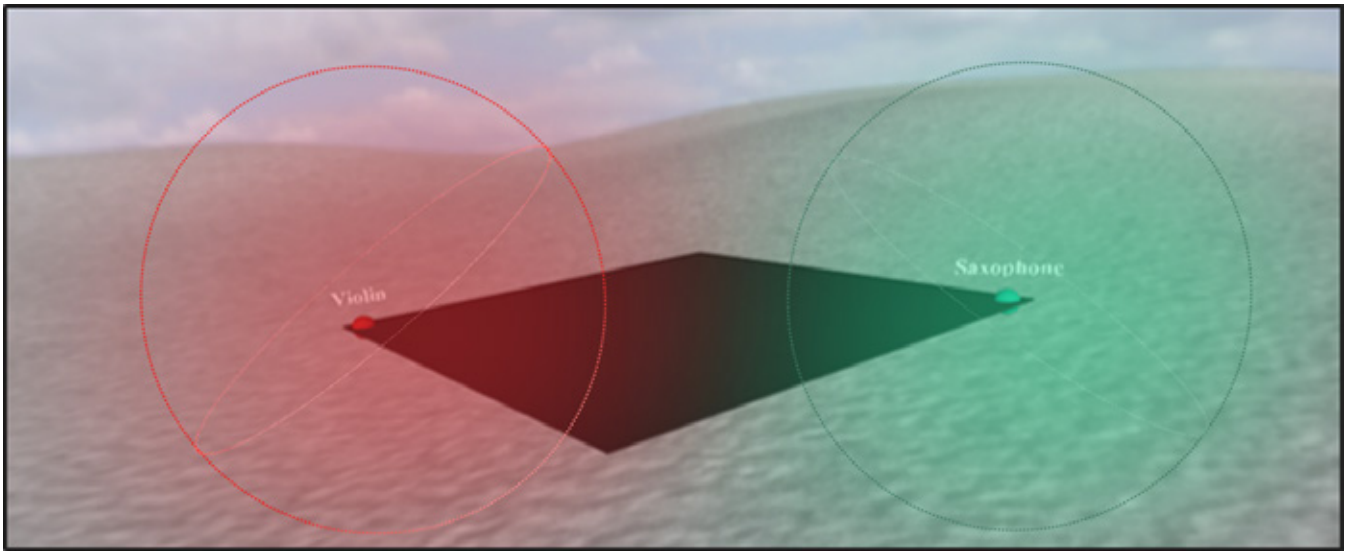


Fig. 5. Forth example: spatial sound through camera animation in an X3D scene with two sound sources.

```
    this.addField_SFNode("delay", x3dom.nodeTypes.DelayNode),
    this.addField_SFFloat("playbackRate", 1),
    this.addField_SFNode("metadata",x3dom.nodeTypes.X3DMetadataObject)
}
```

As it is mentioned above, the AudioSound includes transform field which is linked with the Transform type node, in order to connect them using the DEF/USE attribute. Additionally, the source attribute is to determine the sound source, through the AudioSource enhanced X3D node. Moreover, the panner attribute links with the PannerNode, which is reliable for the representation of audio source position and behavior. After that, the filter is intended to insert a simple low-order filter. Lastly, the attributes delay and playbackRate are used for delay and the rate of input sound data.

Furthermore, AudioSource have been registered with the follow structure:

Registration of AudioSource Node in X3DOM

```
x3dom.registerNodeType("AudioSource"){
this.addField_SFString("description", ""),
this.addField_SFBool("loop", !1),
this.addField_SFNode("metadata",
x3dom.nodeTypes.X3DMetadataObject),
this.addField_MFString("url", []),
this.addField_SFTime("pauseTime", 0),
this.addField_SFFloat("pitch", 1),
this.addField_SFTime("resumeTime", 0),
this.addField_SFTime("startTime", 0),
this.addField_SFTime("stopTime", 0) }
```

The description field specifies a textual description of the audio source, the loop attribute arranges for the repetition of sound source. The url field points to the location of the sound source of interest and the pitch is a multiplication factor applied to sound sampling and playback. Lastly, in X3D structure (pauseTime, startTime, stopTime) had the defaults value (0, 0, 0). In our implementation, the respective attributes was implemented with the default values (-1, 0, -1), for the case that pauseTime and stopTime do not need to be used.

B. JavaScript Controller

The role of JavaScript Controller is to emulate the ScriptProcessorNode node of Web Audio API, which provides the ability of web audio synthesis and process, directly in JavaScript. Equally important is the fact that one or more AudioSound components, which incorporates AudioSource, PannerNode, BiquadFilterNode, DelayNode and Transform nodes, could be included in HTML. Accordingly, the structure of JavaScript Controller starts with a repetitive detection in the HTML DOM, until find an AudioSound node (or an Inline node. In this case, the scene is an X3D file and all nodes are loaded by Inline node via the "url" of X3D file.), essentially until recognize an X3DOM audio element. Once this being achieved, a new 3D Sound Object will be created. Then, registered nodes are invoked, following certain order, AudioSourceNode, AnalyserNode, WaveShaperNode, BiquadFilterNode, ConvolverNode, GainNode, PannerNode and AudioDestinationNode. Those of the above nodes will be recognized, they will be appended in an array, in order to be available. Together with that, 3D Sound Object was initialized and was updated, through the exchange of data with HTML DOM. In addition to, ArrayBuffer and XMLHttpRequest are used, for the purpose of to load and play sound. In case that, visual and aural conditions will be changed, HTML DOM will be respectively updated.

IV. Evaluation

In this section the results of our implementation are given. In order to verify the validity of our method, we carried out several experiments – examples, in which new registered nodes were used (http://medialab.teicrete.gr/minipages/x3domAudio/index.html). Our tests have been investigated with the use of Google Chrome 41.0, Firefox Mozilla 36.0 and Opera 28.0. This choice was based on the fact that these browsers support the Web Audio API [56]. However, in our development has anticipated the fact that the browser cannot support spatial sound for any reason (for example not support Web Audio API). In this case, the implementation was adapted and produces the result without sound spatiality. Additionally, the browser Opera cannot load .mp3 files, in this case, our work controls this condition and if it is true, returns warning message. Next, a description is following, which presents these examples in detail.

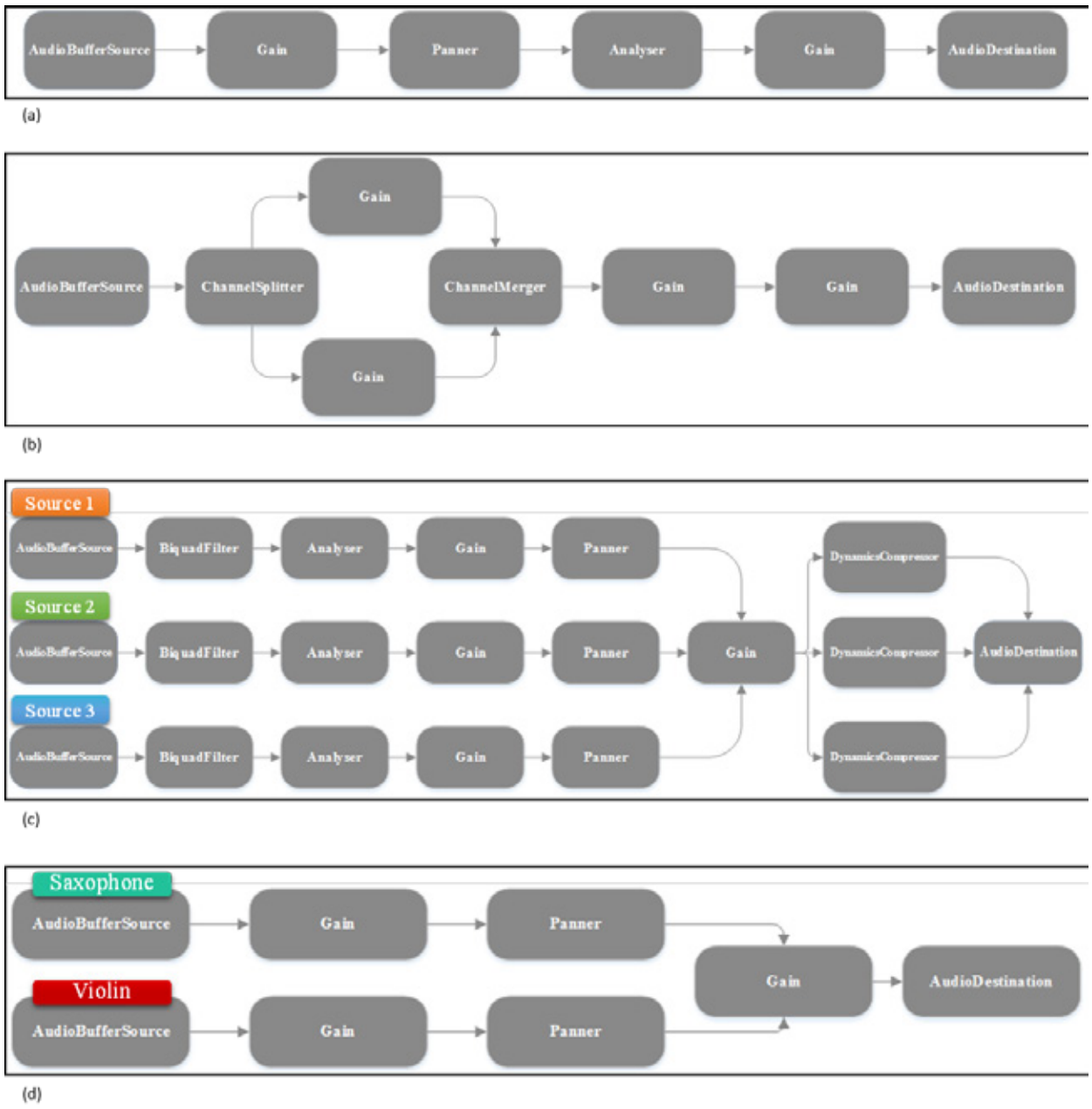


Diagram 1: (a) The Node Diagram of Single Audio Source Example (First Example). (b) The Node Diagram of Split Channels Example (Second Example). (c) The Node Diagram of Spatial Sound Effects and Filters Example (Third Example). (d) The Node Diagram of Web3D Spatial Audio Camera Animation Example (Forth Example).

A. Examples-Experiments

1) Single Audio Source

The first example1 evaluates the attenuation of one sound source, in a web 3D scene, through the X3DOM. Specifically, it includes the implementation of a single sound source, which is represented by a 3D object. The spatiality of the sound is expressed by a process, in which when the user approaching nearby to the sound source the volume is increased and accordingly when removed therefrom is reduced. In addition to this and depending on the side of the sound source that the user observes, the sound is emitted from the corresponding speaker. Apart from the 3D scene, we have also added an analyser slider. The analyser gives the possibility to receive real-time generated data, without any change from the input to output sound information. Through this process we achieved the audio visualization of the sound source.

In order to achieve what is described in the first example, a subset of new nodes (registered by us in X3DOM framework) were utilized in HTML DOM. Specifically, AudioSound is used as the parent element of any other new node. Transform, PannerNode and AudioSource constitute the children of AudioSound. The role of Transform node is to offer multiple DEF/USE copies of X3DOM in the new registered AudioSound node.

Moreover, PannerNode spatializes the input audio based on the relative position of the source and the listener. Finally, the AudioSource is responsible to load the sound file.

The Diagram 1a summarizes the background of implementation. Particularly, the graph illustrates how the nodes are connected and in

1http://www.medialab.teicrete.gr/minipages/x3domAudio/singleAudio.xhtml

which sequence they are implemented for this specific example.

2) Split Channels

The second example2 assesses the capability of the audio channels split, through our implementation (Fig 4). This X3D scene includes a simple sound source which can be moved right and left. Depending on the position of the sound source, the user can hear the produced sound from the corresponding output speaker. In this case, the new registered sound nodes of the X3DOM are called directly from the HTML file. In particular, we started with the AudioSound, which is the parent element of any other new node, as mentioned earlier. Transform, PannerNode and AudioSource which can represent different kinds of filters, comprise the children of the AudioSound. Accordingly, there is a source that can be passed through a PannerNode for the spatialization of the input audio.

Also, Diagram 1b outlines the nodes connection of the second example through the implementation.

3) Spatial Sound Effects and Filters

In the same way, a third example3 is investigated during the evaluation process, in order to introduce effects and filters in a X3DOM environment. This example includes an X3D scene with three sound sources. Each of them is visualized by a 3D object (in our case is a sphere) that depicts the sound effects. Specifically, we have added filters through of them we are able to manage the different sound effects in an impressive way. Filters can be composed of a number of attributes, frequency, detune, gain and the factor quality which also known as Q.

Basically, the filters are classified in some specific types, depending on the sound effects that produce. In detail, there is the Low-pass filter which can create more muffled sound. Another one is the High-pass filter, which is used to generate tinny sound. Equally important is the Band-pass filter, which cuts off low and high frequencies and passes through only these within a certain range. On the contrary, the Notch filter has exactly the opposite operation of the Band-pass filter. Then is the Low-shelf filter, its role is to change the amount of bass in a sound, as a result the frequencies that are lower than the current frequency get a boost, while them that are over it remain unchanged. Next, the High-shelf filter is responsible for the quantity of treble in a sound. Moreover, Peaking filter is used in order to handle the amount of midrange in a sound. Lastly, there is the All-pass filter, whose role is to introduce phased effects.

In order to implement the said to our example it has been utilized the nodes of the previous example in a similar way (the new registered sound nodes of the X3DOM are called directly from the HTML file). Consequently an AudioSound node for each sound source is used. AudioSound is comprised of Transform, PannerNode, AudioSource and BiquadFilterNode. Also, an analytical menu is provided in the web page, so the user can change the parameters of the filters which are described above.

The Diagram 1c shows in which way the registered nodes in X3DOM created and used in this scenario-example.

4) Web3D Spatial Audio Camera Animation

In the last example4 (Figure 5), we evaluate the attenuation of two different sound sources, while the camera (the user) is moving in the 3D scene. Through the immersion in the X3D scene the user could attend a rational navigation. Whenever the camera moves in the direction of an existing sound source, the sound strength of this source increases, while the sound strength of the other (the second one) decreases and vice versa. Through this process, great realism of the scene is achieved, since it emulates the spatial sound in real world.

In a like manner as in the previous examples, the new registered sound nodes of the X3DOM are called directly from the HTML file. Therefore, an AudioSound node (which contains Transform, PannerNode and AudioSource, as its children) for each sound source is used.

Furthermore, Diagram 1d represents the new registered nodes flow of this example.

Consequently, the experiments indicated that our implementation corresponded as it was expected, without any major problem. Furthermore, some tests repeated with the use of a considerable number of sound sources at the same time, with effectual results.

2 http://www.medialab.teicrete.gr/minipages/x3domAudio/splitChannels.xhtml
3 http://www.medialab.teicrete.gr/minipages/x3domAudio/filters.xhtml
4 http://www.medialab.teicrete.gr/minipages/x3domAudio/spatialAudioCamera.xhtml

V. Conclusions

Even though the X3DOM advantages in comparison with X3D, it does not require plugins, the X3DOM could not adequately handle the spatial sound. This was a major drawback for the design of realistic and interactive 3D scenes. Our proposal is a sufficient way to solve this problem and adds the spatial sound in X3DOM, ensuring that the quality of 3D scenes will be increased.

Substantially, the implementation of already existing X3D sound nodes was insufficient. For this reason, we included the structure of Web Audio API nodes and we adapted them in the X3DOM framework.

Lastly, based on the examples-experiments, it can be concluded that interactive Web3D scene can be composed with the use of new registered nodes and the results confirm our methodology.

References

[1] Raggett, D. (1995). Extending WWW to support platform independent virtual reality. Techical Report.

[2] KHRONOS GROUP. Retrieved from https://www.khronos.org/news/press/khronos-releases-final-webgl-1.0-specification [accessed May 2015].

[3] Eicke, T. N., Jung, Y., & Kuijper, A. (2015). Stable dynamic webshadows in the X3DOM framework. Expert Systems with Applications. Retrieved from http://dx.doi.org/10.1016/j.eswa.2014.11.059

[4] Daly, L., & Brutzman, D. (2007). X3D: Extensible 3D Graphics Standard. 24, pp. 130 - 135. IEEE. doi:10.1109/MSP.2007.905889

[5] McIntosh, P., Hamilton, M., & Schyndel, R. v. (2005). X3D-UML: Enabling Advanced UML Visualisation Through X3D. Association for Computing Machinery, Inc.

[6] Behr, J., Jung, Y., Keil, J., Drevensek, T., Zoellner, M., Eschler, P., & Fellner, D. (2010). A Scalable Architecture for the HTML5/ X3D Integration Model X3DOM. Web3D '10 Proceedings of the 15th International Conference on Web 3D Technology. doi:10.1145/1836049.1836077

[7] Garbe, K., & Herbst, I. (2008). Extending X3D with Perceptual Auditory Properties. Virtual Reality Conference, 2008. VR '08. IEEE. doi:10.1109/VR.2008.4480787

[8] Ding, H., Schwarz, D., Jacquemin, C., & Cahen, R. (2011). Spatial audio: graphic modeling for X3D. Web3D '11 Proceedings of the 16th International Conference on 3D Web Technology. doi:10.1145/2010425.2010427

[9] W3C. (2013). Web Audio API. Retrieved from http://www.w3.org/TR/webaudio/ [accessed May 2015].

[10] Wyse, L. (2014). Interactive Audio Web Development Workflow. ACM. doi:10.1145/2647868.2655064

[11] Wyse, L., & Subramanian, S. (2013). The viability of the web browser as a computer music platform (Vol. 37). (C. M. Journal, Ed.)

[12] Dibra, D., Otero, N., & Pettersson, O. (2014). Real-Time Interactive Visualization Aiding Pronunciation of English as a Second Language. IEEE. doi:10.1109/ICALT.2014.131

[13] Extensible 3D (X3D) ISO/IEC 19775-1:2008. Retrieved from http://www.

web3d.org/documents/specifications/19775-1/V3.2/Part01/Architecture.html [accessed May 2015]
[14] Ding, H., Schwarz, D., Jacquemin, C., Cahen, R. (2011). Spatial audio: graphic modeling for X3D. Web3D '11 Proceedings of the 16th International Conference on 3D Web Technology
[15] Parisi, T. (2014). Programming 3D Applications with HTML5 and WebGL. O'Reilly Media, Inc
[16] Spala, P., Malamos, A. G., Doulamis , A., & Mamakis, G. (2011). Extending MPEG-7 for efficient annotation of complex web 3D scenes (Vol. 59). Multimedia Tools and Applications, Springer US. doi:10.1007/s11042-011-0790-5
[17] Stewart, J. A., Dumoulin, S. J., & Noël, S. (2006). X3D Conformance Testing Factors in Creating Aviable Test Suite. Electrical and Computer Engineering, 2006. CCECE '06. doi:10.1109/CCECE.2006.277506
[18] Bouras , C., Panagopoulos , A., & Tsiatsos, T. (2005). Advances in X3D multi-user virtual environments. Multimedia, Seventh IEEE International Symposium. doi:10.1109/ISM.2005.28
[19] Behr, J., Eschler, P., Jung, Y., & Zollner, M. (2009). X3DOM: a DOM-based HTML5/X3D integration model. Web3D '09 Proceedings of the 14th International Conference on 3D Web Technology. doi:10.1145/1559764.1559784
[20] X3DOM. Retrieved from http://www.x3dom.org/ [accessed May 2015]
[21] Baglivo, A., Ponti, F. D., Luca, D. D., & Fanini, B. (2013). X3D/X3DOM, Blender Game Engine and OSG4WEB: open source visualisation for cultural heritage environments. Digital Heritage International Congress (DigitalHeritage).
[22] Stamoulias, A., Malamos, A. G., Zampoglou, M., & Brutzman, D. (2014). Enhancing X3DOM declarative 3D with rigid body physics support. Proceedings of the Nineteenth International ACM Conference on 3D Web Technologies. ACM.
[23] Mora-Lumbreras , M. A., Flores-Pulido, L., González-Contreras, B. M., & Portilla , A. (2012). Incorporating 3D Sound in Different Virtual Worlds. Web Congress (LA-WEB). doi:10.1109/LA-WEB.2012.17
[24] Kapetanakis, K., Panagiotakis, S., Malamos, A. G., & Zampoglou, M. (2014). Adaptive Video Streaming on top of Web3D A bridging technology between X3DOM and MPEG-DASH. Telecommunications and Multimedia (TEMU), 2014 International Conference, IEEE. doi:10.1109/TEMU.2014.6917765
[25] Wu, J.-R., Duh, C.-D., Ouhyoung, M., & Wu, J.-T. (1997). Head Motion and Latency Compensation on. Lausanne Switzerland: ACM VRST '97.
[26] Lee, H. C., Kim, H. B., Lee , M. J., Kim , P. M., Suh, S. W., & Kim, K. H. (1998). Development of 3D sound generation system for multimedia application. Computer Human Interaction. IEEE.
[27] Cowan, B., & Kapralos, B. (2013). Spatial sound rendering for dynamic virtual environments. Digital Signal Processing (DSP). IEEE.
[28] Laitinen, M. V., Pihlajamäki, T., Erkut, C., & Pulkki, V. (2012). Parametric time-frequency representation of spatial sound in virtual worlds. ACM Transactions on Applied Perception (TAP). doi:10.1145/2207216.2207219
[29] Smus, B. (2013). Web Audio API. O'Reilly Media, Inc
[30] Spuy, R. v. (2015). Sound with the Web Audio API. In Advanced Game Design with HTML5 and JavaScript. IEEE.
[31] Smus, B. Retrieved from Developing Game Audio with the Web Audio API: http://www.html5rocks.com/en/tutorials/webaudio/games/ [accessed May 2015]
[32] Roberts, C., Wakefield, G., & Wright, M. (2013). The Web Browser As Synthesizer And Interface. NIME.
[33] Walker, W., & Belet, B. (2015). Birds of a Feather (Les Oiseaux de Même Plumage): Dynamic Soundscapes using Real-time Manipulation of Locally Relevant Birdsongs.
[34] openAL. Retrieved from http://openal.org./ [accessed May 2015]
[35] Seo, B., Htoon, M. M., Zimmermann, R., & Wang, C.-D. (2010). Spatializer: A Web-based Positional Audio Toolkit. ACE. ACM.
[36] Dirksen, J., Learning Three.js – the JavaScript 3D Library for WebGL, (2015)
[37] Three.js Retrieved from http://threejs.org [accessed May 2015]
[38] GidHub, webaudiox, Retrieved from
[39] https://github.com/jeromeetienne/webaudiox [accessed May 2015]
[40] GidHub, howler.js. Retrieved from https://github.com/goldfire/howler.js [accessed May 2015]
[41] GidHub, pedalboard.js. Retrieved from
[42] http://dashersw.github.io/pedalboard.js/ [accessed May 2015]
[43] GidHub, wad, Retrieved from https://github.com/rserota/wad [accessed May 2015]
[44] GidHub, fifer, Retrieved from https://github.com/f5io/fifer-js [accessed May 2015]
[45] DevBattles. Learn Web Audio API. Retrieved from http://www.devbattles.com/sand/post-69-Learn-Web-Audio-API [accessed May 2015]
[46] Toyosak, A., Flahive, L., & Diaz, J. (2015). Music Story 2: a Motion-Graphic-based Music Video Creator.
[47] Mahadevan, A., Freeman, J., & Magerko, B. (2015). EarSketch: Teaching computational music remixing in an online Web Audio based learning environment. Web Audio Conference.
[48] Paradis, M., Clarke, R. G., & Melchior, F. (2015). VenueExplorer, Object-Based Interactive Audio for Live Events. WAC. Paris
[49] Wyse, L. (2015). Spatially Distributed Sound Computing and Rendering Using the Web Audio Platform. 1st Web Audio Conference. Paris.
[50] Pendharkar, C., Bäck, P., & Lonce, W. (2015). Adventures in scheduling, buffers and parameters: Porting a dynamic audio engine to Web Audio.
[51] Schnell, N., Saiz, V., Barkati, K., & Goldszmidt, S. (2015). Of Time Engines and Masters An API for Scheduling and Synchronizing the Generation and Playback of Event Sequences and Media Streams for the Web Audio API. WAC. Paris.
[52] Clark, C., & Tindale, A. (2014). Flocking: A Framework for Declarative Music-Making on the Web. ICMC|SMC. Athens.
[53] Zampoglou, M., Spala, P., Kontakis, K., Malamos, A. G., & Ware, J. A. (2013). Direct Mapping of X3D Scenes to MPEG-7 Descriptions. Web3D '13, ACM. doi:10.1145/2466533.2466540
[54] Pohja, M. (2003). X3D and VRML Sound Components.
[55] Extensible 3D (X3D) Sound component. Retrieved from http://www.web3d.org/documents/specifications/19775-1/V3.3/Part01/components/sound.html [accessed May 2015]
[56] Ahmad, L., Boukerche, A., Hamidi, A., Shadid, A., Pazzi, R. (2008). Web-based e-learning in 3D large scale distributed interactive simulations using HLA/RTI. Parallel and Distributed Processing, IEEE
[57] Repplinger, M., Loffler, A., Schug, B., Slusallek, P. (2009). Proceedings of the 14th International Conference on 3D Web Technology, ACM
[58] Browsers support Web Audio API. Retrieved from http://caniuse.com/#feat=audio-api [accessed May 2015]

Andreas Stamoulias was born in Athens, Greece, in 1987. He received a B.Sc. in Applied Informatics and Multimedia in 2011 and a M.Sc. in Informatics Engineering, in 2014, both at the Technological Educational Institute of Crete. Since 2011, he has been a Researcher at the Multimedia Content Laboratory, dept. of Informatics Engineering, T.E.I. of Crete. His main interests are 3D graphics and multimedia web applications.

Eftychia Lakka was born in 1982. She received her BSc degree in Computer Science from the University of Ioannina, Greece (2006) and her M.Sc. in Synthesis of Images and Graphic Designs from the University of Limoges, France (2011). She has been as Research Associate at the Advanced Knowledge, Image & Information Systems Laboratory, dept. of Informatics, T.E.I. of Athens (2010-2014) and at the Multimedia Content Laboratory, dept. of Informatics Engineering, T.E.I. of Crete (2015-now).

Athanasios G. Malamos was born in 1969. He received his BSc degree in Physics from the University of Crete (1992) and his PhD from the Technical University of Crete in 2000. From 1997 to 2002 he was a Research Assistant and a researcher in the ICCS National Technical University of Athens. Since 2002 he is with the Technological Educational Institute of Crete at the Department of Informatics Engineering as an Assistant Professor (2002-2006) and as an Associate Professor (2006 until present). He is the coordinator of the Multimedia Content Lab, and has supervised many research projects granted with EU and National research funds. He has served as program committee member and reviewer for several international conferences and workshops. Dr. Malamos is a reviewer for IEEE, Springer as other international journals. He is member of the IEEE Computer Society, ACM SIGGRAPH and the WEB3D consortium. His research interests include multimedia semantics, AI and graphics

Building Real-Time Collaborative Applications with a Federated Architecture

Pablo Ojanguren-Menendez[1], Antonio Tenorio-Fornés[1], and Samer Hassan[2]

[1] GRASIA research group of Complutense University of Madrid, Madrid, Spain,
[2] Berkman Center for Internet & Society (Harvard University, US), Cambridge, US

***Abstract* —Real-time collaboration is being offered by multiple libraries and APIs (Google Drive Real-time API, Microsoft Real-Time Communications API, TogetherJS, ShareJS), rapidly becoming a mainstream option for web-services developers. However, they are offered as centralised services running in a single server, regardless if they are free/open source or proprietary software. After re-engineering Apache Wave (former Google Wave), we can now provide the first decentralised and federated free/open source alternative. The new API allows to develop new real-time collaborative web applications in both JavaScript and Java environments.**

***Keywords—Apache* Wave, API, Collaborative Edition, Federation, Operational Transformation, Real-time**

I. INTRODUCTION

Since the early 2000s, with the release and growth of Wikipedia, collaborative text editing increasingly gained relevance in the Web . The wiki software [1] (such as MediaWiki, TikiWiki and others), which enabled scalable collaborative edition of documents, rapidly became popular. Nowadays, we can see thousands of wikis used by researchers, institutions, enterprises, and a wide diversity of communities to crowdsource the knowledge of the participants. Just Wikia [2], a wiki service provider, accounts for 300K wiki communities with 135M monthly visitors.

Writing texts in a collaborative manner implies multiple challenges, especially those concerning the management and resolution of conflicting changes: those performed by different participants over the same part of the document. That is, if Alice and Bob edit the same sentences at the same time, we should make sure none of their contributions is lost. In fact, in a scenario where we have hundreds or thousands or contributors over the same pages, such conflict is not rare. These conflicts are usually handled with asynchronous techniques as in version control systems for software development [3] (e.g. SVN, GIT), resembled by the popular wikis. In these environments, the software automatically merges contributions over different sections, but users are forced to "take turns" to edit the same sentences (or otherwise manually merge the others' contributions to theirs).

However, some synchronous services for collaborative text editing have arisen during the past decade. These allow users to write the same document in real-time collaboration (simultaneously), as in Google Docs [4] and Etherpad [5]. They tend to sort out the conflict resolution issue through the Operational Transformation [6] technology which has grown to become the de-facto standard in real-time collaborative systems. These services are typically centralised: users editing the same content must belong to the same service provider. However, if these services were federated, users from different providers would be able to edit contents simultaneously. Federated architectures provide multiple advantages concerning privacy and power distribution between users and owners, and avoid the isolation of both users and information in silos [7].

The rest of this paper is organised as follows: first, the state of the art of Operational Transformation frameworks is outlined in Section 2. Section 3 depicts the re-engineering approach and the technologies and tools that were used. Section 4 covers the main concepts of the original Wave Platform, and the changes that were performed are explained in detail. Afterwards, the results are discussed in Section 5. Finally, conclusions and next steps are presented in Section 6.

II. State of the Art of Real-time Collaboration

The development of Operational Transformation (OT) algorithms started in 1989 with the GROVE System [8]. During the next decade many improvements were added to the original work and an International Special Interest Group on Collaborative Editing (SIGCE) was set up in 1998. During the 2000s, OT algorithms were improved as long as mainstream applications started using them [9].

In 2009, Google announced the launch of Wave [10] as a new service for live collaboration where people could participate in conversation threads with collaborative edition based on the Jupiter OT system [11]. The Wave platform also included a federation protocol [12] and extension capabilities with robots and gadgets [13]. Allegedly because of lack of fast user adoption, in 2010 Google shut down the Wave service. However, as initially promised, Google released the main portions of the source code to the Free/Open Source community, and handed its ownership to the Apache Foundation. Since then, the project belongs to the Apache Incubator program and it is referred as Apache Wave [14]. Eventually, Google has included Wave's technology on several products, such as Google Docs and Google Plus. Despite its high technological potential, the original final product had a constrained purpose and a hardly reusable implementation.

Other web applications became relevant during that time, such as the Free/Libre/Open Source Software (FLOSS) Etherpad. However, it was mostly after the Google Wave period when FLOSS OT-based frameworks appeared, allowing the integration of real-time collaborative edition of text and data within third-party applications. The most relevant examples are outlined as follows.

TogetherJS [15] is a Mozilla FLOSS project that uses the WebRTC protocol for peer-to-peer communication among web browsers, together with OTs for concurrency control of text fields. It does not provide storage and it needs a server in order to establish communications. It is a JavaScript library and uses JSON notation for messages.

ShareJS [16] is a server-client FLOSS platform for collaborative edition of JSON objects as well as plain text fields. It provides a client API through a JavaScript library.

Goodow [17], is a recent FLOSS framework replicating the Google Drive Real-Time API with additional clients for Android and iOS, while providing its own server implementation.

DOI: 10.9781/ijimai.2015.356

On the other hand, *Google provides a Real-Time API* as part of its Google Drive SDK . It is a centralised (non-FLOSS) service handling simple data structures and plain text.

In general, these solutions are highly centralised. Despite they claim collaboration, users from different servers cannot work or share content. Besides, they mostly provide concurrency control features without added value services like storage and content management. And all of them just allow collaborative edition of simple plain text format.

III. Re-engineering: technologies and tools

This section summarises the procedure followed to re-engineer and build a generic Wave-based collaborative platform, together with the technologies used. First, it introduces the software and technologies that have been generalised, Apache Wave and Wave in a Box, and afterwards the technologies used to develop and test the performed extensions. The description of how and where the results are shared and published conclude this section.

A. Assessment of Apache Wave & Wave in a Box

Wave in a Box is the FLOSS reference implementation of the Apache Wave platform, which supports all former Google Wave protocols and specifications [18] and includes both implementations of the Server and the Client user interface. Most of its source code is original from Google Wave and was provided by Google, although it was complemented with parts developed by community contributors. It enables real-time collaboration over rich-text conversations in a federated infrastructure. It was designed to be an extensible platform through the use of gadgets and robots.

The existing source code is written in Java and the Google Web Toolkit (GWT) [19]. GWT is a FLOSS framework which allows to write Java code and translate it to JavaScript in order to be used in a Web browser. This approach is used to write all Wave components shared between server and client. User interface components are developed in GWT and they are strongly coupled to the Wave's business logic.

The lack of technical documentation forced to perform a preliminary extensive source code inspection, identifying main packages and interfaces and developing text documentation and diagrams. It was concluded that from a logical point of view, Wave concepts could be reused for general purposes, and that technically the source code was organised in layers properly decoupled.

B. Development & Testing frameworks

Both, server and client components of the Wave in a Box software have been extended. In particular, extensions to the server's storage system have been added to support the NoSQL database MongoDB [20] and some HTTP RESTful services have been also created. Part of new source code in client components has been written avoiding GWT dependencies in order to be reused in any Java runtime environment without adaptations. On top of this code, the JavaScript client API has been developed with some GWT specific code.

Concerning software testing, the JavaScript framework Jasmine [21] was used in addition to existing Java unit tests. The test suite attacks all JavaScript API functions in a web browser environment. These are end-to-end tests where all components of the Wave architecture are verified, from client API methods, to server's storage routines.

C. Contributions

The development has been tracked and released in an open and public source code repository [22]. It includes documentation and different examples about how to use the API.

Besides, during the development process, several contributions have been made to the Apache Wave FLOSS community, in the form of source code patches, documentation and diagrams.

IV. Generalising the Wave Federated Collaborative Platform

This section shows the fundamentals of the Wave platform and how they have been used to turn Wave into a general-purpose platform unlike the former conversation-based one.

A. Original Wave Data Models & Architecture

This subsection describes how original Wave data models work from a logical point of view. This allows further understanding of the presented work.

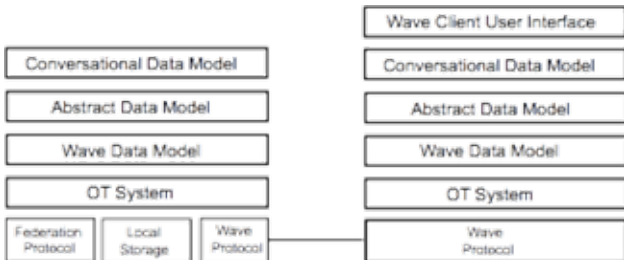


Fig. 1. Apache Wave Architecture, including data model layers.

1) The Wave Content Model

There are three different logical data models in the original Wave systems (Fig. 1). The Wave data model [23] is the basic level of data abstraction in the system providing a basic storage entity, Documents, and two aggregated entities: Wavelets and Waves.

Documents are XML documents where arbitrary data can be stored. They are logically grouped in a *Wavelet* which provides access control for the contained documents. Finally, Wavelets are grouped logically in *Waves*. A Wave is basically a unique identifier -for a particular domain- referencing a set of Wavelets which controls the access to a group of XML Documents.

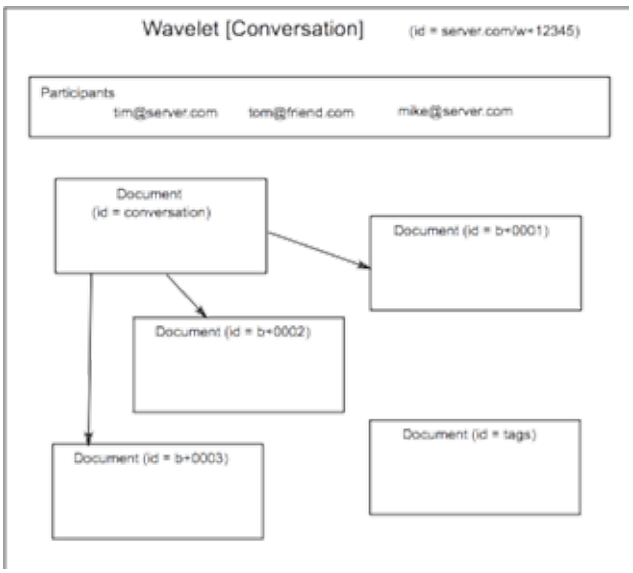


Fig. 2 Example of a Wavelet structure (Wave Data Model) representing a wave conversation (Wave Conversational Data Model)

The actual way to store these entities, and the Document's XML in particular, is through the historical set of changes performed to them. These changes are represented with a special set of character-based operations over a document: the Operational Transformations (OT) .

In the cases of having different users changing an entity at the same time, the OT's applied to the data entity through a special concurrency control algorithm ensures a consistent state of the entity, among all users, after all OT's have been applied. The OT system is responsible to implement such functionality. The implementation of the Wave Data Model allows to react when changes are performed over these entities thanks to this operation-based design.

2) The Abstract Data Model

In summary, the Wave Data Model enables only real-time collaborative editing of structured text (XML). However, it was convenient for the Wave system to handle non textual data as well. The Abstract Data Model provides a set of basic data structures –maps, lists and strings or Abstract Data Types (ADT)– which are represented as XML within Documents. This way, these data structures can be used by different users concurrently whereas they inherit the consistency properties of the underlaying OT system. Besides, the data model translates incoming OT's from the underlying data model in meaningful mutation events for data structures like "element is added", "element is removed", etc.

3) The Conversational Data Model

On top of these two layers, the Conversational Data Model [24] is placed. It provides the data entities and business logic of the original Google Wave product, focused on conversations.

A conversation is handled by a Wavelet, and each message is stored as a Document. The structure of messages is also stored in a Document but using the Abstract Data model instead: the logical structure of the thread can be seen as maps and lists of Documents' identifiers. The Conversational Data Model codifies the content's type of each Document within its identifier (Fig. 2).

These layers are deployed in a client-server architecture. The server side or "Service Provider" provides mainly OT history storage, OT system and federation control with other servers using the XMPP protocol [25]. Additional services like indexing and robots rely on the rest of already introduced data model layers. On the other hand, client side is responsible of the application logic and the user interface, therefore it handles all data layers as well.

The implementation of this architecture is a Java/GWT software originally developed by Google. This technology allows to use almost completely the same source code for all layers in both, server and client modules. Java source code is translated to optimised JavaScript by the GWT compiler. Just a few and specific parts tied to the execution environment are different between server and client, such as networking and random number generation. The server-client communication between follows the Wave Client-Server Protocol. It defines a set of operations and JSON data entities to exchange Operational Transformations for Waves, Wavelets and Documents.

B. General-Purpose Collaboration: Generalising the Wave Data Model & Architecture

Previous section outlined the original Wave's data models and architecture. This section introduces how they can be used in a generic way thanks to the new Wave Content Model, and the Wave Content API.

1) The Wave Content Model

The Wave Content Model is a new general-purpose data model built on top of both existing Wave and Abstract Data Models. It provides a more convenient set of data abstractions and relationships to work with Abstract Data Types. This new data model allows to see a Wavelet as a dynamic tree of nested data objects: maps, lists, text strings and rich text documents. These objects are stored in different Documents of the Wavelet whereas the new data model manages the organization of them and their relationships among the Documents properly (Fig. 3).

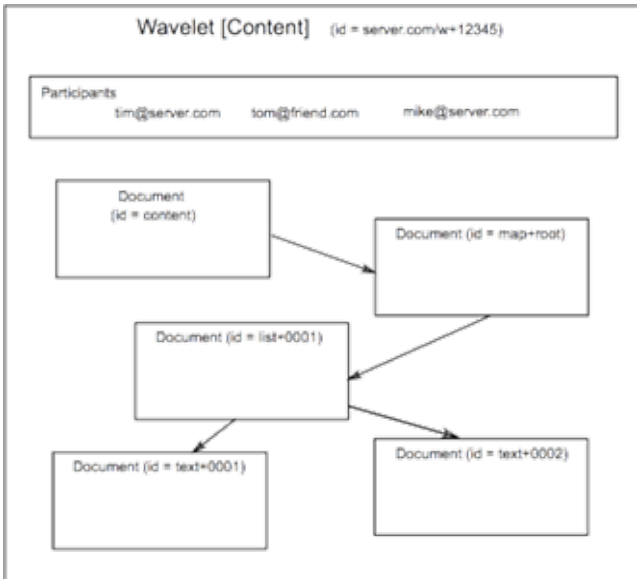


Fig. 3 Example of a Wavelet structure (Wave Data Model) representing a collaborative data object (Wave Content Model)

The Wave Content Model is implemented as a class hierarchy (Fig.4) controlling each possible data type –map, list, string and text– plus a controller class for the whole Wavelet, following the Composition Pattern [26].

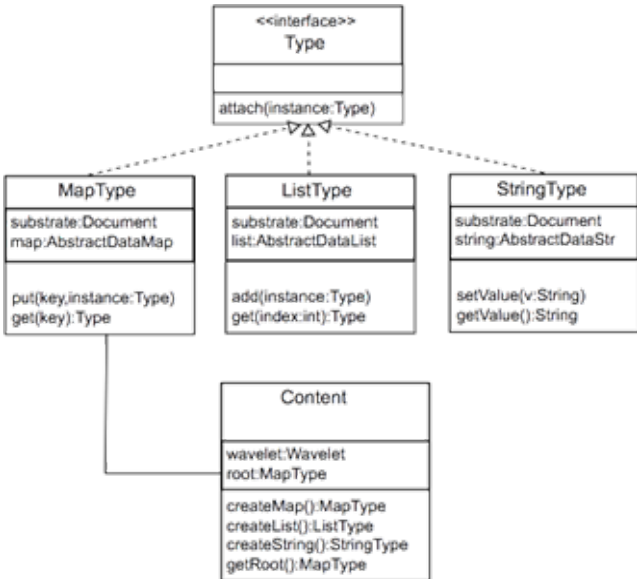


Fig. 4 Class hierarchy implementing the Wave Content Model.

A data class instance, or data objects, handles one single underlaying abstract data type instance over a single Document. New instances are initially unhooked from any Wavelet, so they must be attached to an existing parent instance. Attach process creates the underlying

substrate Document, the right Abstract Data Type handler and stores the new Document identifier as reference in the parent instance. This classes allow to register callback methods to be notified on model mutations.

With this approach, Wavelets -and Waves- became generic and dynamic data containers where multiple users can create and modify a nested data structure at the same time ensuring its consistency over the time.

In comparison with the former architecture stack, in the presented approach the Conversational Data Model has been removed and replaced by the Wave Content Model. Of course, the existing user interface layer is also removed (Fig. 5).

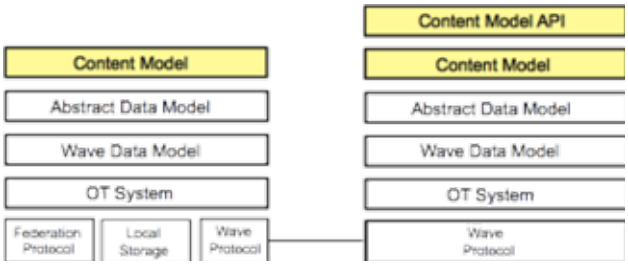


Fig. 5. New Apache Wave Architecture, including new content model

2) The Wave Content API

The new Wave Content Model allows to see Waves as real-time collaborative data structures. However, additional effort is required to expose this model to third-party applications in a handy manner.

According to the technology used in the Apache Wave implementation, just new Java or GWT web applications could use new content data model directly. With the aim of offering these new capabilities to any web application, a JavaScript API has been built.

Although GWT eventually translates Java code into JavaScript, this is not suitable to be consumed directly by non-GWT JavaScript code in a web-browser environment due to the following facts (among others): GWT-generated JavaScript, which is obfuscated by the compiler, does not provide references to objects with suitable names; GWT Exceptions do not flow out of the GWT code, so they must be translated and adapted to external code properly.

Java Script Native Interface (JSNI) and Overlay Types are features of GWT allowing to write arbitrary native JavaScript code integrated transparently within Java code. These features have been used to develop a native JavaScript layer which exposes functionality of the GWT-generated objects of the Wave Content Model. This is an implementation of the Proxy Pattern.

Additional functionality is also required in the JavaScript API. First, users no longer will use the former user interface to get registered or logged in. Therefore, the API provides replacement methods for making HTTP calls to create and authenticate users.

Management of the Wave life cycle now is provided through the API to clients. They can open or create Waves by calling API's methods. Moreover they can be aware of changes in the model registering callback functions in the API.

3) Content Search Index

Clients are able to query Waves stored in the Server Provider thanks to a new query service. Original Wave server implementation stores Wavelets as a sequence of OT's. This approach prevents to look into actual data of Documents to perform operations, for example executing search queries, regardless of the storage engine used.

A secondary storage is used now in order to provide a query service. Anytime the Server Provider commits a change to the main storage, an asynchronous indexing process takes care of the changed Wavelet: a full view of its Wave Content Model is generated in memory and a Visitor Pattern is used to transverse data objects generating an equivalent JSON document.

This process is optimised in two different ways: first, the number of times the indexing process runs is decreased by queuing committed changes sequentially and processing them in groups according their time closeness. Second, loading and transversing the full content model in memory is avoided by pruning. Each received change references to its target Document, which stores unequivocally one data object in the data model. This information is used to skip data model branches without changes in any of its data objects.

Finally, JSON documents are stored in the NoSQL database. The API encapsulates the database query interface and filters queries according to the current logged in user: a user cannot retrieve Wavelets where she is not a participant.

V. Discussion

This paper introduces the first federated platform for real-time collaboration available nowadays. However, using Wave involves some issues, mainly due to the limitations of the source code and its technologies.

There are several critiques concerning the complexity of the Wave OT system regarding two main issues: the complexity of the Operational Transformation system put in place [16] and the large length of the source code with around 500 thousand lines [27]. These facts together with the lack of good documentation causes the maintenance of the source code to be a tough task, requiring highly skilled developers in object-oriented programming with enough mathematical background. However, any OT system is inherently complex. To design a flexible and comprehensive set of operational transformations –such as Wave's– in order to provide an actually usable functionality is hard in any case. Besides, to implement control algorithms is a hard task, even if nowadays they are properly formalised.

Some existing OT implementations use a simpler approach. These OT systems are generally based in the JSON language, having a smaller set of OT operations just defined to operate at the language level. In contrast, Wave's OT system has significantly superior capabilities. It includes business logic operations in the system, such as add and remove participants to a Wavelet. But the most relevant features are to include XML tags and text annotations as part of the OT language. The first allows to handle any XML dialect, while the latter enables contextual meta data over that XML. These characteristics are used in the Wave's rich text format, which, for example, allows to embed arbitrary objects within the text, from images to widgets, just using new XML tags for them.

Operation's semantics and syntax of the introduced API follows the same style of the Google Drive Real-Time API: starting from a root map, new data objects must be created by a factory and then attached to the existing data tree. On the other hand, JSON based OT systems work seamlessly in JavaScript environments, allowing direct manipulation of the data. It is hard to conclude which approach is more appropriated, but the first seems more generic concerning the API implementation in different programming languages, as it is not as tied to JavaScript. Moreover, data structures of JSON documents and new Wavelet's inner structure are equivalent, so it would not be hard to develop adapters. However, currently there is no actual data about the developers preference, i.e. how comfortable are they with each approach.

Performance issues must be taken into account in the new Wave

Content Model. The first consideration is whether the new changes have a negative influence in the general performance of the platform in comparison with the original architecture. Regarding the client, no special impact in performance is expected as long as data objects of the new content model are created in memory only when access to them is required. On the server's side, no changes have been done affecting performance critical aspects of the OT system like in memory recreation of Wavelets and delta-based storage. However, current design of the JavaScript API duplicates some data structures of the underlaying data model to simplify the implementation. Internal improvements in this area could be performed, although they do not affect current or future use of the API.

The GWT development framework is sometimes seen as a disadvantage regarding efficiency and code complexity in comparison with development of native JavaScript software with modern native frameworks [28]. It is true that GWT was produced in a time when JavaScript tools and frameworks were not as advanced as today. However, it is a very stable and mature FLOSS project, and it is supported by Google. Moreover, the GWT compiler generates highly optimised code and it solves the issue of managing dual-language applications.

Client-Server communications relies massively on WebSockets [29] because changes in Wavelets are transmitted in both directions continuously. Protocol implementation is provided by an embedded Jetty HTTP server instance, a classic *Servlet* container which has been improved to support new HTTP features recently. It might be more efficient to use a non-blocking IO server [30] in order to improve vertical scalability. In addition, to use an embedded Jetty instance, prevents the deployment of the code into standard Java server containers.

Finally, it is necessary to assess the use of XMPP as a federated communication protocol among servers. It has been almost a standard for distributed communications in chat applications during more than a decade. However, the previous adoption from big players, such as Google and Facebook, has dropped. Moreover, it seems a heavy protocol to be used in small devices, and to support new features apart from chatting, especially in comparison with new decentralised protocols .

VI. Concluding Remarks and Future Work

A federated platform to develop web applications with real-time collaborative editing capabilities has been presented in the previous sections. It has been developed as a generalisation of the Apache Wave platform, the FLOSS project formerly known as Google Wave.

Nowadays there is no other federated (or distributed) platform for real-time collaboration of data and rich-text.

The provided API is a functional alternative to existing collaborative platforms. It provides a full-stack of software ready to be deployed, including functionalities only comparable with the proprietary Google Drive Real-Time API. Additional features such as the participation model, content storage and search index are part of the platform whereas they are missed in the rest of OT systems.

The API is offered in JavaScript and it can be used in any Web application. But thanks to the Java code base, it would be really easy to have versions for Java and Android applications. In such case, it would be an alternative to the lack of a Google Drive Real-Time API native client for Android.

From a wider perspective, this work opens new challenges in the context of decentralised collaboration:

In the introduced model, access and modification of content (and its structure) is granted to all participants in a Wavelet. However, this might not be enough for some sort of applications where read but not write permissions could be required for some users, e.g. a participant's profile information should not be written by anyone else whereas it must be readable by friend participants.

But also a fine-grain access control could be required beyond the current per-document access control. For instance, in a content Wavelet representing a poll, a user might be allowed to change her vote, but not to change others participants votes.

Under some circumstances integrity of the data model should be enforced, for instance allowing one and only one vote in the previous example. Or in a list of chess moves, enforcing the order and correctness of them.

Content Wavelets are highly flexible data entities for model application where the inner structure allows to define parent-child relationships of data elements. However, in any application, relationships among Wavelets or among inner objects of different Wavelets emerge naturally, so mechanisms to handle them must be explored, e.g. typifying Wavelets, object identification, etc.

Furthermore, in a scenario where several applications make use of the distributed data objects (for instance accessing profile information of users), the use of standard formats for data representation would be required. Technologies such as the Semantic Web [31] and Linked Data [32] provide an example of how distributed data can be organised and linked in a manner that allows further operations such as querying in a decentralised environment.

Current trends in software are driven by the mobile ecosystem. There, code and data are separated: *apps* running in devices, while retrieving data from a remote storage. Nowadays, it is easier to consider these apps managing data generated from different users and stored in different remote servers but eventually combining them in the device.

This work shows the unexplored high potentials of Google's original development, in spite of its complexity and lack of documentation. Thus, this work steps out engineering challenges for the reuse of parts of Apache Wave. The result is a platform ready to explore new challenges in decentralisation of data and services. We certainly hope this work will pave the way for other researchers and developers.

Acknowledgment

This work was partially supported by the Framework programme FP7-ICT-2013- 10 of the European Commission through project P2Pvalue (grant no.: 610961).

References

[1] B. Leuf and W. Cunningham, The Wiki Way: Collaboration and Sharing on the Internet. {Addison-Wesley Professional}, 2001.

[2] "Collaborative communities for everyone! - Wikia." [Online]. Available: http://www.wikia.com/Wikia.

[3] B. Berliner, "CVS II: Parallelizing software development," USENIX Association., pp. 341–352, 1990.

[4] Google Inc. "Google Docs." [Online]. Available: https://docs.google.com.

[5] The Etherpad Foundation, "Etherpad." [Online]. Available: http://etherpad.org/.

[6] Sun, S. Xia, C. Sun, and D. Chen, "Operational Transformation for Collaborative Word Processing," in Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, New York, NY, USA, 2004, pp. 437–446.

[7] C. A. Yeung, I. Liccardi, K. Lu, O. Seneviratne, and T. Berners-lee, "Decentralization: The future of online social networking," presented at the In W3C Workshop on the Future of Social Networking Position Papers, 2009.

[8] C. A. Ellis and S. J. Gibbs, "Concurrency Control in Groupware Systems,"

in Proceedings of the 1989 ACM SIGMOD International Conference on Management of Data, New York, NY, USA, 1989, pp. 399–407.
[9] "ACE - a collaborative editor." [Online]. Available: http://sourceforge.net/projects/ace/.
[10] A. Ferrate, Google Wave: Up and Running. O'Reilly Media, Inc., 2010.
[11] D. A. Nichols, P. Curtis, M. Dixon, and J. Lamping, "High-latency, Low-bandwidth Windowing in the Jupiter Collaboration System," in Proceedings of the 8th Annual ACM Symposium on User Interface and Software Technology, New York, NY, USA, 1995, pp. 111–120.
[12] Baxter, A. and Bekmann, J. and Berlin, D. and Gregorio, J. and Lassen, S. and Thorogood, S., "Google Wave Federation Protocol Over XMPP." Google Inc., 2009.
[13] G. Trapani and A. Pash, The Complete Guide to Google Wave. 3ones Inc, 2010.
[14] "Apache Wave Incubating." [Online]. Available: http://incubator.apache.org/wave/.
[15] Mozilla Labs, "TogetherJS." [Online]. Available: https://togetherjs.com/.
[16] J. Gentle, "ShareJS," Nov-2011. [Online]. Available: http://sharejs.org/.
[17] T. Chuanwu "Goodow - Google Docs–style collaboration via the use of operational transforms," GitHub. [Online]. Available: https://github.com/goodow.
[18] "Google Wave Protocol." [Online]. Available: http://www.waveprotocol.org/.
[19] R. Dewsbury, Google Web Toolkit Applications. Pearson Education, 2007.
[20] K. Chodorow, MongoDB: The Definitive Guide. O'Reilly Media, Inc., 2013.
[21] "Jasmine: Behavior-Driven JavaScript." [Online]. Available: http://jasmine.github.io/.
[22] P. Ojanguren, "SwellRT, a real-time federated collaboration framework." [Online]. Available: https://github.com/P2Pvalue/swellrt.
[23] A. North, "Wave model deep dive," 2010. [Online]. Available: https://cwiki.apache.org/confluence/display/WAVE/Wave+Summit+Talks
[24] G. North, A. J., "Google Wave Conversation Model," Oct-2009. [Online]. Available:
[25] http://wave-protocol.googlecode.com/hg/spec/conversation/convspec.html
[26] P. Saint-Andre, "Extensible Messaging and Presence Protocol (XMPP): Core," RFC Editor, RFC6120, Mar. 2011.
[27] E. Gamma, R. Helm, R. Johnson, and J. Vlissides, Design Patterns: Elements of Reusable Object-Oriented Software Pearson Education, 1994.
[28] "The Apache Wave (Incubating) Open Source Project on Open Hub." [Online]. Available: https://www.openhub.net/p/apache_wave.
[29] T. Burnham, CoffeeScript: Accelerated JavaScript Development. Pragmatic Bookshelf, 2011.
[30] I. Fette and A. Melnikov, "The WebSocket Protocol," RFC Editor, RFC6455, Dec. 2011.
[31] Gregor Roth, "Architecture of a Highly Scalable NIO-Based Server" 2007. [Online]. Available: https://today.java.net/pub/a/today/2007/02/13/architecture-of-highly-scalable-nio-server.html.
[32] T. Berners-Lee, J. Ora, L. Ora and others, "The semantic web," Scientific american, vol. 284, no. 5, pp. 28–37, 2001.
[33] C. Bizer, T. Heath and T. Berners-Lee, "Linked data-the story so far," Semantic Services, Interoperability and Web Applications: Emerging Concepts, pp. 205–227, 2009.

Pablo Ojanguren (Oviedo, 1979) holds a Engineering degree in Computer Science (2003) and a MSc in Software Engineering (2006) from the Universidad de Oviedo (Spain). It is also a certified Project Manager Professional. He is currently senior software engineer and researcher in the EU-funded FP7 P2Pvalue project on the development of webtools for Commons-based peer production. He has been running different IT positions in international companies as Accenture, BBVA and YellowPages Group with special focus in content management systems, enterprise integration patterns and IT project management. Pablo's main research area is decentralised architectures in social issues as commons-based peer production, peer-to-peer participation, digital democracy and data privacy.

Antonio Tenorio-Fornés (Madrid) holds an Engineers's Degree on Computer Science (2012) and a Master in Computer Science Research (2013) by the Complutense University of Madrid (Spain). He is currently doing his PhD research on democracy tools for Commons-based Peer Production Communities and working as researcher and engineer in the GRASIA research group of Complutense University of Madrid as part of the EU-funded FP7 P2Pvalue project. His research interests include decentralized technologies, Commons-based Peer Production communities, Artificial Intelligence, Multi-agent Systems, Agent-Based Social Simulation and declarative programing languages among others. Antonio Tenorio-Fornés (Madrid) holds an Engineers's Degree on Computer Science (2012) and a Master in Computer Science Research (2013) by the Complutense University of Madrid (Spain). He is currently doing his PhD research on democracy tools for Commons-based Peer Production Communities and working as researcher and engineer in the GRASIA research group of Complutense University of Madrid as part of the EU-funded FP7 P2Pvalue project. His research interests include decentralized technologies, Commons-based Peer Production communities, Artificial Intelligence, Multi-agent Systems, Agent-Based Social Simulation and declarative programing languages among others.

Samer Hassan (Madrid, 1982) holds an Engineering degree in Computer Science (2006), a MSc in Artificial Intelligence (2007) and a PhD in Social Simulation (2010) from the Universidad Complutense de Madrid (Spain), together with a Diploma in Political Science (2006) from the Spanish National Distance Education University (UNED, Spain). He is currently Fellow at the Berkman Center for Internet & Society (Harvard University, US) and Assistant Professor at the Universidad Complutense de Madrid (Spain). He has carried out research in distributed systems, social simulation and artificial intelligence from positions in the University of Surrey (UK) and the American University of Science & Technology (Lebanon). Coming from a multidisciplinary background in Computer Science and Social Sciences, he has more than 45 publications in those fields. Engaged in free/open source projects, he co-founded the Comunes Nonprofit and the Move Commons webtool project, and has been accredited as grassroots facilitator. He's involved as UCM Principal Investigator in the EU-funded FP7 P2Pvalue project on the development of webtools for Commons-based peer production. His research interests include Commons-based peer production, online communities, distributed architectures, social movements & cyberethics. Dr Hassan currently belongs to the GRASIA research group, the Berkman Center for Internet and Society, the Editorial Board of the Society for Modelling & Simulation newsletter, and has belonged to the European Social Simulation Association and the Center for Research in Social Simulation. He has been member of Scientific or Organising Committees of 45 international conferences.

Step Characterization using Sensor Information Fusion and Machine Learning

[1]Ricardo Anacleto, [2 3]Lino Figueiredo, [3]Ana Almeida, [1 4]Paulo Novais and [3]António Meireles

[1]ALGORITMI research group at University of Minho, Portugal
[2]Electrical Engineering Department, Institute of Engineering of Porto, Portugal
[3]Knowledge Engineering and Decision Support (GECAD) University of Minho, Portugal
[4]Computer Science at the Department of Informatics, in the School of Engineering of the University of Minho, Portugal

***Abstract* — A pedestrian inertial navigation system is typically used to suppress the Global Navigation Satellite System limitation to track persons in indoor or in dense environments. However, low-cost inertial systems provide huge location estimation errors due to sensors and pedestrian dead reckoning inherent characteristics. To suppress some of these errors we propose a system that uses two inertial measurement units spread in person's body, which measurements are aggregated using learning algorithms that learn the gait behaviors. In this work we present our results on using different machine learning algorithms which are used to characterize the step according to its direction and length. This characterization is then used to adapt the navigation algorithm according to the performed classifications.**

***Keywords* — Pedestrian Inertial Navigation System, Indoor Location, Learning Algorithms, Information Fusion**

I. Introduction

LOCATION information is an important source of context for ubiquitous systems, as it can be explored to improve life quality since emergency teams [1] can respond more precisely if the team members location is known, tourists can have better recommendations [2], the elderly can be better monitored [3], parents can be more relaxed with their children in shopping malls [4] and presence control systems can produce better reslts [22].

The major limitation of these systems is related to retrieving individual's location, which nowadays is only based on a GNSS (Global Navigation Satellite System), restricting the use of these systems to environments where GNSS signals are available. However, GNSS signals are not available inside buildings, in urban canyons, in the underground, underwater and in dense forests. Consequently location-aware applications sometimes cannot know the user location. Therefore, developing complementary localization technologies for these environments would unleash the use of many applications as presented above [23].

There are already some proposed systems that retrieve location in indoor environments. However, most of these solutions require a structured environment [5]. Therefore, these systems could be a possible solution for indoor environments, but in a dense forest or in urban canyons they are very difficult to implement.

To suppress structured environment limitations, a Pedestrian Inertial Navigation Systems (PINS) can be used. Typically, a PINS is based on an algorithm that involves three phases: step detection, step length estimation and heading estimation. A PINS uses accelerometers, gyroscopes, among other sensors, to continuously calculate via dead reckoning the position and orientation of a pedestrian. These sensors are based on MEMS (Microelectromechanical systems), which are tiny and lightweight sensors, making them ideal to integrate into the person's body. Unfortunately, large deviations of inertial sensors can affect performance, so the PINS big challenge is to correct the sensors deviations.

In the previous works of the research team, the step detection was improved by using an algorithm that combines an accelerometer and force sensors placed on the pedestrian's foot [6]. This approach led to better results [7] on the estimation of the pedestrian displacement. However, it still exists an error of 0.4% in step detection and an error of 7.3% in distance estimation.

We have found that a PINS solution only based on one IMU (Inertial Measurement Unit), composed by an accelerometer and a gyroscope, is not accurate enough. Thus, we believe that using several IMU in the person's body, combined with an information fusion strategy, will improve the accuracy of a PINS.

Information fusion is a multi-disciplinary research field with a wide range of potential applications in areas such as defense, robotics, automation and pattern recognition. During the past two decades, extensive research and development on multiple sensor data fusion has been performed for the Department of Defense of the United States of America [8]. This subject has been and will continue to be an ever-increasing interest field in research community, where it is intended to develop more advanced information fusion methodologies and architectures.

In the case of PINS, the MEMS sensors have some limitations and low accuracy, which does not happen on more expensive sensors like the ones used on aviation and military applications. To reduce the sensors complexity and thereby its cost, the information from a set of simple and low-cost sensors can be combined. This leads to the creation of a less expensive system, which captures accurate and reliable information about the pedestrian movements. Moreover, this fusion turns the system more fault tolerant [21].

Information fusion combined with learning techniques are being used in different INS fields to assist in displacement estimation. In robotics, Faceli et al. [9] use these techniques to improve the accuracy of distance measurements between a robot and the objects present in the environment by 7%.

These techniques are also used in autonomous driving vehicles. Stanley [10] software relied on machine learning and probabilistic reasoning techniques. Its IMU combined with artificial intelligence techniques were able to maintain accurate pose of the vehicle during GPS outages of up to 2 minutes.

In land vehicle applications, Caron et al. [11] and Noureldin et al. [12] propose machine learning techniques like neural networks, which introduce context variables and errors modelling for each sensor. Authors conclude that with an adequate modelling an accuracy

DOI: 10.9781/ijimai.2015.357

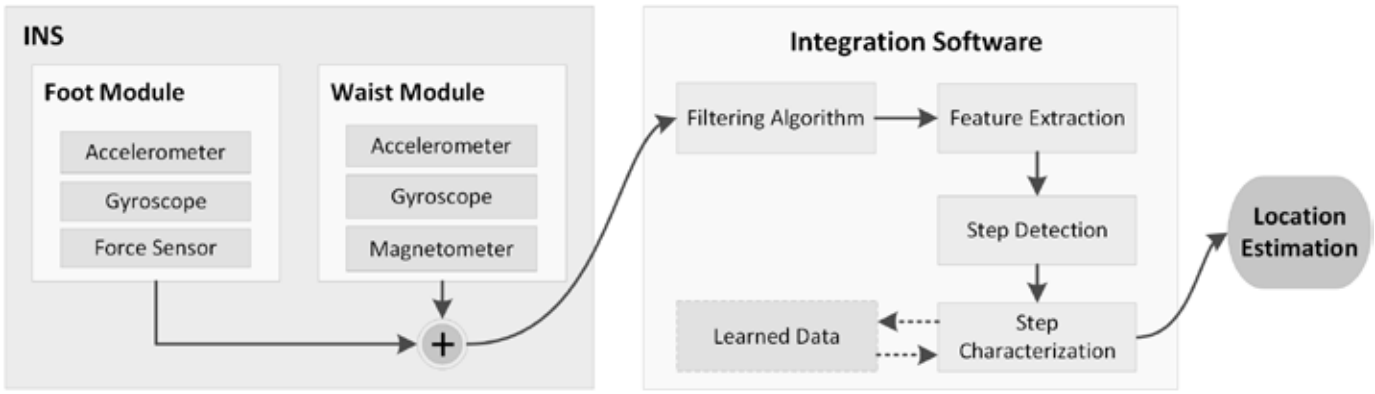


Figure 1 - System architecture.

improvement of 20% can be achieved. Recently, Noureldin et al. [13] have improved the previous results by considering past samples of INS position and velocity errors. Bhatt et al. [14] propose a hybrid data fusion methodology using Dempster-Shafer theory augmented by a trained Support Vector Machine (SVM), which corrects the INS errors. The proposed methodology has shown an accuracy improvement of 20%.

Since these experiences presented good results in the respective area, we wanted to explore similar techniques but applied to PINS. Our proposal applies an information fusion from several IMU spread in the person's body, and learning algorithms that based on contextual and past examples can improve the PINS accuracy. However, we needed to understand which is the best machine learning model to be used in a PINS.

This goal is addressed throughout the document, where the system architecture is presented in Section II. In the following three sections, Section III, IV and V, are presented the machine learning algorithms that were used to characterize a step. This characterization is applied to limit the typical error growing of PINS. In these sections is presented a comparison between a SVM and a Neural Network. In Section III are presented the algorithms that characterize the step according to the type of terrain, normal (flat) or stairs (i.e. ascending or descending).

In Section IV are presented the algorithms that characterize the step as forward or backward, and in Section V are presented the algorithms that classify a step according to its size (i.e. short, normal and long). In each section is presented an evaluation made to each algorithm. Finally, in Section VI are discussed the conclusions and the future work.

II. System Architecture

The proposed system is composed by two low-cost IMU, developed by the authors [6], and an "Integration Software" (described in Sections III, IV and V). The "Integration Software" starts by filtering the signals obtained from the sensors, then some features are extracted, which are used to detect a step and thereby to characterize it according to some previously learned data. Finally, the displacement is estimated based on the collected information. This architecture is represented in Figure 1.

When referring to a low-cost IMU it implies different things for researchers, since for some a thousand euros IMU is considered low-cost. However, in a PINS a low-cost IMU should cost less than 100€. This price restriction, implies the use of MEMS sensors that are truly low-cost.

The first IMU (Waist IMU), represented in Figure 2, is placed on the abdominal area and is composed by a STMicroelectronics L3G4200D gyroscope [15], a Analog Devices ADXL345 accelerometer [16] and a Honeywell HMC5883L magnetometer [17].

The second IMU (Foot IMU) is placed on the foot and is represented in Figure 3. It is composed by an Analog Devices ADXL345 accelerometer [16], a STMicroelectronics L3G4200D gyroscope [15] and two Tekscan FlexiForcer A201 force sensors [18]. Typically, an accelerometer is used to detect and quantify the foot movement, and the gyroscope is valuable to transform this acceleration data from the sensor frame to the navigation frame.

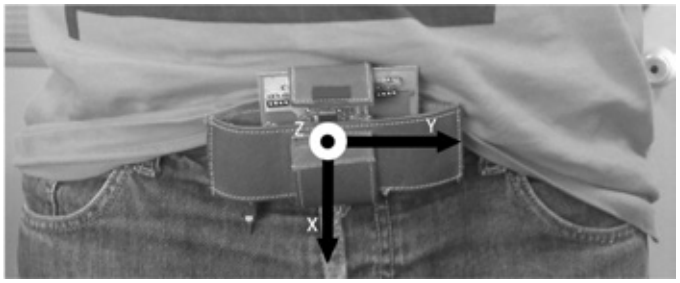


Figure 2 - Waist BSU with the corresponding axis

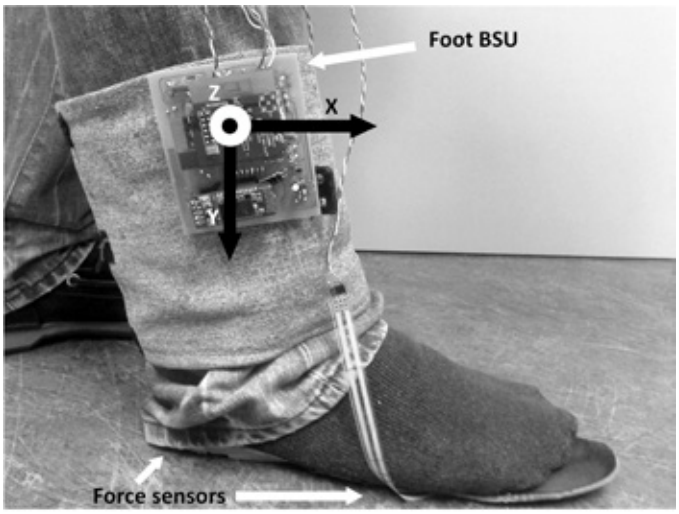


Figure 3 - Foot BSU with the corresponding axis

Force sensors were included since they can improve the detection of the moments when the user touches his feet on the ground, as well as, the correspondent contact force. The combination of force sensor data with accelerometer data improves the accuracy of the step length estimation [7]. One force sensor was placed on the front part of the foot and the other on the heel, as shown in Figure 3.

Although the pattern of the acceleration can be used to classify a step, sometimes the accelerometer produce a signal that does not follow any pattern, which turns to be useless to correctly classify a step.

These random readings can be surpass by using several sources of data combined with learning algorithms. The probability that two

sources of data give erroneous acceleration patterns at the same step is much reduced. The fusion between all the sensors information can improve the number of correct classifications.

III. Step Terrain

The first step characterization that is performed, is about the type of terrain where it was given. There are three possibilities: (i) in a normal (flat) terrain; (ii) or in ascending; (iii) or descending stairs.

For this characterization it was used the data from three sensors: (i) foot accelerometer (y-axis); (ii) foot gyroscope (z-axis); (iii) and waist accelerometer (x-axis).

The y-axis of the foot accelerometer provides relevant information about the foot elevation, which is essential to distinguish between ascending or descending stairs, since the forces are the opposite. However, from the several tests performed it was noticed that the main distinction that can be made using this sensor data is between ascending stairs and the other types of terrain. When ascending a stair the foot has to perform a higher elevation than in the other two cases. Regarding the other two types of terrain, descending stairs and normal, the data obtained from this sensor is very similar. The main difference is at the end of the step that, in the case of descending stairs, a higher acceleration is sensed since the foot touches the ground with a higher force than in the normal terrain type.

The z-axis of the foot gyroscope provides information about the foot rotation in each type of terrain. The foot rotation is much more noticeable in the ascending and descending stairs terrains. When ascending stairs it has an upward rotation peak and then a downward rotation peak, and it is the opposite when descending stairs. The data from this sensor is very important to make the distinction between these two types of terrains. Regarding the normal terrain, the pattern is similar to the descending stairs. However, the sensed rotation is much softer. Nonetheless this sensor provides a good accuracy on making the distinction between the three types of terrain.

Finally, the x-axis of the waist accelerometer provides similar data as the foot accelerometer. In ascending stairs a higher acceleration is sensed, in both positive and negative scales. When descending stairs this acceleration is much lower than in the other two types of terrain. The acceleration sensed in the normal terrain is within the other two. It provides similar data to distinguish between a flat surface and descending stairs. However, when ascending stairs it provides distinguishable data.

Considering the data provided by these signals, it can be established that, combining their data, they are suitable to be used to differentiate each possible characterization terrain. Since the strengths of each signal can be combined to achieve a final consensus.

To perform this characterization the learning algorithms were fed with a total of 72 inputs (24 inputs per each sensor). Each sensor signal was divided into 6 equal parts, and for each one of these parts the maximum, minimum and mean values were obtained, as well as, the slope. The slope was calculated based on the first and on the last measurement of each part. This data gives a total of 24 inputs per each sensor that are fed into the learning algorithm.

The division of the signal was made because giving to a classifier a complete signal can be very heavy and confusing to the algorithm to identify the patterns of the signal and therefore estimate the correct label for that pattern. Thus, it is reduced the dimensionality of the problem domain for the purposes of improving the performance of the algorithms and to decrease the computational load.

It was decided to divide the signal in 6 parts, because, during a step, each sensor signal is typically composed by 30 measurements. Thus, in order to have an average of 5 measurements per iteration the signal was divided into 6 equal parts. More parts will divide the signal too much, and fewer parts will pass insufficient information to the learning algorithm. Thus, the 6 was the number of parts that have best represented each one of the signals.

The learning algorithms were trained with a total of 970 samples, 358 samples of ascending stairs steps, 358 samples of descending stairs steps and 254 samples of normal terrain steps. To validate the algorithms a total of 170 samples were used (62 ascending, 62 descending and 46 normal). To test the algorithms a total of 540 samples were used (180 ascending, 260 descending and 100 normal), and a 10-fold cross-validation using these datasets was also performed.

A. SVM

Since in this characterization there are three possible classes (i.e. normal, ascending or descending stairs), and the SVM models can only classify two at each time, three SVM models (SVM Model 1, SVM Model 2 and SVM Model 3) were created. From the executed tests it was identified that the best results were achieved using a kernel, configured as a 3th order polynomial. This architecture is represented in Figure 4.

The models were trained with the same data, but with different class labels vectors. In this case there are three vectors. The first vector, which is used by the SVM Model 1, indicates that the ascending stairs steps belong to the positive class and the others to the negative. The second vector, which is used by the SVM Model 2, indicates that the descending stairs steps are the positive entries and the other the negatives. The third vector, which is used by the SVM Model 3, indicates that the normal terrain steps are the positive classifications and the others the negative. Meaning that the positive class of each classifier is ascending stairs, descending stairs and normal terrain, respectively.

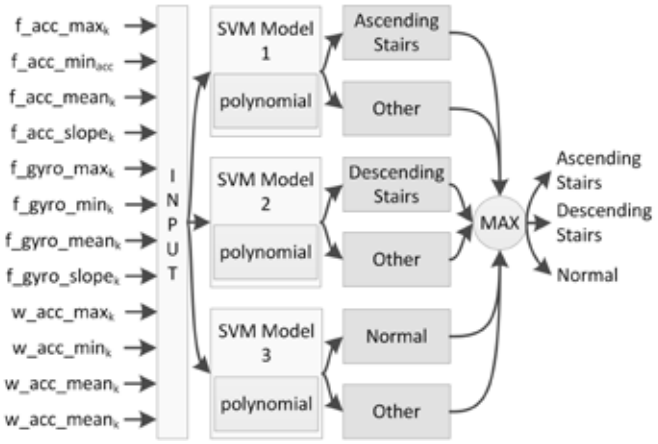


Figure 4 - SVM architecture for step terrain characterization

The score of the new observations are then estimated using each classifier. This will create a vector with three scores, one per each classifier. The index of the element with the highest score is the index of the class to which the new observation most likely belongs. For example, if the first index has the highest value, then the step is classified as ascending stairs. Thus each new observation is associated with the classifier that gives to it the maximum score.

After the learning phase, a 10-fold cross validation to the model was performed. The SVM Model 1 presented no error, the SVM Model 2 presented an error of 0.8% and the SVM Model 3 presented an error of 2.6%.

B. Neural Network

In Figure 5 is represented the design of the implemented neural network that classifies the type of terrain. The neural network receives as input (j) the 72 features previously presented. This input is passed to the Hidden Layer, which is composed by 144 neurons. Then, the Output Layer returns the final result about the type of terrain where the step was given.

The neural network parameters namely, the number of neurons in the hidden layer, the learning rate and the number of iterations, were tuned by trial and error. The learning rate was defined as 0.01 and the number of iterations as 36.

The mean squared error of the best validation performance was 7.97×10^{-7} with a gradient of 9.80×10^{-7} at epoch 36.

The error given by the neural network is very low, where during the training phase more than 98% of the results are very close to zero error. The highest error for an instance was of 1.50×10^{-5}.

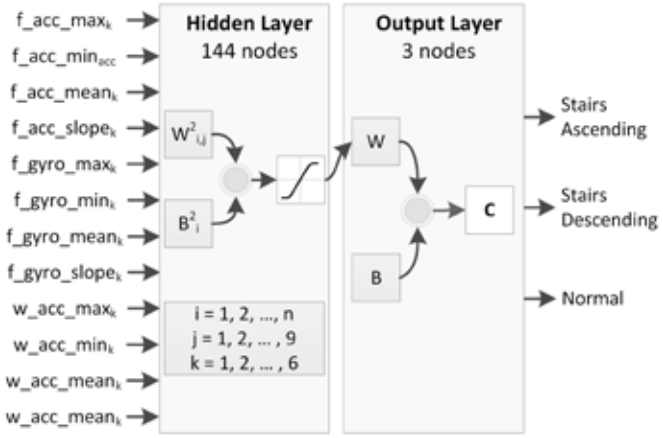


Figure 5 - Neural Network architecture for step terrain characterization

C. Evaluation

The implemented algorithms that characterize the type of terrain, were evaluated using a dataset of 800 steps performed by two pedestrians (400 steps for each pedestrian).

The test scenario is represented in Figure 6, which involves a complex path with a set of straight walks and a set of stairs.

The results obtained for this scenario can be seen in Table 1. This table presents for each algorithm, the categorization accuracy (in percentage).

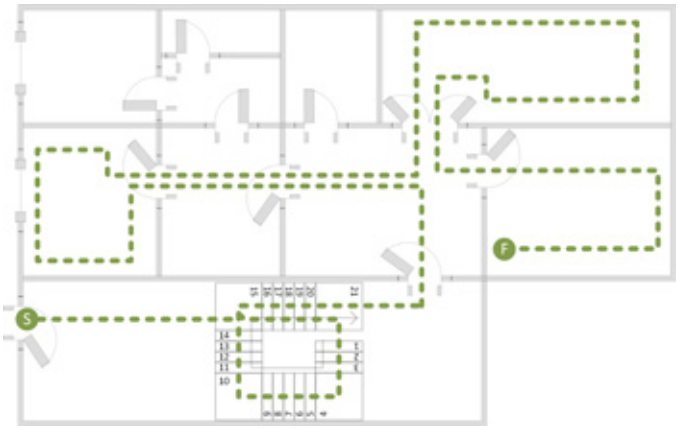


Figure 6 - Evaluation Scenario

For all the algorithms are presented the results obtained, in separate for each BSU and for the combination of the data of both BSU. This allows to identify which one of the BSU has higher accuracy in each characterization type.

TABLE 1 - ACCURACY RESULTS FOR STEP TERRAIN CHARACTERIZATION

Method	Ascending		Descending		Normal	
	Waist BSU	Foot BSU	Waist BSU	Foot BSU	Waist BSU	Foot BSU
SVM	97.5%	99.4%	94.2%	99.4%	85.1%	94.9%
N.N.	98.3%	100%	94.1%	100%	87.2%	94.9%
SVM Fusion	99.4%		99.5%		96.2%	
N.N. Fusion	100%		100%		98.7%	

Considering the obtained results it can be concluded that the ascending stairs class is the easiest to classify. The normal terrain class is sometimes confused with the descending stairs class, so it is with this misclassification that most errors occur.

Regarding the BSUs, the foot BSU gives more accurate data, since the foot is closer to the ground. The waist BSU can give a good indication about the vertical movement of the body. However, it obtains similar data when descending stairs and in normal terrain. Thus, it presents worst results in these classifications.

Interpreting the results obtained for each algorithm using each BSU in separate, it can be concluded that the Neural Network achieves better results on both BSU locations.

Analyzing the results obtained for each algorithm when considering the fusion of both BSU, the Neural Network presented the best results, achieving a mean accuracy of 99.4%, having 100% of accuracy on predicting the ascending and descending stairs classes.

Also, it can be concluded that through the sensors complementarity the type of terrain was categorized with higher accuracy.

From our tests it was identified that a learned dataset 5 times smaller, than the used one, is sufficient to achieve similar results. Making the learning procedure simpler and faster to a pedestrian perform before using our system.

Concluding, the evaluation results show that both BSU give similar results on detecting each type of terrain, but with their integration better results can be achieved.

IV. Step Direction

The second characterization performed to a step is about the direction that it can take. There are two possibilities, a forward step, which is the most natural to a human perform, or a backward step.

During this research, by analysing the datasets collected from all the walks, it was identified that step direction can be characterized by combining the data obtained from two sensors placed in different BSU, the foot accelerometer and the waist gyroscope. In the case of the accelerometer, the one placed in the pedestrian's foot gives more accurate results than the one on the waist. However, in the case of the gyroscope, the best results can be achieved with the one placed on the waist, since it give us the pelvic rotation, which combined with the accelerometer data is important to determine the direction of a step.

To classify the step direction 25 features where extracted from the sensors measurements, where 24 are obtained from the accelerometer data and 1 from the gyroscope data. To extract the features from the accelerometer signal, it was divided into 6 equal parts. For each one of these parts the maximum, minimum and mean values were obtained, then it was calculated the slope.

The other input is obtained from the gyroscope signal, which represents the motion of the waist. If the pelvic as a positive rotation, then the value 1 is assigned to the input, if it is a negative rotation the value 0 is assigned to the input. For example, for a forward left step the rotation will be positive and for a backward left step the rotation will be negative, for the right foot it is the opposite.

To train the learning algorithms the following number of samples was used: 450 samples (190 forward and 260 backward) for training, 100 samples for validation (40 forward and 60 backward) and 100 samples (50 forward and 50 backward) for testing. A 10-fold cross-validation using these datasets was also performed.

A. SVM

The design of the implemented SVM approach can be seen in Figure 7. It receives as input the 25 features previously presented. This input is passed to the Hidden Nodes that estimate the best separating hyperplane between the two classes, which maximizes the margin between the two classes. This division is performed using a "linear" kernel. Then, the Output Layer returns the final result about the step direction. A degree of confidence for the two possible results is given by this algorithm.

During the cross-validation the algorithm achieved an accuracy of 100% for both classes, forward or backward.

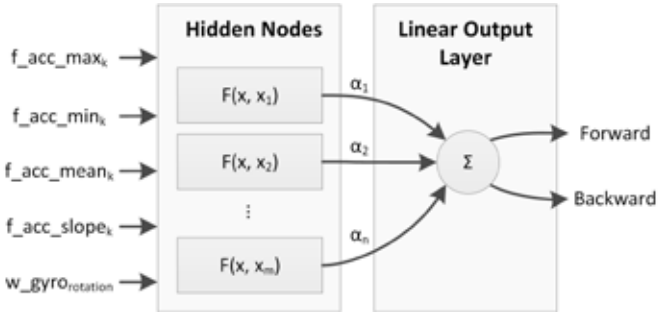


Figure 7 - SVM architecture for step direction characterization

B. Neural Network

The design of the implemented neural network can be seen in Figure 8. The neural network receives as input (j) the 25 features previously presented. This input is passed to the Hidden Layer, which is composed by 10 neurons. Then, the Output Layer returns the final result about the step direction.

The learning rate was defined as 0.01 and the number of iterations was defined as 31.

The mean squared error of the best validation performance is 2.08×10^{-8} with a gradient of 9.58×10^{-7} at epoch 35. The error is very low when training the network, but even lower when validating and testing the established neural network. Also, more than 90% of the results are very close to zero error. The highest error for an instance, during the network training, was of 0.15.

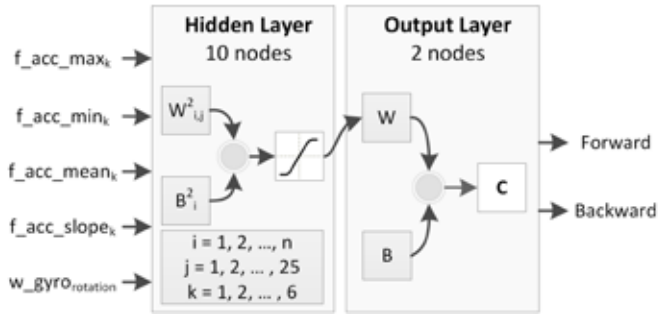


Figure 8 - Neural Network architecture for step direction characterization

C. Experimental Results

The step direction characterization algorithms were evaluated using a dataset of 240 steps performed by two pedestrians (120 steps for each pedestrian).

The test scenario is the same as in the previous characterization, where it was walked in forward and in backward movements. Two runs in this scenario, for each pedestrian, were performed.

The results obtained for this scenario can be seen in Table 2. From the obtained results, it can be concluded that, to perform this characterization, the data obtained from the foot BSU sensors are better, than the data obtained from the waist BSU sensors. This mainly happens because when the user is moving the foot is a more stable platform than the waist. A lot of unwanted accelerations are sensed by the waist, which leads to a poor characterization of the step direction, but there are some features that can be retrieved to help other sources to properly characterize the step.

Regarding the step direction characterization the backward one is the most difficult to classify. Mainly because for a human a forward step is a more natural movement to perform than a backward one.

The step direction is a simple characterization to be performed to a step, so the results were the expected. The learning algorithms proved to have an accuracy of 100%.

For this characterization it was not detected any difference in behaviour between the two learning algorithms.

TABLE 2 - ACCURACY RESULTS FOR STEP DIRECTION CHARACTERIZATION

Method	Ascending		Descending	
	Waist BSU	Foot BSU	Waist BSU	Foot BSU
SVM	98.6%	100%	96.7%	100%
N.N.	99.4%	100%	95.5%	100%
SVM Fusion	100%		100%	
N.N. Fusion	100%		100%	

V. Step Length

The third, and final, characterization performed to a step is about the length class. There are three possibilities, a short, a normal or a long step. These intervals must be defined from a set of exercises for a pedestrian in specific. Based on the collected data, and on the average of the collected steps, it was considered that short steps are the ones with a maximum distance of 30cm, the normal steps size ranges between 30cm and 45cm, and the long steps have a size longer than 45cm.

The x-axis of the foot accelerometer measures the acceleration that is sensed in the horizontal movement of the foot. The quantification of this acceleration is important, because it is correlated with the performed displacement. As the duration and the peak of the acceleration is higher, the longer is the step.

The force sensor data gives reliable information about the amount of time that the foot is not in contact with the ground and, about the force intensity that is made when touching the ground, as well as, when lifting the foot from the ground. The amount of time that the foot is in the air, can be correlated with the acceleration. A higher acceleration value combined with a longer duration of the foot in the air, indicates that a longer step was made.

The x-axis of the gyroscope data gives reliable information about

the rotation that was performed by the pelvic, where a higher rotation corresponds to a longer step. However, in corners this rotation can be higher for the same type of step. Thus, the combination of this data with the data given by the foot BSU sensors is important to achieve more accurate results. Many errors can occur by using the gyroscope data by itself.

The implemented learning algorithms have as input the foot force sensor and accelerometer data, and the waist gyroscope data. A total of 29 features are fed into the learning algorithms to classify the step length, where 24 are retrieved from the foot accelerometer, 3 from the foot force sensor and 2 from the waist gyroscope.

The foot accelerometer signal was divided into 6 equal parts, as shown in the previous implemented neural networks. This gives a total of 24 inputs that are fed into the learning algorithm.

The next 3 features are obtained from the force sensor signal. The first one is the number of measurements that exist until the maximum force value occur (foot touches the ground). A stronger impact means that the step was longer. The other feature is the force applied when the foot lifts up from the ground. This gives information about the impulse that was performed in the leg in order to perform some horizontal movement. Typically, when the impulse is higher the step is longer. The last feature retrieved from the force sensor is the number of measurements with value of zero, which corresponds to the amount of time that the foot is in the air.

From the gyroscope signal two features are extracted. The first one is the amplitude of the signal, which is correlated to the size of a step. Typically, a higher rotation means that the step is longer. The second feature is the length of the signal, which correlated with the amplitude, gives important information about the size of the step.

From the several tests performed, these features were the ones that had the best results in classifying the possible length of a step.

The learning algorithms were trained with a total of 855 samples, 435 samples of a short step, 225 samples of a normal step and 195 samples of a long step. To validate the algorithms a total of 171 samples were used (87 short, 45 normal and 39 long). To test the algorithms a total of 114 samples were used (58 short, 30 normal and 26 long). As in the other characterizations a 10-fold cross-validation using these datasets was also applied.

A. SVM

The design of the implemented SVM approach can be seen in Figure 9. This approach receives as input the 29 features previously presented.

In this characterization three SVM models (SVM Model 1, SVM Model 2 and SVM Model 3) were created. After some testing, it was identified that the best results were achieved with the following configuration for each model:

- SVM Model 1 was configured to classify the short steps using a "rbf" (radial basis function or Gaussian) kernel configured with an automatic scale;
- SVM Model 2 was configured to classify the normal steps with a "polynomial" kernel, configured as a 2nd order polynomial;
- SVM Model 3 was configured to classify the long steps using a "linear" kernel.

The models were trained with the same data, but with different class labels vectors. In this case there are three vectors. The first vector, which is used by the SVM Model 1, indicates that the short steps belong to the positive class and the others to the negative. The second vector, which is used by the SVM Model 2, indicates that the normal steps are the positive entries and the others the negatives. The third vector, which is used by the SVM Model 3, indicates that the long steps are the positive classifications and the others the negatives.

The score of the new observations are then estimated using each classifier. This will create a vector with three scores, one per each classifier. The index of the element with the highest score is the index of the class to which the new observation most likely belongs. For example, if the first index has the highest value, then the step is characterized as short. Thus, each new observation is associated with the classifier that gives to it the maximum score.

After the learning phase, a 10-fold cross validation to the model was performed. The SVM Model 1 had no error, the SVM Model 2 presented an error of 7% and the SVM Model 3 presented an error of 0.8%.

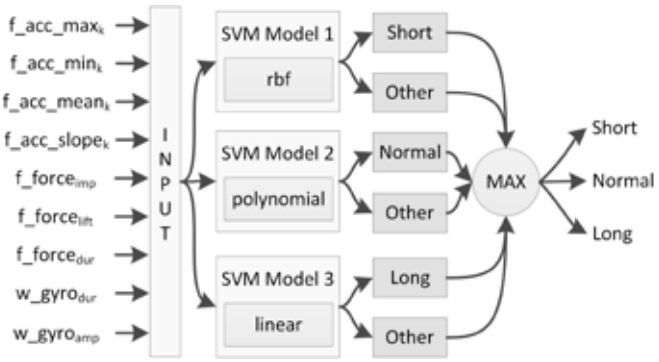


Figure 9 - SVM architecture for step length characterization

B. Neural Network

The implemented neural network to classify the step length is shown in Figure 10. The neural network receives as input (j) the 29 features previously presented. This input is passed to the Hidden Layer, which is composed by 60 neurons. Then, the Output Layer returns the final result about the step length.

As in the other characterizations the learning rate was defined as 0.01, and the number of iterations as 25.

The mean squared error of the best validation performance was 7.55×10^{-5} with a gradient of 4.59×10^{-4} at epoch 25.

During the training phase more than 95% of the results were very close to zero error, where the highest error for an instance, during the network training, was of 0.20.

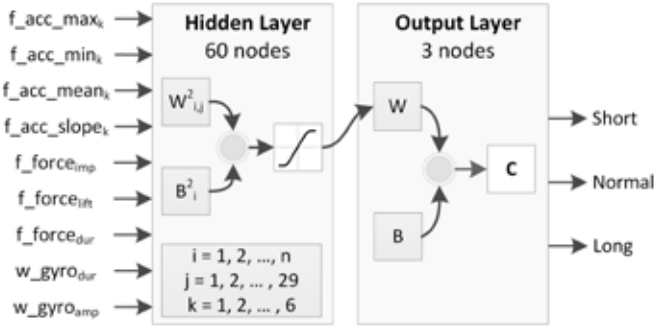


Figure 10 - Neural Network architecture for step length characterization

C. Evaluation

The step length characterization algorithms were evaluated using a dataset of 300 steps performed by two pedestrians (150 steps for each pedestrian).

The test scenario is represented in Figure 6 and the obtained results can be seen in Table 3.

TABLE 3 - ACCURACY RESULTS FOR STEP LENGTH CHARACTERIZATION

Method	Short		Normal		Long	
	Waist BSU	Foot BSU	Waist BSU	Foot BSU	Waist BSU	Foot BSU
SVM	98.4%	96.1%	80.3%	81.9%	92.6%	90.7%
N.N.	98.3%	100%	86.7%	83.3%	96.2%	100%
SVM Fusion	100%		86.0%		98.1%	
N.N. Fusion	100%		90.0%		96.2%	

Considering the obtained results, it can be concluded that for a short step both BSU present similar results. For this classification the learning algorithms presented an accuracy of almost 100%.

For a long step, there is not an evident difference between BSUs, since when considering the data from each BSU the Neural Network had the best results. However, when using the data from both BSU, SVM has the best performance.

The normal step is the most difficult to classify. The main reason for this phenomenon is because it sits between the other two classes. For this classification the combination of both BSU gives better results than using the data from each BSU individually. Nonetheless, none of the misclassifications given by the algorithms was to the opposite class, meaning that a short step was never classified as a long step and vice-versa.

Analysing the obtained results, it can be concluded that through the sensors complementarity the step length was categorized with higher accuracy. Also, it can be concluded that the learning of the gait parameters enables a more precise characterization of a step. The Neural Network gave the best results, having a mean accuracy very close to 96%.

From our tests it was identified that a learned dataset 10 times smaller is sufficient to achieve similar results. Making the learning procedure simpler and faster to a pedestrian perform before using this system.

The tests and the evaluation results, have shown that both BSU give similar results on detecting short steps, but in the case of normal steps the foot BSU has a higher accuracy. However, the long steps are better detected by the waist BSU.

Combining the data from both BSU, the weaknesses of one are suppressed by the advantages of the other, thus improving the overall results.

VI. Conclusion

Develop a PINS to be used by pedestrians in their daily life is a huge challenge. Many approaches already have been proposed, but must of them rely on a structured environment that usually is infeasible to implement and the others don't provide the necessary accuracy.

To suppress some of these limitations we propose a PINS based on low-cost sensors and on fusion and learning techniques. The sensors are placed on the foot and on the waist of a pedestrian, and their information is combined to achieve more accurate location estimation results. The data from both IMU was heavily explored in order to provide an acceptable level of performance, since one IMU can complement the other in the different activities that a pedestrian can perform.

The proposed system characterizes the step according to the activity that the pedestrian is performing. This characterization starts by estimating the type of terrain where the step was given. Then it estimates if the step was a forward or a backward one. This is very important to correctly estimate the pedestrian displacement, since they are opposite directions. The third classification is regarding the step length. This characterization fits into one of three categories: short, normal or long. With this classification we limit the displacement estimation according to the bounds of each category.

The inclusion of the step characterization module, through the use of more than one IMU and the neural network algorithm, led to an improvement, compared to the previous results [7], in displacement estimation of 52%. In the same scenario the error has decreased from 7.3% to 4.8%.

In the future we want to divide the step length characterization into more classes, to verify if it improves the displacement estimation accuracy. Also, we want to implement more step characterization characteristics.

References

[1] J. Elwell, "Inertial navigation for the urban warrior," in Proceedings of SPIE, vol. 3709, 1999, pp. 196 – 204.

[2] J. Lucas, N. Luz, M. Moreno, R. Anacleto, A. Almeida, and C. Martins, "A hybrid recommendation approach for a tourism system," Expert Systems with Applications, vol. 40, no. 9, pp. 3532 – 3550, Jul. 2013.

[3] J. Ramos, R. Anacleto, A. Costa, P. Novais, L. Figueiredo, and A. Almeida, "Orientation system for people with cognitive disabilities," in Ambient Intelligence - Software and Applications, ser. Advances in Intelligent and Soft Computing. Springer Berlin Heidelberg, Jan. 2012, no. 153, pp. 43 – 50.

[4] R. Anacleto, N. Luz, A. Almeida, L. Figueiredo, and P. Novais, "Shopping center tracking and recommendation systems," in Soft Computing Models in Industrial and Environmental Applications, 6th International Conference SOCO 2011, ser. Advances in Intelligent and Soft Computing. Springer Berlin Heidelberg, 2011, no. 87, pp. 299 – 308.

[5] J. Hightower and G. Borriello, "Location systems for ubiquitous computing," IEEE Computer, vol. 34, no. 8, pp. 57 – 66, 2001.

[6] R. Anacleto, L. Figueiredo, A. Almeida, and P. Novais, "Person localization using sensor information fusion," in Ambient Intelligence - Software and Applications, ser. Advances in Intelligent Systems and Computing. Springer International Publishing, Jan. 2014, no. 291, pp. 53–61.

[7] R. Anacleto, L. Figueiredo, A. Almeida, and P. Novais, "Localization system for pedestrians based on sensor and information fusion," in 17th International Conference on Information Fusion (FUSION), July 2014, pp. 1–8.

[8] M. Liggins, D. Hall, and J. Llinas, Handbook of Multisensor Data Fusion: Theory and Practice, Second Edition, 2nd ed. CRC Press, 1997.

[9] K. Faceli, A. d. Carvalho, and S. O. Rezende, "Combining Intelligent Techniques for Sensor Fusion," Applied Intelligence, vol. 20, no. 3, pp. 199–213, May 2004.

[10] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, K. Lau, C. Oakley, M. Palatucci, V. Pratt, P. Stang, S. Strohband, C. Dupont, L.-E. Jendrossek, C. Koelen, C. Markey, C. Rummel, J. van Niekerk, E. Jensen, P. Alessandrini, G. Bradski, B. Davies, S. Ettinger, A. Kaehler, A. Nefian, and P. Mahoney, "Stanley: The robot that won the DARPA Grand Challenge," Journal of Field Robotics, vol. 23, no. 9, pp. 661–692, Sep. 2006.

[11] F. Caron, E. Duflos, D. Pomorski, and P. Vanheeghe, "GPS/IMU Data Fusion Using Multisensor Kalman Filtering: Introduction of Contextual Aspects," Information Fusion, vol. 7, no. 2, pp. 221–230, Jun. 2006.

[12] A. Noureldin, T. B. Karamat, M. D. Eberts, and A. El-Shafie, "Performance Enhancement of MEMS-Based INS/GPS Integration for Low-Cost Navigation Applications," IEEE Transactions on Vehicular Technology, vol. 58, no. 3, pp. 1077–1096, Mar. 2009.

[13] A. Noureldin, A. El-Shafie, and M. Bayoumi, "GPS/INS integration utilizing dynamic neural networks for vehicular navigation," Information Fusion, vol. 12, no. 1, pp. 48–57, Jan. 2011.

[14] D. Bhatt, P. Aggarwal, V. Devabhaktuni, and P. Bhattacharya, "A novel hybrid fusion algorithm to bridge the period of GPS outages using low-cost INS," Expert Systems with Applications, vol. 41, no. 5, pp. 2166–2173, Apr. 2014.

[15] STMicroelectronics, "L3g4200d: Three axis digital output gyroscope," 2014. Available: http://www.st.com/st-web-ui/static/active/en/resource/technical/document/datasheet/CD00265057.pdf.
[16] A. Devices, "Adxl345 digital accelerometer," 2014. [Online]. Available: http://www.analog.com/static/imported-files/data sheets/ADXL345.pdf.
[17] Honeywell, "3-axis digital compass ic hmc5883l," 2014. Available: http://www51.honeywell.com/aero/common/documents/myaerospacecatalog-documents/Defense Brochures-documents/HMC5883L 3-Axis Digital Compass IC.pdf.
[18] Tekscan, "Flexiforce sensors for force measurement," 2014. [Online]. Available: http://www.tekscan.com/flexible-force-sensors.
[19] J. Saunders, V. Inman, and H. Eberhart, "The major determinants in normal and pathological gait," The Journal of Bone & Joint Surgery, vol. 35, no. 3, pp. 543 – 558, 1953.
[20] C. Vaughan, B. Davis, and J. O'connor, Dynamics of human gait. Human Kinetics Publishers Champaign, Illinois, 1992.
[21] C. Mateos, C. Ruiz, R. Crespo and A. Sanz, "Relative Radiometric Normalization of Multitemporal Images", International Journal of Artificial Intelligence and Interactive Multimedia, Vol 1, no. 3, pp. 54-59, 2010
[22] S. Rios-Aguilar, "Intelligent Position Aware Mobile Services for Seamless and Non-Intrusive Clocking-in", International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 2, no. 5, pp. 48-50, 2014
[23] J. Espada, V. García-Díaz, R. Crespo, B. G-Bustelo, J. Lovelle, "Improving the GPS Location Quality Using a Multi-agent Architecture Based on Social Collaboration", Pratical Applications of Intelligent Systems, Advances in Intelligent Systems and Computing, Springer Berlin Heidelberg, Vol. 279, pp. 371-379, 2014

Ricardo Anacleto. Since 2011 it is a PhD student in the Doctoral Program MAP-i in the areas of Mobile Computing and Localization Systems. It teaches in ESTGF - School of Technology and Management Felgueiras, teaching in the area of Computer Science (Mobile and Ubiquitous Computing). Currently, he is also a member of the ALGORITMI research group at University of Minho.

Lino Figueiredo received the Ph.D degree in Electrical and Computer Engineering from the Faculty of Engineering, University of Porto, Portugal, in 2005. He is a Professor with the Electrical Engineering Department, Institute of Engineering of Porto and member of the Knowledge Engineering and Decision Support (GECAD) research group. His current research interests include ambient intelligence, Wireless intelligent sensors, Simulation and modeling of traffic systems and intelligent transportation systems.

Ana Almeida obtained a PhD in Production end Systems Engineering in 2003, from University of Minho. Her research interests include AI applications, Intelligent Systems, Mobile Systems and Affective Computing. She is a member of the Knowledge Engineering and Decision Support (GECAD) research group. She has participated in more than fifteen projects and integrated the organizing/program/scientific committee of several scientific conferences.

Paulo Novais is an Associate Professor with Habilitation of Computer Science at the Department of Informatics, in the School of Engineering of the University of Minho (Portugal) and a researcher at the ALGORITMI Centre in which he is the coordinator of the research group Intelligent Systems Lab, and the coordinator of the research line in "Ambient intelligence for well-being and Health Applications".

António Meireles is a PhD student in Electrical Engineering and Signal Processing at University of Aveiro and researcher at research group GECAD. He has a Master degree in Electronics Engineering from ISEP, a post-graduation in Biomedical Engineering from University of Porto and management from Porto Business School. During his career he was project leader in different research and development projects and semiconductor development. His main interests are electronics and digital signal processing applied to biomedical devices.

Agile values and their implementation in practice

Eva-Maria Schön[1], Maria J. Escalona[1], Jörg Thomaschewski[2]

[1]*University of Seville, Spain*
[2]*University of Applied Sciences Emden/Leer, Germany*

***Abstract* — Today agile approaches are often used for the development of digital products. Since their development in the 90s, Agile Methodologies, such as Scrum and Extreme Programming, have evolved. Team collaboration is strongly influenced by the values and principles of the Agile Manifesto. The values and principles described in the Agile Manifesto support the optimization of the development process. In this article, the current operation is analyzed in Agile Product Development Processes. Both, the cooperation in the project team and the understanding of the roles and tasks will be analyzed. The results are set in relation to the best practices of Agile Methodologies. A quantitative questionnaire related to best practices in Agile Product Development was developed. The study was carried out with 175 interdisciplinary participants from the IT industry. For the evaluation of the results, 93 participants were included who have expertise in the subject area Agile Methodologies. On one hand, it is shown that the collaborative development of product-related ideas brings benefits. On the other hand, it is investigated which effect a good understanding of the product has on decisions made during the implementation. Furthermore, the skillset of product managers, the use of pair programming, and the advantages of cross-functional teams are analyzed.**

***Keywords* — Agile Development, Agile Values, Scrum, Cross Functional Teams, Human Needs.**

I. Introduction

Agile Methodologies are commonly used in our time [1]. Compared to traditional process models (e.g., waterfall model [2]) they promise benefits such as on-time delivery and customer satisfaction [3]. This is based on the assumption that the scope of the product to be developed is not yet fully defined at the beginning of the process and a response to changes is more important than following a fixed plan [4]. Therefore, in the application of agile approaches, there is no initial specification that describes all requirements up to the smallest detail. Requirements are often documented in form of User Stories [5] or Persona Stories [6] [7]. Those will be developed iteratively during the development process. With increasing progress in the development process, the scope of the product becomes clearer. For this purpose, the product, based on a defined vision, is developed iteratively at the beginning [8] [6] [9]. Such a flexible approach has the advantage that the knowledge, which is gained by the project members during the development process of the product, may influence the product development in a positive way. Thus, a product can be developed to meet the expectations of users and other stakeholders.

In literature, many case studies describe how agile approaches can be optimized for practical use [10]. Here, on one hand useful tools and on the other hand the integration of different approaches from other domains are described. In particular, the combination of Human Computer Interaction and Agile Methodologies shows a variety of the best available practices [11] [12] [13]. These kinds of hybrid approaches are often used for development of products in the field of Interactive Multimedia (e.g. eLearning tools [36], consumer products, digital services). As user interaction plays an important role for these products, user participation during development is necessary in order to develop products with a good user experience. In practice these hybrid approaches are often based on recommendations of the authors and are not validated experimentally. Further empirical research is therefore appropriate as, inter alia, used for Silva da Silva et al. [10].

This article focuses on the cooperation in an agile team as well as the understanding of roles and responsibilities. To this end, best practices were reviewed with a questionnaire study. The main contribution of this article is to give optimization for agile approaches based on validated theses. The contents are aimed at both agile practitioners who are interested in quantitative statements about theses - based on their daily work, as well as to the management that wants to adopt Agile Methodologies and faces the challenge of creating existing conditions.

First, a brief overview of the emergence of agile values and principles is given. Following that, the research objectives and the methodology used for the analysis are described. Subsequently, the study and its implementation are discussed. Finally, the results and their conclusions for agile product development processes are debated.

II. Agile software development

Already in the mid-80s, it has been shown that a sequential approach to product development is not well suited due to the lack of flexibility [14]. Thus, in addition to the traditional process models, such as the waterfall model [2], new process models have been developed. For one thing, these are iterative process models such as Rational Unified Process [15], for the other it is about Agile Methodologies such as Scrum [16], Extreme Programming [17] and Feature-Driven Development [18]. Even, some initiatives are bringing together classical methodologies or tendency with agile principles, like approach presented in [19] or in [20]. In particular, agile approaches bring a high level of flexibility, which has not been there previously, and are suitable for the development of complex products [21]. Their application becomes widely spread nowadays, with Scrum playing an important role [1].

In 2001, the *Agile Alliance* [4] created the *Manifesto for Agile Software Development*. The *Agile Manifesto* includes values and principles that help to optimize the software development process. Most of the principles even play an important role in today´s agile community [22].

The values and principles provide no rules, but rather describe the attitude of the Agile Methodologies that should be used. They follow the aim that communication among those involved in the project and reactions to changes are in the foreground. In particular, the relations between the people involved in the process are underlined as very

DOI: 10.9781/ijimai.2015.358

important *("Individuals and interactions over Processes and Tools"* [4]).

Another important principle is named self-organizing teams [4]. The teams are supported by the organization - in which they operate - for the time of the execution of their tasks. It is not prescribed how they implement their tasks [16]. Their environment places confidence in the teams to own the skills to implement their tasks in a self-organized way [4]. This type of work can lead to a greater satisfaction among those people who are involved. Satisfaction and positive experiences can be the fulfillment of *psychological needs* [23]. Hassenzahl et al. [24] describe that the fulfillment of psychological needs (e.g., *competence, relatedness, popularity, stimulation, security*) lead to a state of *well-being*. In the self-organized work, the need for autonomy and competence is satisfied and thus the state of well-being occurs.

In the context of Agile Methodologies, specific methods are often used. *Pair programming* is a best practice that has its origins in the agile software development [17]. In pair programming, two developers work together on the same task. Williams et al. [25] describe that the results produced with pair programming have better quality and time to market is reduced. Furthermore, they noted that the developers had more fun working together on the problem-solving process. This can be attributed to the fact that the developers feel safer by the 4-eyes principle during the execution of their tasks. The collaborative development of the solution gives the developers a sense of security. In this way, the human desire for security is fulfilled, and reaches the state of well-being. However, the use of pair programming should always be seen in the context of people working with it. In some cases, pair programming can be perceived as extremely inefficient, very exhausting, and as a waste of time [26].

III. Study

In the following the research objectives, the study design and the implementation are discussed.

A. Research objectives and methodology

The aim of this study is to examine current ways of working in product development using agile approaches from trenches. Both, the cooperation in the project team and the understanding of the roles and tasks need to be analyzed. First theses are prepared for this project, which are then verified by a questionnaire study. The theses have been formulated on the basis of assumptions that are often encountered in practice.

The following six hypotheses are investigated[1]:

(1) The collaborative development of product-related ideas has the advantage that the team develops a better understanding of the product.

(2) A good understanding of the product helps the developers to make better decisions during implementation.

(3) The product manager should have the ability to create a rudimentary concept of the product, which is then elaborated in more detail.

(4) The concept of pair programming can also be transferred to the creation of a design concept.

(5) The concept of pair programming can also be transferred to the implementation of quality assurance measures.

(6) In project teams composed of members with different professional backgrounds, team members learn more than in teams composed of members from the same field.

In the original study, further theses have been examined. The complete questionnaire can be found in Schön [27].

[1] The original questions were written in German (see Appendix A).

B. Construction of the questionnaire

In the design of the questionnaire appropriate guidelines have been used for the design of good online questionnaires [28] [29]. It is important to keep the amount of items small, because the response rate is higher for short questionnaires compared to long questionnaires. In addition, long questionnaires usually have a higher dropout rate for episode [30]. The questionnaire was written in German and is divided into three sections: introduction, body and conclusion. The preface contains the instruction objective of the survey, privacy, duration, contact information, and some questions about the differentiation of the target group. The main section includes items formulated to verify the propositions. The final part consists of open questions and a final page, by thanking the participants and contact information for questions and suggestions.

In a pretest of the questionnaire, it was tested and revised in five iterations by various test persons of the target group. For this purpose, qualitative interviews were carried out with these people. Those were asked to answer the questionnaire. Thus, the idea of the test persons could be collected for analysis; the method *Think Aloud* [31] has been applied. Based on the pretest, the average time to answer the questions was set up to 10-15 min. The willingness to complete a questionnaire with this period is relatively high, compared temporally to elaborate ones [32].

C. Implementation of the study

To carry out the study, an online survey with the survey tool *Limesurvey*[2] was placed. The online survey was conducted during the period 2014-03-18 – 2014-04-08 (duration of three weeks). The target group of the survey has been selected from the IT industry with expertise in the subject area Agile Methodologies. Participants were recruited through personal networks, entries in thematically relevant groups in social media networks, and in forums of the university network *oncampus*[3]. The study has been carried out within an interdisciplinary group of participants. This has the advantage that the theses are evaluated from different perspectives. Information regarding the professional experience of the participants is shown in Table 1. In addition, Table 2 shows the company type of the participants.

Overall, the questionnaire was been filled out 175 times. Of these, 129 questionnaires were completed (dropout rate = 26.28%). 98% of those stopped after the questions of the introductory part. This leads to the assumption that these participants did not feel addressed as a target group.

The results of 129 completed questionnaires were filtered for analysis in order to obtain the answers of the participants who have already had practical experience with agile approaches. For the evaluation of the results, the participants were included who have already used an agile approach (N = 93). The key aspects of the participants - in the last two years - were wide-ranging and covered the following subject areas[4]: project management (39), software architecture (30), quality assurance (26), back-end development (25), front-end development (23), user experience design (19), infrastructure (7), technical sales (7) and operations (6).

TABLE I
PROFESSIONAL EXPERIENCE (N=93)

Experience	Answers
Young Professional (less than 1 year)	3.23%
Professional (between 1-8 years)	47.31%
Senior (more than 8 years)	49.46%

[2] www.limesurvey.org
[3] www.oncampus.de
[4] Participants had the opportunity to make multiple choices

TABLE II
TYPE OF COMPANY (N=93)

Company type	Answers
Service provider	64.52%
Product manufacturer	25.81%
Freelancer	7.53%
Other	2.15%

IV. Results

To capture the personal opinion 5-point Likert items were used. There are the following gradations *totally agree, tend to agree, neutral, disagree, and totally disagree*. In one question, the participants also had the opportunity to make no statement at all.

The survey results that are considered for the evaluation come from the participants who have already used an agile approach and who have worked with Scrum (N=93) in the last two years. Two questions on *pair programming* have been shown only to those participants who have confirmed that they have a notion of *pair programming*. As a result there is a smaller sample of these items (N=81).

A. Presentation of results

For clarity, the results are first presented in tabular form (Table 3). For this purpose, the two positive responses (*totally agree, tend to agree*) are counted as the sum of the theses. Subsequently, each item will be considered in detail.

In the following figures (Figure 1 – 5), the results (see Table 3) of the questionnaire are shown in detail.

TABLE III
OVERVIEW OF THE STATEMENTS

Item	Theses	Agree-ment	Number of Participants
1	The collaborative development of product-related ideas has the advantage that the team develops a better understanding of the product.	91%	N=93
2	A good understanding of the product helps the developers to make better decisions during implementation.	94%	N=93
3	The product manager should have the ability to create a rudimentary concept of the product, which is then elaborated in more detail.	86%	N=93
4	The concept of pair programming can also be transferred to the creation of a design concept.	76%	N=81
5	The concept of pair programming can also be transferred to the implementation of quality assurance measures.	70%	N=81
6	In project teams composed of members with different professional backgrounds, team members learn more than in teams composed of members from the same field.	80%	N=93

Collaborative development (Item 1)

The collaborative development of product-related ideas has the advantage that the team develops a better understanding of the product.

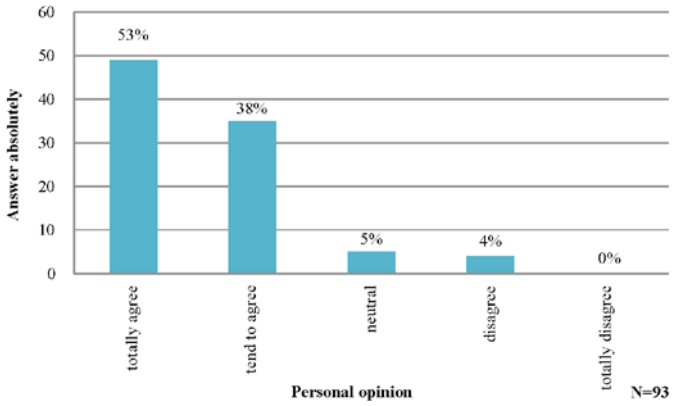


Figure 1: Collaborative development

91% of the sample N=93 agreed to this statement. The high agreement makes clear that it is considered useful in practice to involve the team in the ideation process.

Good understanding (Item 2)

A good understanding of the product helps the developers to make better decisions during implementation.

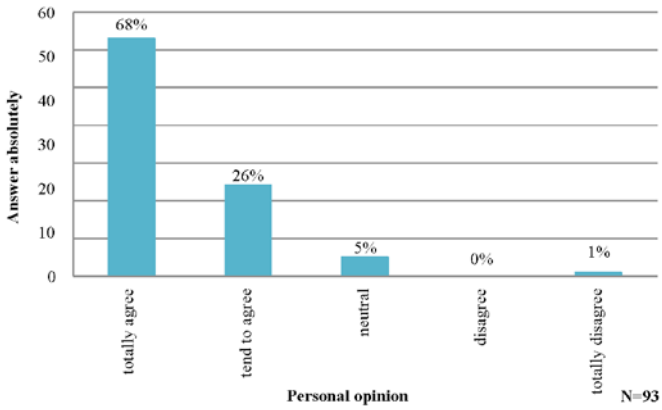


Figure 2: Understanding the product helps during implementation

This item received the highest popularity by the participants. 94% of the sample (N=93) agreed with this statement. Therefore, the developers can develop a good understanding of the product, it is important to include them in the ideation process (see Figure 1). In addition, visual artifacts, such as e.g. a sketch support the communication process and contribute to a better understanding [33].

Skills of the product managers (Item 3)

The product manager should have the ability to create a rudimentary concept of the product, which is then elaborated in more detail.

With this item we have examined the skillset of the product managers. 86% agreed to the sample N=93. In the selection of the person for the role of the product manager, it should be ensured that this person is able to develop a rudimentary concept of the product.

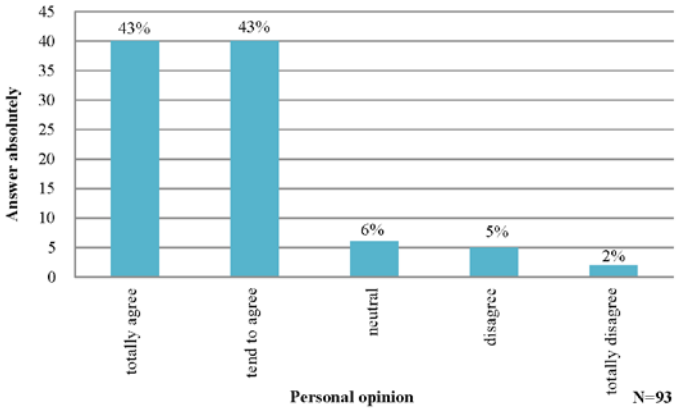


Figure 3: Skillset of product managers

Pair Concepting und Pair Testing (Item 4, 5)

Item 4 and Item 5 are displayed only to those participants who have confirmed that they have a notion of pair programming. This results in a sample of N=81. 76% of the participants agreed to item 4, and 70% agreed to item 5. The concept of pair programming can therefore not only be used in programming, but can also optimize the operation in other domains, such as in the conceptual design (Item 4) or quality assurance (Item 5).

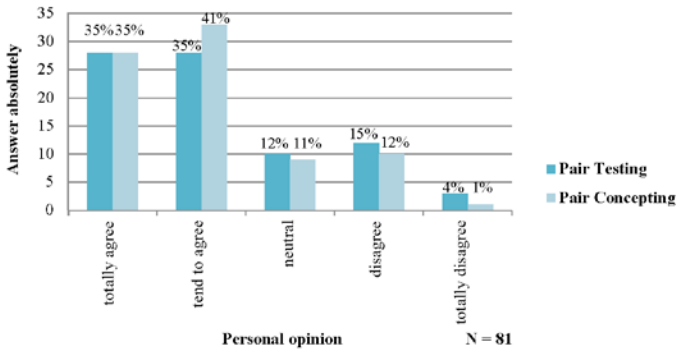


Figure 4: Pair Concepting and Testing

Since the concept of pair programming is closely linked to programming and which has not yet been established in the fields of design and testing, the agreement probably does not refer to its own experience but also reflects the willingness, the concept promising in these two domains to use.

Cross functional teams (Item 6)

In project teams composed of members with different professional backgrounds, team members learn more than in teams composed of members from the same field.

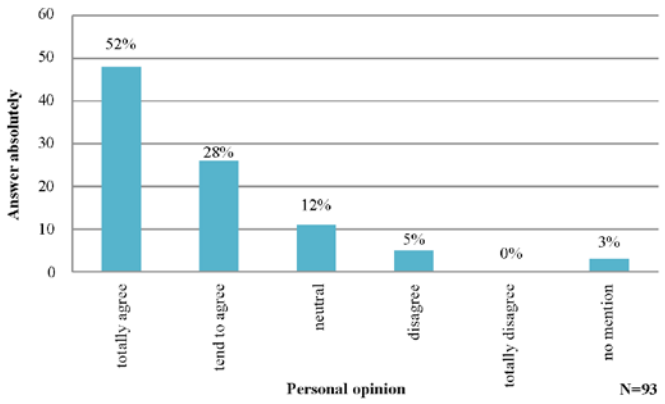


Figure 5: Cross functional teams

In addition to the 5-point Likert scale, participants had the opportunity to make no mention of this statement, it was used by 3 participants (3%). 80% of the sample, N=93 agreed to this statement.

B. Conclusion for agile product development processes

The survey clearly supports all of the six theses. The results thus confirm the assumptions that are made in practice. In addition, they provide important insights for Agile Product Development Processes. It has been shown that the collaborative development of product-related ideas contribute to a better understanding of the product (see, Item 1). For Agile Development Processes this entails that a collaborative ideation process should take place as early as possible in the development process. Here, it is advantageous to include developers, because a good understanding of the product helps them to make better decisions during implementation (see, Item 2). Furthermore, the results show that the product manager should have the ability to create a rudimentary concept of the product (see Figure 3). This finding is very significant for the selection of a suitable candidate (in Scrum product owner) who takes over the role of the product manager. In practice, the product owner is often supported by a product ownership team [34] in carrying out one`s tasks. The team may consist, for example of a business analyst, a user experience designer, and a lead developer. These collaboratively develop in an iterative product discovery requirements (see Figure 6). The product discovery is used to define the strategic direction of the product and to evaluate different ideas [35].

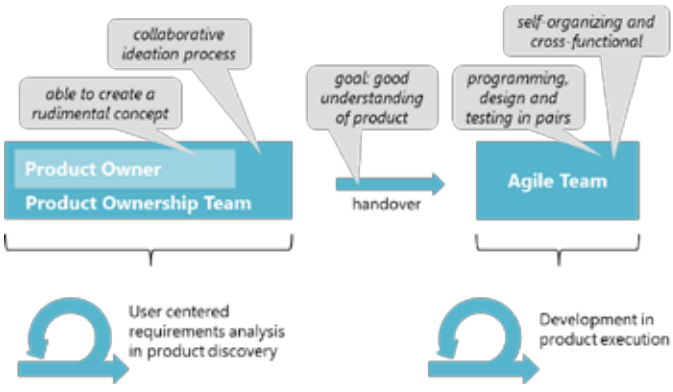


Figure 6: Implications for an Agile Product Development Process

The method pair programming shows that best practices of Agile Product Development promote cooperation in the agile team. The results of the survey show that the concept of pair programming is also seen on the creation of a design concept and the implementation of

quality assurance measures as appropriate (see Item 4, Item 5). Thus, the advantages of the method can be used across domains. Furthermore, cross-functional teams have the advantage that the team members learn more from each other than in purely functional teams (see Figure 6).

V. Conclusion

In this article, the current operation has been studied in agile product development processes. For this purpose, an empirical study was conducted. By means of the analysis of collaboration in the agile team of the understanding of roles and responsibilities has occurred. With the results, existing best practices of agile product development could be confirmed. In addition, qualitative optimization for agile product development processes could be derived. The optimization applies in particular for middle-sized projects, where the user interface plays an important role. Apart from that, the size of the company is irrelevant. Another significant point is that scaling agile is not considered in the aim of this study. The aim of this study was to analyze the collaborative work on team level.

The use of process models and methods should be evaluated with focus on the people affected. Compared to traditional approaches, agile approaches base the values and principles of the Agile Manifesto. If the people concerned do not practice these, the success in the implementation of these approaches fail.

The relationship between agile values with regard to the fulfillment of psychological needs can be further investigated in future studies. For this purpose, best practices such as pair programming can be used. In addition, the study has been carried out in the German-speaking area. It is also possible to conduct an international study, because of the wide spreaded use of agile approaches.

Appendix A

In the following the original German questions are listed:

(1) Die gemeinsame Entwicklung von produktbezogenen Ideen hat den Vorteil, dass das Team ein besseres Verständnis vom Produkt entwickelt.

(2) Ein gutes Verständnis vom Produkt hilft den Entwicklern dabei, bessere Entscheidungen während der Implementierung zu treffen.

(3) Der Produktverantwortliche sollte die Fähigkeit besitzen ein rudimentäres Konzept vom Produkt zu erstellen, welches anschließend detaillierter ausgearbeitet wird.

(4) Das Konzept vom Pair Programming lässt sich ebenfalls auf das Erstellen eines Design-Konzeptes übertragen.

(5) Das Konzept vom Pair Programming lässt sich ebenfalls auf die Durchführung von qualitätssichernden Maßnahmen übertragen.

(6) In Projektteams, die sich aus Mitgliedern mit unterschiedlichem fachlichem Hintergrund zusammensetzen, lernen die Teammitglieder mehr voneinander als in Teams, die aus Mitgliedern derselben Fachrichtung bestehen.

Acknowledgment

This research has been supported by MeGUS Project (TIN2013-46928-C3-3-R) of the Ministerio de Ciencia e Innovación, Spain.

References

[1] Komus, A., Kuberg, M., Atinc, C., Franner, L., Friedrich, F., Lang, T., Makarova, A., Reimer, D., Pabst, J. (2014) Status Quo Agile 2014 www.status-quo-agile.de.

[2] Royce, W. (1970) Managing the Development of Large Software Systems

[3] Dybå, T., Dingsøyr, T. (2008) Empirical studies of agile software development: A systematic review, Information and Software Technology 50, 833–859

[4] Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R., Mellor, S., Schwaber, K., Sutherland, J., Thomas, D. (2001) Manifesto for Agile Software Development, www.agilemanifesto.org. Accessed 10 August 2015

[5] Cohn, M. (2004) User stories applied For agile software development. Addison-Wesley signature series, Addison-Wesley, Boston

[6] Winter, D., Holt, E.-M., Thomaschewski, J. (2012) Persona driven agile development Build up a vision with personas, sketches and persona driven user stories, Proceedings of the 7th Conference on Information Systems and Technologies (CISTI)

[7] Hudson, W. (2013) User stories don't help users, Interactions 20, 50–53

[8] Schwaber, K. (2004) Agile project management with Scrum, Microsoft Press, Redmond, Wash.

[9] Holt, E.-M., Winter, D., Thomaschewski, J. (2012) Von der Idee zum Prototypen Werkzeuge für die agile Welt. In Usability Professionals 2012. German UPA e.V., Stuttgart

[10] Silva da Silva, T., Martin, A., Maurer, F., Silveira, M. (2011) User-Centered Design and Agile Methods: A Systematic Review. In 2011 AGILE Conference, 77–86.

[11] Beyer, H. (2010) User-centered agile methods. Synthesis lectures on human-centered informatics #10, Morgan & Claypool Publishers, San Rafael, Calif.

[12] Patton, J. (2008) Twelve (12) emerging best practice for adding user experience work to agile software development, http://agileproductdesign.com/blog/emerging_best_agile_ux_practice.html. Accessed 10 August 2014

[13] Sy, D. (2007) Adapting usability investigations for agile user-centered design, Journal of usability studies 2, 112–132.

[14] Takeuchi, H., Nonaka, I. (1986) The New New Product Development Game, Harvard Business Review

[15] Kruchten, P. (2003) The rational unified process An introduction. The Addison-Wesley object technology series, Addison-Wesley, Boston

[16] Sutherland, J., Schwaber, K. (2013) The Scrum Guide The Definitive Guide to Scrum : The Rules of the Game, www.scrum.org/Portals/0/Documents/Scrum%20Guides/2013/Scrum-Guide.pdf. Accessed 10 August 2014

[17] Beck, K., Andres, C. (2005) Extreme programming explained Embrace change, Addison-Wesley, Boston, MA

[18] Palmer, S. R., Felsing, J. M. (2002) A practical guide to feature-driven development. The Coad series, Prentice Hall PTR, Upper Saddle River, NJ

[19] Torrecilla-Salinas, C. J., Sedeño, J., Escalona, M. J., Mejías, M. (2015) Estimating, planning and managing Agile Web development projects under a value-based perspective. Information and Software Technology, 61, 124-144

[20] Ros, J.J. (2015) BIMODAL IT. El arte de trabajar a dos velocidades. III Jornadas is TMF. Seville, Spain.

[21] http://itsmf.es/index.php?option=com_content&view=article&id=1696. Accessed 10 July 2015

[22] Schwaber, K. (1997) SCRUM Development Process, in OOPSLA Business Object Design and Implementation Workshop, Sutherland, J., Casanave, C., Miller, J., Patel, P., Hollowell, G., Eds. Springer London, London, 117–134

[23] Williams, L. (2012) What Agile Teams Think of Agile Principles, Communications of the ACM Volume 55 Issue 4, 71-76

[24] Sheldon, K. M., Elliot, A. J., Kim, Y., Kasser, T. (2001) What is satisfying about satisfying events? Testing 10 candidate psychological needs, Journal of Personality and Social Psychology 80, 325–339

[25] Hassenzahl, M., Diefenbach, S. (2012) Well-being, need fulfillment, and Experience Design, DIS 2012 – June 11-12, Newcastle (UK)

[26] Williams, L., Kessler, R. R., Cunningham, W., Jeffries, R. (2000) Strengthening the case for pair programming, IEEE Softw. 17, 19–25

[27] Tessem, B. (2003) Experiences in Learning XP Practices: A Qualitative Study, Goos, G., Hartmanis, J., van Leeuwen, J., Marchesi, M., Succi, G.,

Eds. Springer Berlin Heidelberg, Berlin, Heidelberg, 131–137

[28] Schön, E.-M. (2014) Menschzentriertes Vorgehensmodell für einen agilen Produktentwicklungsprozess, Masterthesis, HS Emden/Leer

[29] Gräf, L. (2002) Assessing Internet Questionnaires: The Online Pretest Lab, Batinic, B., Reips, U.-D., Bosnjak, M., Eds. Hogrefe & Huber Publishers, Seattle, 73–93

[30] Bortz, J., Bortz-Döring, Döring, N. (2009) Forschungsmethoden und Evaluation Für Human- und Sozialwissenschaftler; mit 87 Tabellen. Springer-Lehrbuch, Springer-Medizin-Verl., Heidelberg

[31] Tuten, T. L., Urban, D. J., Bosnjak, M. (2002) Internet Surveys and Data Quality: A Review, Batinic, B., Reips, U.-D., Bosnjak, M., Eds. Hogrefe & Huber Publishers, Seattle, 7–27

[32] van Someren, M. W., Barnard, Y. F., Sandberg, J. A. (1994) The think aloud method A practical guide to modelling cognitive processes. Knowledge-based systems, Academic Press, London, San Diego

[33] Bosnjak M., Batinic, B. (2002) Understanding the Willingness to Participate in Online-Surveys – The Case of e-mail Questionnaires, Batinic, B., Reips, U.-D., Bosnjak, M., Eds. Hogrefe & Huber Publishers, Seattle, 111–116

[34] Buxton, B. (2008) Sketching user experiences Getting the design right and the right design, Morgan Kaufmann, Amsterdam

[35] Patton, J. (2009) Becoming a Passionate Product Owner, A Certified Scrum Product Owner Course, www.agileproductdesign.com/training/passionate_product_owner.html. Accessed 10 August 2014

[36] Cagan, M. (2008) Inspired How to create products customers love, SVPG Press, Sunnyvale, Calif.

[37] Cortés, J.A., Lozano, J.O. (2014) Social Networks as Learning Environments for Higher Education, in IJIMAI: International Journal of Interactive Multimedia and Artificial Intelligence, Special issue on Multisensor User Tracking and Analytics to Improve Education and other Application Fields, Vol. 2, No. 7, 63-69

Eva-Maria Schön received her MSc in Computer Science and Media Applications at the University of Applied Science Emden/Leer in 2014. She is a PhD student at the University of Seville (Spain) and also works as a Lead Consultant at CGI (Germany). She focuses on agile product management and human centered design. Her research interests are agile software development, requirements engineering and human computer interaction.

María José Escalona Cuaresma received the PhD degree in computer science from the University of Seville, Spain, in 2004. Since 1999, she has been a lecturer and researcher in the Department of Computer Languages and Systems at the University of Seville. She is the director of the Web Engineering and Early Testing research group. Her current research interests include the areas of requirement engineering, web system development, model-driven engineering, early testing and quality assurance. She also collaborates with public companies like Consejeria de Cultura and Servicio Andaluz de Salud in quality assurance.

Jörg Thomaschewski became full professor at the University of Applied Sciences Emden/Leer, Germany from September 2000. His research interests are in the fields “Internet Applications” focusing on human computer interaction, e-learning and software engineering. He is author of various online modules, e.g. “Human Computer Communication”, which is used in the Virtual University (Online) at six university sites. He has wide experience in usability training, analysis and consulting.

A network based methodology to reveal patterns in knowledge transfer

Orlando López-Cruz, Nelson Obregón N.

[1] *Universidad El Bosque, Bogotá,*
[2]*Universidad Javeriana, Bogotá*

***Abstract* — This paper motivates, presents and demonstrates in use a methodology based in complex network analysis to support research aimed at identification of sources in the process of knowledge transfer at the interorganizational level. The importance of this methodology is that it states a unified model to reveal knowledge sharing patterns and to compare results from multiple researches on data from different periods of time and different sectors of the economy. This methodology does not address the underlying statistical processes. To do this, national statistics departments (NSD) provide documents and tools at their websites. But this proposal provides a guide to model information inferences gathered from data processing revealing links between sources and recipients of knowledge being transferred and that the recipient detects as main source to new knowledge creation. Some national statistics departments set as objective for these surveys the characterization of innovation dynamics in firms and to analyze the use of public support instruments. From this characterization scholars conduct different researches. Measures of dimensions of the network composed by manufacturing firms and other organizations conform the base to inquiry the structure that emerges from taking ideas from other organizations to incept innovations. These two sets of data are actors of a two-mode-network. The link between two actors (network nodes, one acting as the source of the idea. The second one acting as the destination) comes from organizations or events organized by organizations that "provide" ideas to other group of firms. The resulting demonstrated design satisfies the objective of being a methodological model to identify sources in knowledge transfer of knowledge effectively used in innovation.**

***Keywords* — Knowledge Transfer; Technological Innovation; Technology Transfer; Social Networks Analysis.**

I. Introduction

THIS paper is intended to introduce and show the application of a methodology to identify the underlying structure of the process of knowledge transfer at the inter-organizational level, as one of the main resources to create innovation.

Even the search for methods to understand and design effective processes to incept innovation is an active thread in contemporary research, the dominant research paradigms used to inquiry on knowledge and technology transfer (TT) continue to be those traditional descriptive research methods of the natural and social sciences.

The strength of the proposed methodology is its simplicity. Even that, it seeks to provide a unified model to reveal knowledge sharing patterns about different periods of time and different sectors of the economy.

This methodology complements the statistical methodology provided by the NSD on SITD and SITDS surveys, by unify the way to identify sources and recipients links when knowledge is needed to be transferred procuring to generate innovation in firms.

To present the results of the application of the proposed methodology, this paper is structured as follows: Section II introduces the concepts of knowledge and knowledge sharing. Section III describes SITD and SITDS as the origin of raw data to search for structures of knowledge transfer. Section IV explains the reason why when investigating on technology transfer is unavoidable to address the technological knowledge transfer topic. Section V introduces a categorization of technology transfer models which resulted from a literature review.

In section VI, the mathematical foundations for representation and analysis of social networks are introduced. Sections VII and VIII describe the proposed methodology to be demonstrated in use. Section IX shows the case of the manufacturing sector where the proposed methodology is demonstrated in use. The results are shown in Section X, and discussion and conclusion are presented in sections XI and XII. Finally, the references used in the preparation of the paper are listed.

II. Nature of Knowledge Sharing

A. Knowledge

It is understood that when innovation is a desirable outcome, knowledge is a critical supply to organizations [1, 2]. Knowledge is a multidimensional concept that has been studied since ancient Greeks, and becomes relevant in modern organizations specially when data processing and information systems have not explained nor commanded organizations viability. The role of knowledge in organizations emerged as a key concept over data and information [3].

Under the Shannon and Weaver's communication paradigm [4], there are three levels of communication complexity: syntactic, semantic and pragmatic. These allow to describe knowledge across three organization boundaries: information- processing boundary at a syntax level, interpretive boundary, and pragmatic boundary [5]. These, in turn, allow transferring, translating and transforming knowledge [5]. These determines a structure composed by the link between organizations sharing knowledge and organization.

In this context, knowledge is shared at three different non-exclusive levels. While data are conceived as a set of facts or a symbolic record of facts, without interpretation, information is understood as those data in a context with a sender and a receiver [6]. When available information conducts to both comprehension and action, on the context of that information, emerges knowledge [7]. Knowledge is not tradable (i.e. knowledge may not be subject of a sale transaction), therefore it must be "thought" [8] and learned.

B. Knowledge sharing

In this work, knowledge sharing is the act by which humans share knowledge in a community or set of organizations [8]. This does not imply nothing about the intention to get economic benefits or satisfy

DOI: 10.9781/ijimai.2015.359

individual interests.

In this context, knowledge transfer, is a possibility when sharing knowledge, but it depends on conditions of the context where knowledge appears, conditions of the context where may appear, and a possible link between those two contexts.

III. Data to search for sharing structures

In order to search for structures enabling knowledge sharing, raw data gathered by the National Department of Ststistics of Colombia (NDS) was used. This Department is the head of the National Statistics System which is integrated by governmental organizations and autonomous organizations, according to Colombia national regulations and laws.

By 1996, the NDS conducted the first *national survey on technological development* (STD). It was conducted on firms of the manufacturing sector of the economy (Fig.1). The second survey was conducted nine years later, in 2005.

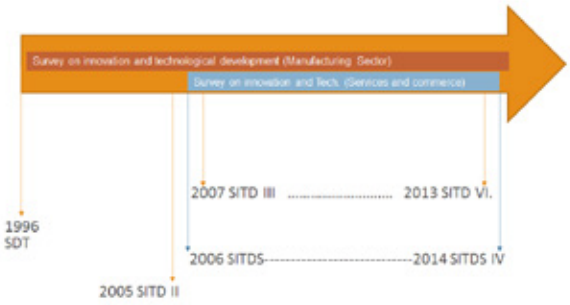


Fig. 1. Timeline showing the years when the survey on innovation and development technology for the manufacturing sector (SIDT) and the corresponding survey for the services and commerce sector (SITDS) has been conducted. The very first SITD was just a survey on development technology (SDT) in 1996. SITD and SITDS are conducted every two years. SITD in odd years and SITDS in even years.

Since 2005, the survey is known as the *survey on innovation and technological development* (SITD) and is being conducted every odd year and data gathered correspond to the two preceding years. For instance, SITD-2015 gathers data from 2013 and 2014.

SITD is characterized by a survey on the manufacturing sector of the economy. Therefore, since 2006 the NDS conducts an additional survey to gather data from the sector of services and commerce. The national *survey on innovation and technological development-services* (SITDS), conducted by NDS every even year since 2006, gathers data corresponding to two years preceding the survey. This means that SITDS-2014 is composed by data from 2012 and 2013.

SITD as well as SITDS aim for a (statistically) characterization of innovation dynamics of firms. In addition, both of them aim for an analysis of the usage of public instruments at the production and services sectors of the economy.

IV. Economic development: From technology transfer to knowledge transfer

In economics, "development" refers to the process of improvement of the economic, political, and social well-being of people [9-12]. Not all societies sustains the same development rhythm and economies experiment cyclic behaviors [13-15]. This leads to different levels or stages of economic development of countries.

In order to continue catching up with developed countries, less developed countries import technology. In turn, organizations in recipient countries face the challenge to incorporate imported technology to their production processes in order to increase productivity.

But, even though technology were a commodity, transference is not a gentle process [16, 17]. First, recipients must choose where to take technology from or where to get ideas to improve their processes. Then, find out the way to incorporate technology to the production processes. In this way, TT needs to be more than movement of physical assets.

Accepting that TT is the process by which commercial technology, a technological innovation and the related knowledge is disseminated [18, 19], then the issue is what is understood as technology.

Technology include tools, techniques, materials and power resources developed by humans to achieve their goals [20, 21]. Technology may include physical artifacts, but ever includes knowledge [21]. TT is valuable to increase productivity in so far as transferred knowledge allows to modify production processes. This has been an issue for economists [22] but on a retrospective or an explanatory point of view.

V. Technological knowledge Transfer categories

A. Technological knowledge transfer

Transferring technology is not achieved without transferring associated knowledge [23]. Knowledge most not be reified, this means that knowledge should not be understood as an object outside individuals and their environment as a historical and cultural context. Knowledge is different form information and data: knowledge is a human act [8] related to a human process inside human being's mind [24], therefore, knowledge appears at the orbit of personal thoughts and experiences [25, 26].

Technological artifacts are not just the physical object. Besides any material instantiation of technology, they include mental models [8] that make sense the usage of the artifact by a community. Consequently, any TT must be thought as knowledge transfer also.

B. A categorization of technological knowledge transfer

Some authors introduce categories of TT. A four-categories model [27] of TT, according to strengths, limitations and focusing consists of 1) a category of traditional models, composed by appropriability, dissemination, and communication models, 2) 1990's models including the Gibson and Smilor model [28], Sung and Gibson model [29], and Rebentish and Ferreti model [30], 3) Knowledge-based model [27, 31, 32], and 4) organizational learning models [33] (Table I). The last two categories may be thought as two sides of the same coin. However, in this paper are set apart.

Traditional models are reductionist (Table I). In the appropriability model, the technology developer assumes a passive role: quality technologies sell themselves. It is linear because states as in imperative for research the production of technological developments that are realized in the market.

The dissemination model focuses on diffusion [19]. In this model, knowledge freely flows from the expert to the non-expert user. It is supposed that a researcher lacks of prejudices to disclose his innovations and the potential user is eager to know innovations. Linearity comes from the unidirectional communication to the user, who appears to be aligned at the end of the TT process as "final user".

In contrast, the usage model of knowledge includes a three.-way communication between researcher, developer and user of the technology. This model does not explain TT beyond organization boundaries.

The communication model perceives TT as an on-going process where individuals share ideas. This model follows the network

paradigm where feedback is ubiquitous. Knowledge in this model is understood as an independent object, universally applicable, similar to scientific knowledge.

Models of TT in the 1990's category (Table I) are focused to solve limitations in traditional models. In the Gibson and Smilor model [28] TT is a passive process defined by three levels of involvement, from the relationship researcher-user point of view: (I) technology development, (II) technology acceptance, and (III) technology application. It emphasizes in level I, taking market pressures and research quality causing TT. This makes this model similar to the appropriability model.

TABLE I
CATEGORIES FOR TECHNOLOGICAL KNOWLEDGE TRANSFER MODELS [a]

Category	Model	Time
Traditional	Appropriability	1945-1950
	Dissemination	1960-1970
	Knowledge usage	1980's
	Communication	1980-1991
1990's	Gibson & Smilor	1991
	Rebentisch & Ferreti	1995
	Sung & Gibson	2000
Knowledge-based vision	KBV	1985-2000
Organizational-Learning based vision	OL	1991-2007

[a] Possible models for technological transfer appeared before the 20th century are not included in this table.

The Sung and Gibson model [29] is an enhancement of the Gibson and Smilor model [28] to four levels: (I) Creating knowledge and technology (II) deployment, (III) Implementation, and (IV) Commercialization. Knowledge created at level I spreads through publications, teleconferences, and massive media, not including the user. Therefore, this is a passive process.

The Rebentish and Ferreti model [30] is developed from the point of view of the transferor. It emphasizes knowledge embodied in assets being transferred (explicit knowledge [34, 35]). Even though its focus on the relationship between technology-organizational context, and the organization capacity to adopt new technology, the model is linear biased because omits actors of the TT process.

The Knowledge-Based View (KBV) [2, 36] of TT (Table I) understands firms as sets of knowledge and knowledge in firms produce competitive advantage. The aim is to get a sustainable advantage by means of knowledge. This view conceives knowledge as an intangible resource, hardly transferable on an intra-firm basis. A firm may be modeled as an instrument to transfer and develop knowledge next to other related organizations, not just as a knowledge repository. The model acknowledges insight and individual abilities as (tacit) knowledge that is difficult to assemble and transfer, but that allows to bring together operational issues in learning [37 p.430].

Organizational Learning (OL) [33] (1) as a cognitive process with and without intention, sets up links between actions in the past and in the future, forming a stock of organizational knowledge and memory. (2) as a change in behavior and improvement in organizational effectiveness, make changes operative through modifications in individual, group and firm behavior, and (3) both as cognitive process and change in organizational behavior, OL allows behavior changes to improve organizational performance and to develop new knowledge.

As stated before, KBV and OL are two sides of the same coin. KBV and OL act as a knowledge acquiring unit and as an individual, group and firm behavior modification unit.

VI. Representation of Social Networks

The technical analysis conducted in this study is based in network analysis, a discipline which stands on the innovation of Jacobo Levy Moreno: the sociogram [38]. This was applied in the measurement of the interpersonal relations in small groups, known as sociometry [39] since 1934 [40]. The relations are studied using graph theory. This allows to represent, design, and calculate properties on networks

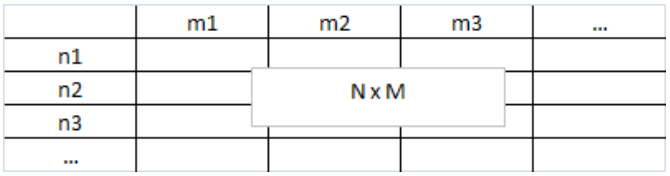


Fig. 2. A two.mode network where the set of nodes n1,n2 ,n3,... may represent the source of information or "ideas" for innovation and the set of nodes m1, m2, m3,… the destination of those information and "ideas".

[41].

A network (a graph) may be understood in its basics as a collection of points (nodes) joined together in pairs by lines (edges). To design a network model [42] the starting point is the measures of the properties of the network (i.e. properties of the phenomenon to be studied or the state to achieve), for instance: number of edges, number of nodes, degree of vertexes, and clustering coefficient. There are many other properties of networks that may be inferred (calculated) [43]. This model is not a single network, but a probability distribution on many networks (i.e. an assemble model).

If the network is to be designed to comply with some characteristics, generative network models are to be used [41]. In the many generative network models, the preferential attachment models seems to be the most adequate to design the growth of complex networks exhibiting power laws. Preferential attachment models may be the Price model, the Barabási and Albert model, the vertexes copying model, and network optimization model.

Newman [41] affirms that Price was inspired by H.A.Simon works [44] In his paper, Simon does not name the distributions, but Price name them as cumulative advantage distributions, that describe the St. Matthew principle (Mt. 25:29).

A. A mathematical representation of a social network data

Social network data is determined by the substantive concern or theories that support the specific study of a network, and is composed by at least one structural variable of a dataset of the phenomena or field in study [40].

The network under study may be composed by one or more datasets, according to the nature of the data (i.e. the different kind of social network entities involved in the study). The number of different datasets composing data of the network is said to be the "mode" of the network [40]. In this study, there are two distinguishable sets of entities: organizations acting as receivers of information and "ideas" conducting to innovation, and organizations acting as source of those information and "ideas" (Table II). These two sets conform actors of the network that are related each other by means of the link of source/destination of that information.

Appears a two-mode (or 2-mode) network in which clearly there are directed relations (source/destination) (Fig. 2).

TABLE II
EXTERNAL SOURCES OF IDEAS GENERATING INNOVATION

No.	Source of ideas a
1	Professional associations or sectorial associations.
2	Scientific and technological databases.
3	Chambers of Commerce.
4	Technological development center.
5	SENA training center.
6	Research centers.
7	Regional productivity center.
8	Clients.
9	Competitor or firm from the same economical sector.
10	Consultants or experts.
11	R&D departments of other firms.
12	Firm from other economic sector.
13	Trade fairs and exhibitions
14	Technology-based companies and incubators.
15	Public institutions.
16	Internet.
17	Books, journals and catalogs.
18	Technical regulations and standards.
19	Technological parks.
20	Suppliers.
21	Seminars and conferences.
22	Copyright information systems.
23	Industrial property information systems.
24	Universities.

Previous order is not relevant and does not bias the analysis.
[a] Lexicographically ordered in the original in Spanish.

The intersection of rows and columns in the matrix (Fig. 2) may represent the strength of the relation between source and destination or just a binary value to represent if there exist (1) or not (0) a relation between nodes n_i and m_j. This is the case in this study where the aim is to identify the sources of knowledge for innovation that some organizations use.

B. Patterns of links because of knowledge sharing

But it is not statistical data the interest of this study but the structures for innovation in an economical sector. Therefore, patterns are to be discovered from data. In order to do that, a concept capturing "where knowledge comes from" is needed.

The concept is the degree of the node. Since the relation represents source/destination of information and "ideas" for innovation, then the number of edges over a node may represent the importance of a node. Since the degree of the node $d(n_i)$ is the number of edges incident to node n_i, when measuring nodal degree for each of the nodes in the network, a pattern will appear.

If X is the sociomatrix of a two-mode network, x_{ij} is an entry of that matrix, and *g* is the total nodes of the directed-graph of a network, the number of edges incident to node n_i from nodes n_j is defined as the indegree (d_I) of the *i*-esim node (n_i), and may be calculated as $d_I(n_i)$ according to (1):

$$d_I(n_i) = \sum_{j=1}^{g} x_j \quad (1)$$

The higher the value of $d_I(n_i)$ the node n_i is more "important" in the network. In this case, node n_i stands as a relevant source of knowledge in a network (of organizations).

C. The problem to be solved

Up to this point, there seems no problem to be solved. Raw data on the relationships between organizations in an economy is provided by a national survey on manufacturing, services and commerce sectors of the economy. A wide theory on inter-organizational knowledge sharing is being developed, and, graph theory provides a mathematical representation of networks and its properties.

However, looking closely to the task to discover patterns of the relationship between organizations needing to innovate and those providing "ideas" for innovation along time, and the process of understanding the complex innovation path followed by organizations in an economy, there are no way to compare the results of different studies, since every researcher develops an *ad-hoc* methodology group [45] or develops a complicated methodology that prevents its usage [46], which reinforces the trend to use specific methodologies. This will not be a problem unless results need to be compared by the industry to make decisions.

Therefore, both industry and researchers lack of an easy to understand and apply, but powerful methodology, to reveal knowledge flows that produce innovations. This proposal intends to fill this gap.

VII. A methodology to reveal patterns in knowledge transfer

A. What 'methodology' means

In simple terms, a methodology is "a system of principles, practices, and procedures applied to a specific branch of knowledge" [47]. From a process view, a methodology is a device specifying a set of steps and restrictions on the transitions between those steps [48, 49]. In addition, a methodology may be understood as the means that conducts to produce the solution to a problem [50]. In the context of this study, a methodology is an artifact that consists of a list of procedures, each one acting stepping stones, the conditions to proceed between procedures, in order to solve a problem. This artifact may be provided to different researchers, in the same context, to produce consistent results.

B. The relevance of "patterns" in (research) methodologies

Patterns have been present in human knowledge since ancient Greeks. In Euclid's Elements, geometric constructions play the role of patterns in mathematical reasoning. Besides, this geometrical patterns leverages reasoning processes to solve problems.

According to these, any problem is identified by three components. (a) In any problem there must be a thing required or desired (the unknown). Without an unknown, there is nothing to look for, here is nothing to seek [51]. In our research problem the unknown is the structure of relations between sources of ideas for innovation and innovators. In addition, (b) in any problem there must be something given. Without any given (known) that serves as a reference, there is nothing by which the required thing may be recognized [51]. In our research problem, 'data' are SITD and SITDS surveys raw data. This is required because even if we see the required thing we couldn't recognize it. And, finally, (c) in any problem there must be a condition which specifies how the unknown is linked to the given data [51].

The procedure that has been describe above is a 'methodology' to solve (mathematical) problems. The key concerns is that it is, in itself, a pattern. Something that may be imitated even tackling different problems of the same class.

It should be noted that the patterns is not (necessarily) a function. There is no 'parameter' passing. Further, it is not a (simple) procedure that, as a recipe, may be indefinitely repeated giving the same results. It is not a blind computational routine.

It is closer to scripts [52] in the sense that describes the purposes of every role, but the actual staging results in different instances of the same play. In project management, but in a special sense, in software project management projects, it is clear that even when a project manager follows a software development methodology, the manager can never attend the same script in the same way [52] to carry a project out. Nothing guarantees success. However, the pattern is relevant to improve the chances of success.

The structural part of processes, even under changing conditions, preserves the identity of the process itself. Each time a software process is conducted it becomes a new process [53], which recalls for the dynamics of adaptive systems. A short-term gentle adjustment of the structure to an evolutionary environment makes to evolve (co-evolve) the whole that is being studied, either the whole is a software development process [52, 54] or a technology transfer process [28, 29, 31] there exist a structure revealed by patterns [55].

The following sections introduce the procedures of the methodology and the rules to proceed between them.

C. Obtaining data from surveys

Select the periods of time corresponding to the elapsed time of interest. Given that SITD and SITS are available since 2005 and 2006, respectively, including data corresponding to the preceeding two years, on a two-year basis, the researcher should adjust the research window. The public character of data of the NDS guarantee availability in plain format or spreadsheet format.

D. ISIC standardization

When the study includes data from different years for SITD or SITDS, the researcher should make sure that data is classified according to compatible revisions of the International Standard Industrial Classification of All Economic Activities (ISIC). Not all categories of ISIC persists between periods. The researcher should be aware of changes in definitions of each code between revisions (for instance between ISIC Rev 3 and ISIC Rev 4. New codes implies to define dummy codes when constructing a sociomatrix for a previous year.

E. Define dimensions and datasets

Identify the dataset of recipient and the dataset of source of ideas. The source may be further classified according to the boundary of the recipient organization. Blurry boundaries should be documented to state clearly the meaning of (i) internal source of ideas, (ii) external source of ideas. This last category may be even further classified in: other firms, specialized groups, and external relationships).

Define as 'dimension 1' all actors (and its data), corresponding to internal sources. Keep ISIC classification in the following categories: (a) Internal R&D departments, (b) production department, (c) Sales and marketing department, (d) other department of the firm, (e) interdisciplinary groups, (f) management staff, and (g) workers.

As 'dimension 2' include other firms. Keeping ISIC classification of data, the categories are: (a) external R&D department, (b) other related firm, (c) headquarters, (d) clients, (e) competitors, and (f) suppliers.

'Dimension 3', specialized groups, is composed by (a) professional associations or sectorial associations, (b) Chambers of Commerce, (c) agricultural and forest research centers, and (d) technological development center.

Set 'Dimension 4', external relations, as (a) SENA training centers, (b) Consultants and experts, (c) trade fairs and exhibitions, (d) seminars, (e) Books, journals and catalogs, (f) Intellectual property information systems, (g) copyright information systems, (h) Internet and other ICT, and (i) scientific databases.

The first dataset is the matrix resulting from the cross-match between dimension 1 and ISIC. These represent actors of a 2-mode network.

The second dataset is the matrix resulting from the cross-match between dimensions 2, 3 and 4 and ISIC. These represent actors of a 2-mode network.

VIII. Social network data processing

Input datasets into a computer software specialized in network analysis. Software packages often provides a data import utility to accomplish this cumbersome time-consuming activity. Software may be one of the following: UCINET+NetDraw, Egonet, Gephi, Pajek, iGraph, JUNG, Statnet.

Apply the projection operator selecting first dimension 1, and then dimensions 2, 3, and 4. This is to perform an oriented analysis to determine relevant sources (rows). If possible, select MCO minimum square method, as a valued (non-binary) network. The co-occurrence method is adequate for binary networks.

Then, calculate eigenvalues by Single Value Decomposition - SVD). This allows to identify background dimensions of the space (set) sector-by-source of the idea.

Finally, group by degree centrality. For each node of te network calculate degree centrality and graph.

IX. Case of the manufacturing sector

In order to study technological knowledge transfer to industrial firms in the manufacturing sector of the Colombian economy, and determine the structure (main links between sources of ideas of innovations and firms) the proposed methodology was applied.

It is assumed that TT is verified when an incoming idea produces a technological innovation in the recipient firm. Data from SITD-IV serve as data to explore the relations between sources/destination of ideas for innovations. External sources are grouped in twenty four (24) categories (Table II).

Besides, recipient organizations –the organizations producing innovations from incoming ideas- are grouped in 64 subsectors by ISIC Rev.3 (Rev.3 A.C.).

Raw data was classified (rows) according to firm size (small, medium and large enterprise, and then populated a matrix. Columns were sorted buy ISIC Rev.3 code sector. A further classification allowed to set national and foreign firms apart.

The result, each of eight matrixes describes a two-mode network, then they were transformed from a 2-mode network by means of a projection method. This was done by selecrting on of the two datasets and linking the nodes according to the rule: it is connected with another node of the other set. It was used [56], projecting a two-mode network on a weighed one-mode network to preserve the structure of the original network. The networks were drawn [57].

X. Results

Data from social networks were analyzed and organized in relatively homogeneous networks according to national or foreign origin of the firm that originates the idea.

The analysis proceeds by the size of the firm: large, medium or small, keeping the origin (national or foreign origin) of the firm. Fig. 3 shows the 2-mode network for large firms taking ideas from foreign sources.

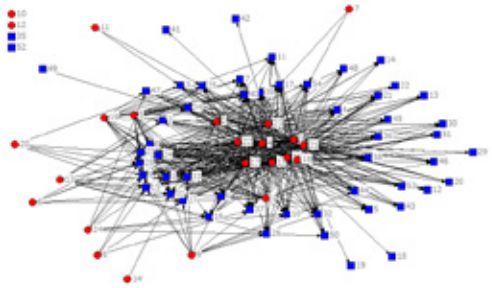


Fig. 3. Two-mode network for large firms with ideas from foreign sources.

Red nodes correspond to the dataset of source of ideas, while blue nodes correspond to the second dataset: manufacturing sector firm category (ISIC). Resulting data shows that sources of higher centrality are as follows (according to their indegree): 1 (151), 2 (152), 3 (153), 4 (154), 15 (175), 16 (181), 17 (191), 18 (192), 21 (203), 22 (210), and 23 (221), as foreign sources for large firms. It is remarkable that foreign sources 10 (160), and 12 (172) are not relevant sources for large firms. Furthermore, firms in subsectors 35 (281) y 52 (359) do not use foreign sources of ideas to incept innovation.

The same conventions apply for the two-mode network of large firms and national sources of ideas (Fig. 4).

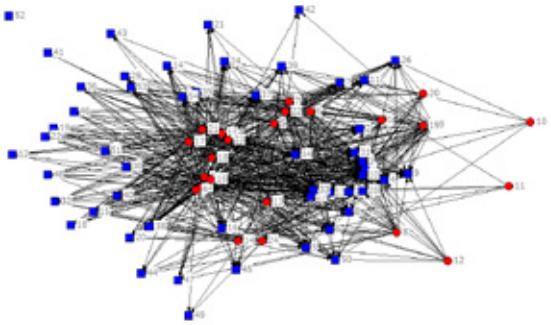


Fig. 4. Two-mode network for large firms with ideas from national sources.

Similarly, to the previous analysis, sources of higher centrality are: 2 (152), 4 (154), 6 (156), 7 (157), 13 (173), 15 (175), 16 (181), 17 (191), 18 (192), 21 (203), 22 (210), and 23 (221).

The case for medium enterprises receiving ideas from foreign sources (Fig. 5) reveal the higher centrality as follows: 2 (152), 3 (153), 4 (154), 16 (181), 18 (192), 21 (203), 22 (210) and 23 (221).

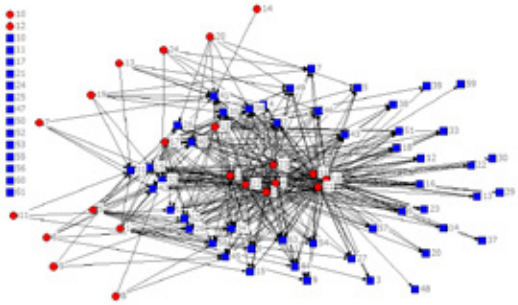


Fig. 5. Two-mode network for medium firms with ideas from foreign sources.

Medium enterprises receiving ideas from national sources (Fig. 6) shows higher centrality for the following national sources: 4 (154), 5 (155), 6 (156), 7 (157), 13 (173), 14 (174), 15 (175), 16 (181), 21 (203), 23 (221), and 24 (222).

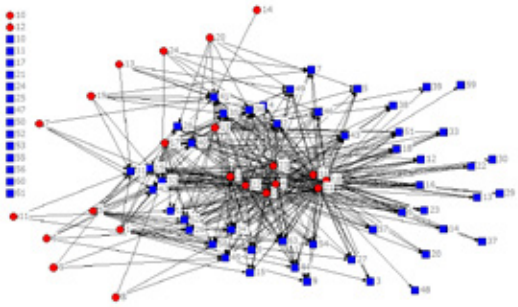


Fig. 6. Two-mode network for medium firms with ideas from national sources.

Small firms reveals a structure of the network where main foreign sources of ideas to produce innovations (Fig. 7) are: 3 (153), 4 (154), 16 (181), 18 (191), 21 (201), and 23 (203).

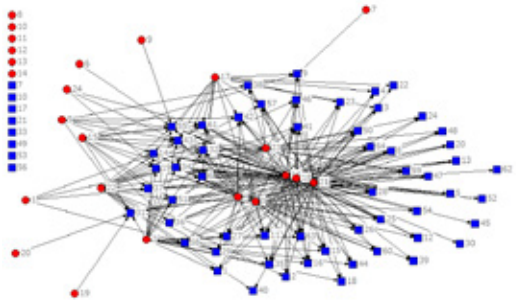


Fig. 7. Two-mode network for small firms with ideas from foreign sources.

When examining small firms linked to their national sources of ideas (Fig. 8) the results are: 2 (152), 7 (157), 16 (181), 17 (182), 18 (191), 21 (201), 23 (203), and 24 (204).

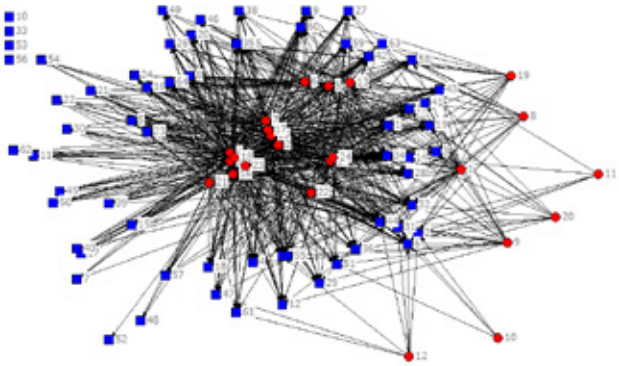


Fig. 8. Two-mode network for small firms with ideas from national sources.

When applying the projection operator over the two-mode network of large firms – foreign sources of ideas (Fig. 3) results that main sources of ideas to generate innovation (Fig. 9) are: 2 (Scientific and technological databases), 16 (Internet), 17 (Technical regulations and standards), and 19 (technological parks) (Table II).

Fig. 9. One-mode network (obtained by projection) for large firms with ideas from foreign sources. The dataset of projection corresponds to the twenty four sources of ideas listed in Table II.

Projection of the two-mode network (Fig. 4) of large firms recipients of ideas from national sources results in a one-mode network (Fig. 10) lets identify the higher centrality nodes, that reveal that main national sources are: 12 (firms from a different sector), 13 (Trade fairs and exhibitions), 14 (Technology-based companies and incubators), 21 (Seminars and conferences) y 23 (industrial property information systems).

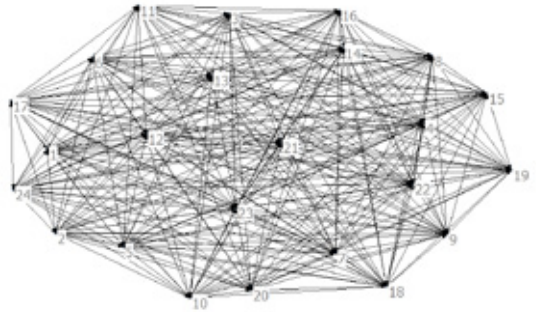


Fig. 10. One-mode network (obtained by projection) for large firms with ideas from national sources. The dataset of projection corresponds to the twenty four sources of ideas listed in Table II.

Projection of the two-mode network (Fig. 5) of medium firms recipient of ideas from foreign sources, in a one-mode network (Fig. 11) reveals that foreign sources: 10 (Consultants or experts) and 12 (Firms form other sectors) are not relevant. In contrast, the rest of foreign sources are probabilistically equal sources of ideas for innovation.

Fig. 11. One-mode network (obtained by projection) for medium firms with ideas from foreign sources. The dataset of projection corresponds to the twenty four sources of ideas listed in Table II.

Projection of the two-mode network (Fig. 6) of medium firms recipient of ideas from national sources, in a one-mode network (Fig. 12) shows that main sources of ideas for innovation are: 15 (Public institutions), 16 (Internet), 17 (Books, journals and catalogs), 18 (Technical regulations and standards), 19 (Technological parks), and y 21 (seminars and conferences).

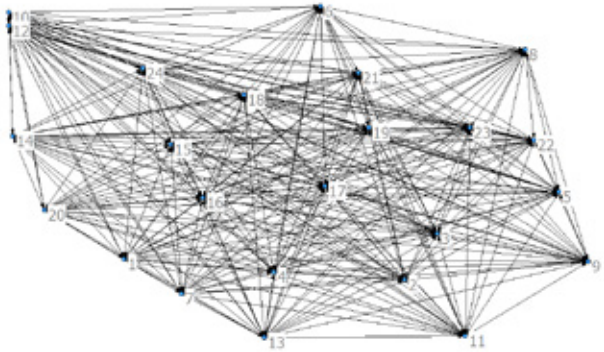


Fig. 12. One-mode network (obtained by projection) for medium firms with ideas from national sources. The dataset of projection corresponds to the twenty four sources of ideas listed in Table II.

The case for small firms, when projecting their two-mode network (Fig. 7), results in a one-mode network (Fig. 13) showing that main foreign sources of ideas for innovation are: 2 (scientific and technological databases), 4 (Technological development centers), 16 (Internet), and 21 (seminars and conferences).

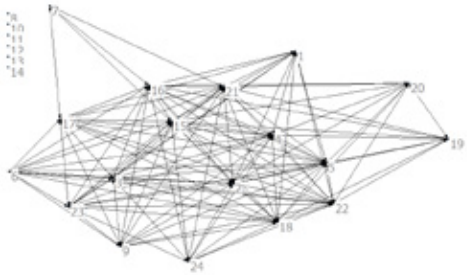


Fig. 13. One-mode network (obtained by projection) for small firms with ideas from foreign sources. The dataset of projection corresponds to the twenty four sources of ideas listed in Table II.

Projection of the two-mode network (Fig. 8) of small firms recipient of ideas from national sources, in a one-mode network (Fig. 14) shows that main sources of ideas for innovation are: 2 (scientific and technological databases), 3 (Chambers of Commerce), 4 (Technological development centers), 15 Public institutions, 16 (Internet), 17 (Books, journals and catalogs), and 21 (Seminars y conferences).

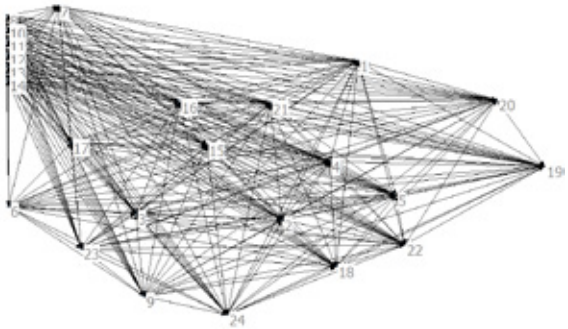


Fig. 14. One-mode network (obtained by projection) for small firms with ideas from national sources. The dataset of projection corresponds to the twenty four sources of ideas listed in Table II.

XI. Discussion

From the analysis, the firms from the manufacturing sub-sector, obtaining ideas to innovate from foreign sources are (ISIC Rev. 3. A. C.): 151 Production, processing and preserving of meat and meat products, 152 Processing and preserving of fruits, legumes, green vegetables, oil and fat, 153 Manufacture of fresh liquid milk, pasteurized, sterilized, homogenized and/or other dairy products, 154 Manufacture of starches and starch products, 175 Manufacturing of fabrics and knitting fabrics, 181 Manufacturing of clothing, except leather based, 191 Leather products, 192 Leather-shoe manufacturing, 203 Manufacture of wooden goods intended to be used primarily in the construction industry, 210 Manufacture of paper and board and paper and board products, and 221 Publishing, printing and reproduction of recorded media.

In addition, main foreign sources of ideas for large firms from those subsectors are scientific and technological databases, technological parks, clients and consultants or experts.

Those firms of subsector 221 Publishing, printing and reproduction of recorded media, prefer to get ideas from international consultants an experts. They complement their knowledge by means of ideas coming from foreign technological parks and information from scientific and technological databases, the specifications received from their international clients.

Ideas coming from technological parks and international consultants for enterprises of 151 Production, processing and preserving of meat and meat products and 153 Manufacture of fresh liquid milk, pasteurized, sterilized, homogenized and/or other dairy products, may allow them the ability to develop preventive actions such as improving competitiveness or even develop competitive advantages- to face foreign competitors in the context of free trade agreements.

However, it seems that the identified sources of ideas for innovation in large firms of the manufacturing sector served to adopt short-term, minimum national competence positions.

A better comprehension of the relationships between the sources of ideas for innovation, according to its origin (foreign or national), and manufacturing firms may provide background to develop policies to encourage firms to adopt an active leading position. Nevertheless, with improved knowledge on the structure of innovation in the manufacturing sector, to strengthen the use of foreign sources of ideas for innovation such as scientific and technological databases, and include the gain of contracting consultants from local universities that may understand local markets better than foreign consultants.

XII. Conclusion

A simple but powerful methodology to investigate the structure of innovation, from SITD and SITDS, based in knowledge transfer in interfirm contexts was developed. This methodology is well founded in social network analysis and easy to use information technology tools.

It was shown that this methodology facilitates a structured generation of information related to the links between sources and receivers of ideas for innovation, allowing to identify preferred sources by firm size, according to the origin of the source (foreign or national).

The methodology that has been introduced and demonstrated in use, satisfies the aim to serve as methodological reference to identify sources of ideas in knowledge transfer to be effectively used in innovation.

Codification of dimensions when constructing datasets may generate biases. In spite of this fact, the methodology remains effective because the relevant findings is on the network structure (identified patterns on the main sources of ideas to generate innovation by any set of firms in ISIC).

These methodology is of interest for foreign companies that may find in Colombia the destination of their products, general activities, and consultancy contracts opportunities.

Finally, in order to compare results from different research processes, a unified but simple methodology facilitates researchers to concentrate on the field of investigation and not in the development of an ad-hoc methodology or the time-consuming learning of a complicated, even complete, methodology.

Further work

This work reports an analysis on two-mode networks emerging from the relations between national and foreign sources of ideas for innovation and manufacturing firms, based on indegree calculation. Additional work may include determining if other centrality and prestige measures may provide accurate information on sources of ideas.

To enhance this methodology, in terms of informing data to stakeholders, a short-term economic oriented analysis may be included. This short-term analysis should include information on international free trade agreements and a transition analysis of foreign commerce policies.

Acknowledgment

Authors express their gratitude to Prof. Rafael Hurtado, at Universidad Nacional de Colombia, for providing useful counselling during the development of this research project. This paper is based on the presentation at CONIITI 2015.

References

[1] J. H. Dyer and N. W. Hatch, "Network-specific capabilities, network barriers to knoledge transfers, and competitive advantage," in *Academy of Management Proceedings*, 2004, pp. V1-V6.

[2] R. M. Grant, "The knowledge-based view of the firm: implications for management practice," *Long range planning,* vol. 30, pp. 450-454, 1997.

[3] M. Boisot and A. Canals, "Data, information and knowledge: have we got it right?," *Journal of Evolutionary Economics,* vol. 14, pp. 43-67, 2004.

[4] C. E. Shannon and W. Weaver, *The mathematical theory of information*. Urbana, Illinois.: University of Illinois Press, 1949.

[5] P. R. Carlile, "Transferring, translating and transforming: an integrative relational approach to sharing and assessing knowledge across boundaries," *Organization Science,* vol. 15, pp. 555-68, 2004.

[6] J. E. Arias Pérez and C. A. Aristizábal Botero, "El dato, la información, el conocimiento y su productividad en empresas del sector público de Medellín," *Semestre Económico,* vol. 14, pp. 95-109, 2011.

[7] G. Durant-Law, "TARDIS: A Journey Through an Enterprise Knowledge Space," Master of Knowledge Management, School of Information and Management, University of Canberrra, Canberra, Australia [http://www.durantlaw.info/sites/durantlaw.info/files/TARDIS_Final_Document.pdf], 2004.

[8] R. McDermott, "Why information technology inspired but cannot deliver knowledge management," *Knowledge and communities,* vol. 41, pp. 21-35, 2000.

[9] G. Dosi, C. Freeman, and S. Fabiani, "The process of economic development: introducing some stylized facts and theories on technologies, firms and institutions," *Industrial and Corporate Change,* vol. 3, pp. 1-45, 1994.

[10] E. E. Hagen, "The process of economic development," *Economic Development and Cultural Change,* pp. 193-215, 1957.

[11] [11] I. Adelman, *Theories of economic growth and development*. Stanford, Calif.: Stanford University Press, 1961.

[12] R. J. Barro, X. Sala-i-Martin, and M. I. T. Press, *Economic growth*. Cambridge; London: The MIT Press, 1999.

[13] N. Kondratieff and W. Stolper, "The Long Waves in Economic Life," *The Review of Economics and Statistics,* vol. 17, pp. 105-115, Nov. 1935 1935.
[14] M. Avella and L. Fergusson, "El ciclo económico, enfoques e ilustraciones. Los ciclos económicos de Estados Unidos y Colombia," *Borradores de Economía,* vol. 284, pp. 1-78, 2003.
[15] R. H. Day, "Irregular growth cycles," *The American Economic Review,* pp. 406-414, 1982.
[16] N. Khabiri, S. Rast, and A. A. Senin, "Identifying main influential elements in technology transfer process: a conceptual model," *Procedia-Social and Behavioral Sciences,* vol. 40, pp. 417-423, 2012.
[17] V. Gilsing, R. Bekkers, I. M. B. Freitas, and M. van der Steen, "Differences in technology transfer between science-based and development-based industries: Transfer mechanisms and barriers," *Technovation,* vol. 31, pp. 638-647, 2011.
[18] UNCTAD, *Transfer of technology*. New York, 2001.
[19] E. M. Rogers, *Diffusion of innovations, 5th ed.* New York. UJ303.484 R64 2003 Colección alterna No.2: Free Press. N.Y., 2003.
[20] A. Hevner and S. Chatterjee, *Design research in information systems: theory and practice* vol. 22: Springer Science & Business Media, 2010.
[21] E. T. Layton Jr, "Technology as knowledge," *Technology and culture,* pp. 31-41, 1974.
[22] K. J. Arrow, "Classificatory notes on the production and transmission of technological knowledge," *The American Economic Review,* pp. 29-35, 1969.
[23] E. M. Rogers, "The nature of technology transfer," *Science Communication,* vol. 23, pp. 323-341, 2002.
[24] E. G. Carayannis, "Knowledge transfer through technological hyperlearning in five industries," *Technovation,* vol. 19, pp. 141-161, 1999.
[25] G. Von Krogh, K. Ichijo, and I. Nonaka, *Enabling knowledge creation: How to unlock the mystery of tacit knowledge and release the power of innovation*: Oxford university press, 2000.
[26] G. Von Krogh, "Understanding the problem of knowledge sharing," *International journal of information technology and management,* vol. 2, pp. 173-183, 2003.
[27] A. Wahab, A. Sazali, H. Uli, and R. Rose, "Evolution and development of technology transfer models and the influence of knowledge-based view and organizational learning on technology transfer," *Research Journal of International Studies,* vol. 12, pp. 79-91, 2009.
[28] D. V. Gibson and R. W. Smilor, "Key variables in technology transfer: A field-study based empirical analysis," *Journal of Engineering and Technology management,* vol. 8, pp. 287-312, 1991.
[29] T. K. Sung and D. V. Gibson, "Knowledge and technology transfer: levels and key factors," in *Proceeding of the 4th International Conference on Technology Policy and Innovation*, 2000, pp. 3-7.
[30] E. S. Rebentisch and M. Ferretti, "A knowledge asset-based view of technology transfer in international joint ventures," *Journal of Engineering and Technology Management,* vol. 12, pp. 1-25, 1995.
[31] S. Abdul Wahab, R. C. Rose, U. Jegak, and H. Abdullah, "A review on the technology transfer models, knowledge-based and organizational learning models on technology transfer," *European journal of social sciences,* vol. 10, 2009.
[32] S. M. Wagner and C. Buko, "An empirical investigation of knowledge-sharing in networks," *Journal of Supply Chain Management,* vol. 41, pp. 17-31, 2005.
[33] G. P. Huber, "Organizational learning: The contributing processes and the literatures," *Organization science,* vol. 2, pp. 88-115, 1991.
[34] I. Nonaka and H. Takeuchi, *The knowledge-creating company : how Japanese companies create the dynamics of innovation*. New York [etc.]: Oxford University Press, 1995.
[35] I. Nonaka, R. Toyama, and A. Nagata, "A firm as a knowledge-creating entity: a new perspective on the theory of the firm," *Industrial and corporate change,* vol. 9, pp. 1-20, 2000.
[36] R. M. Grant, "Toward a knowledge-based theory of the firm," *Strategic management journal,* vol. 17, pp. 109-122, 1996.
[37] C. Dhanaraj, M. A. Lyles, H. K. Steensma, and L. Tihanyi, "Managing tacit and explicit knowledge transfer in IJVs: the role of relational embeddedness and the impact on performance," *Journal of International Business Studies,* vol. 35, pp. 428-442, 2004.
[38] J. L. Moreno, "Sociogram and sociomatrix," *Sociometry,* vol. 9, pp. 348-349, 1946.
[39] J. L. Moreno, "Sociometry in relation to other social sciences," *Sociometry,* vol. 1, pp. 206-219, 1937.
[40] S. Wasserman, *Social network analysis: Methods and applications* vol. 8: Cambridge university press, 1994.
[41] M. E. J. Newman, *Networks : an introduction*. Oxford; New York: Oxford University Press, 2010.
[42] J. Park and M. E. Newman, "The statistical mechanics of networks," *Phys. Rev. E* vol. 70 p. 066117 doi:10.1103/PhysRevE.70.066117, 2004.
[43] M. E. Newman, "The structure and function of complex networks," *SIAM review,* vol. 45, pp. 167-256 2003 (available at arXiv.org/cond-mat/0303516).
[44] H. A. Simon, "On a class of skew distribution functions," *Biometrika,* pp. 425-440, 1955.
[45] R. G. Hurtado and J. E. Mejía, "The structure of investment for technological innovation and development activities in the colombian manufacturing industry," *Innovar,* vol. 24, pp. 33-40, 2014.
[46] J. C. Duque, S. J. Rey, and D. A. Gómez, "Identifying industry clusters in Colombia based on graph theory," *Ensayos Sobre Política Económica,* vol. 27, pp. 14-45, 2009.
[47] DMReview. (2007). *Glossary. SourceMedia (available at www.dmreview.com/glossary/a.html)*.
[48] N. Prakash, "A process view of methodologies," in *Advanced Information Systems Engineering*, 1994, pp. 339-352.
[49] R. Veryard, "What are methodologies good for?," *Data Processing,* vol. 27, pp. 9-12, 1985.
[50] . Bussmann, R. Nicholas, and M. Wooldridge, "On the identification of agents in the design of production control systems," in *Agent-Oriented Software Engineering*, 2001, pp. 141-162.
[51] G. Pólya, *Mathematical discovery; on understanding, learning, and teaching problem solving (1981, Combined edition)*. New York: John Wiley & Sons, 1962,.
[52] V. H. Medina García, S. BolañosCastro, and R. González Crespo, "Process Management Software as a Script, and the Script as a Pattern," *Interbnational Journal of computer and Communication Engineering* vol. 1, pp. 147-150, 2012.
[53] S. J. Bolaños Castro, R. González Crespo, and V. H. Medina García, "Antipatterns: a compendium of bad practices in software development processes," *IJIMAI,* vol. 1, pp. 41-46, 2011.
[54] S. J. Bolaños Castro, R. González Crespo, and V. H. Medina García, "Patterns of software development process," *IJIMAI,* vol. 1, pp. 33-40, 2011.
[55] Pacheco, A., H. Bolivar-Baron, R. Gonzalez-Crespo, and J. Pascual-Espada, "Reconstruction of High Resolution 3D Objects from Incomplete Images and 3D Information", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 2, issue Regular Issue, no. 6, pp. 7-16, 06/2014
[56] S. P. Borgatti, M. G. Everett, and L. C. Freeman, "Ucinet for Windows: Software for social network analysis," 2002.
[57] S. P. Borgatti, "NetDraw software for network visualization," *Lexington, KY: Analytic Technologies,* p. 95, 2002.

Orlando López-Cruz is member of the research group "Riesgo en sistemas naturales y antrópicos", Ph.D.(c) in Engineering, Pontifical Xaverian University, Master in Administration, Universidad Nacional de Colombia, Organizational Control Systems Specialist, Universidad de los Andes, Colombia, Economist, Central University, Colombia. He has held management positions at Organización Sanitas Internacional (Colsanitas), Colsubsidio, and consultant at Gesfor Group. He held teaching positions at Universidad del Rosario, Bogotá, and Central University, Bogotá. Currently, he is associate professor at Universidad El Bosque, Bogotá. Prof. López-Cruz research and consulting interests include Processes and structures design for innovations in organizations, organizational knowledge transfer, engineering complex controls for organizations, design and implementation of agent-based systems for social systems, and adequate implementation of organizational information systems.

Nelson Obregón Neira is member of the research group "Riesgo en sistemas naturales y antrópicos", Ph.D. in Hydrology, University of California, Davis, U.S.A. Master in Civil Engineering, Universidad de Los Andes, Bogotá. Civil Engineer, Universidad Francisco de Paula Santander, Cúcuta, Colombia.He has held teaching and research positions at Pontifical Xaverian University and Universidad Nacional de Colombia. He currently is Director of Instituto Geofisico. Prof. Obregón Neira research and consulting interests include: Intelligent Technologies, turbulent processes, Chaos theory, applied fractal geometry, Data mining and Learning Machine in Geosciences. Computational modeling and applied mathematics (deterministic, statistical and probabilistic mathematics).

Artificial Intelligence applied to project success: a literature review

Daniel Magaña Martínez, Juan Carlos Fernández-Rodríguez

Universidad Antonio de Nebrija, Madrid, Spain

***Abstract* — Project control and monitoring tools are based on expert judgement and parametric tools. Projects are the means by which companies implement their strategies. However project success rates are still very low. This is a worrying situation that has a great economic impact so alternative tools for project success prediction must be proposed in order to estimate project success or identify critical factors of success. Some of these tools are based on Artificial Intelligence. In this paper we will carry out a literature review of those papers that use Artificial Intelligence as a tool for project success estimation or critical success factor identification.**

***Keywords* — Project Management, Artificial Intelligence, Decision Support Systems, Project Success, Critical Success Factors.**

I. Introduction

Project management is the main tool for implementing a company's goals so understanding its key issues is really absolutely vital for Project success [1].

The main challenge of project Management, whether in the times of the Romans, the Renaissance or the present, is to accomplish Project objectives as specified in the main bodies of knowledge PMBOK[2], PRINCE2[3] e ISO21500[4].

Nevertheless, and in spite of the considerable length of time that Project Management has been in existence, and the improvements of the last few years, there are a lot of projects that are still being classified as not successful, in terms of the way they have been managed or because of their results [5], [6].

Some determining factors have been identified in Project Management literature:

1. Projects have always been associated with complexity [7], but they are getting more complex in general, independently of the industry [8], [9].
2. Stakeholders play an important role in project development, it is not just the realm of the project manager and their team. [10]–[14].
3. Projects have always been surrounded by uncertainty and continuous changes that make it really difficult to plan, and accomplish, schedules, resources and budgets [14]–[18].

As identified in the different bodies of knowledge, PMBok[2], Prince2[3] e ISO21500[4] most of the processes are based on expert judgement or on other parametric analytic tools. The fact that expert judgement is one of the most important tools in project management has some limitations:

1. Projects are normally developed in a restricted resource environment so the more complex is a project is, the more accuracy that is needed, and the more difficult it is to apply this expert judgement.
2. Expert judgement is applied by people, by experts so it can lead to bias [19].

On the other hand the bodies of knowledge identify other processes related to learned lessons of the project as part of the methodology for obtaining better results, learning from past experience [2]–[4] and managing all that knowledge in an effective way [20]. However reality shows that, in general, lessons learned are treated in a very superficial way, they are not well documented and they are not communicated so others can take advantage of them in future projects [21]–[23]. This is probably because lessons learned are perceived in a negative or punitive way in many companies [22].

Summarizing, it seems that traditional project tools are not working properly when trying to predict project success. Accordingly new project management tools are emerging in order to improve project success or project success estimation.

The goal of this paper is to review these new proposals that use Artificial Intelligence to improve project success or that can simply predict it. If we were able to predict the future, we would be able to prepare for it. Originally future prediction in project management has been done from the point of view of expert judgement, based on the opinion of those who are analysing the project, the experts. There are studies that try to make a model of this expert knowledge so technology could mitigate the identified risks [24], [25] and consequentially be applied to project management.

II. Literature review methodology

Different references have been found in scientific literature associated with the use of alternative tools to those used in the main bodies of knowledge, some of which are based on Artificial Intelligence. These alternative tools are applied to various project management areas:

1. Estimation of project success
2. Identification of critical success factors
3. Relatedness to project budget
4. Connection to project schedule
5. Project planning
6. Relatedness to risk identification

In this paper we will focus only on those using Artificial Intelligence algorithms to try to conduct a project success estimation o try to identify critical success factors of a project.

In order to carry out this literature review we have done an unstructured initial search just to identify different approaches and goals in the field of artificial intelligence applied to project management. The main objective is to identify the algorithms that are being used for this purpose. The keywords used for this search have been: "p*roject management*", *"artificial intelligence*", "*project success*", "*project success prediction",* and *"critical success factors*".

DOI: 10.9781/ijimai.2015.3510

Subsequently new structured searches have been made for each of the algorithms in order to find more references to complete in this way the bibliography to be analysed. Repeated searches have been made combining these keywords: "*critical success factors*", "*Project success*" with every algorithm identified in unstructured previous searches.

Both searches, unstructured and structured, have been limited to scientific papers, books or book chapters, excluding non-scientific articles. The researche has been done using *Bucea* searching tool at Universidad Complutense de Madrid, Spain.

The result of these searches is 16 references where Artificial Intelligence has been applied to project success estimation or critical success factor identification. Identified references start in 1997 and finish in 2014.

Based on these results, we will perform a structured analysis in order to identify how artificial intelligence algorithms are applied, detect the authors' objectives and list their limitations. Finally we will summarize the authors' conclusions.

Therefore the following questions arise after this literature review process:

1. What are the authors' goals when applying artificial intelligence to project success?
2. What artificial intelligence algorithms have been applied?
3. What limitations are identified when applying those algorithms?

In the next section we will perform a project success concept definition based on existing principle perspectives on scientific literature.

III. Project success

One of the main worries of project management directors is knowing, with some anticipation, if the project they are managing is going to be successful or not. This worry is not a guessing game. It is supported by control and monitoring tools defined in project management frameworks and the bodies of knowledge [2]–[4].

It is not the goal of this paper to review the vast existing literature associated with project success. We would like to conceptualize the project success concept based on the main existing research lines so it can help the reader to understand how artificial intelligence can help in this area.

A project has been, traditionally, categorized as successful if it accomplished the Triple Constraint: scope, budget and schedule [26]. Traditional statistics or parametrical tools were enough for this purpose. However these tools leave aside other qualitative aspects of project management, for example the stakeholders' point of view [27], [28].

Initial studies in this field conclude that we should distinguish project success concepts, focusing on managerial processes of project management on the one hand, and the traditional Triple Constraint of scope, schedule, budget and quality on the other hand where product success criteria is more important from the product point of view. In this regard a project could have been perfectly managed from the project team's point of view, but the product is not in accordance with the stakeholders' expectations [7], [11], [29].

Some authors have concluded that a correct definition of project success or project failure has not been made and this is one of the main reasons why projects are still being considered as failures [27], [30], [31] .

There is a variety of studies proposing new key indicators to create a new framework in order to measure project success based on five dimensions: project performance, project impact on the customer, project impact on the business and its preparation for the future [11], [27], [30], [32].

There is another field of research in literature of those who try to identify the project's critical success factors [26], [33].

IV. Artificial intelligence algorithm's applied to project success

During this literature review, the following algorithms based on artificial intelligence, and applied to project success have been found.

A. Neural Networks

Neural Networks attempt to simulate, to a degree, human way of thinking, and are used, nowadays, for multiple purposes, from credit approval, fraud detection, surveillance systems and other kinds of prediction purposes.

One of the main parts of neural networks consists of learner training so neural network adjust to data patterns and give better results. This training is done comparing neural network results with real and known data and is repeated so it adjusts until results of a very low error rate are achieved.

Neural networks, because of their characteristics, are more accurate than linear models [34], based on regression models, which have been frequently used in project management [35].

B. Fuzzy Cognitive Maps

Fuzzy Cognitive Maps are fuzzy graphical structures that allow the representation of causal reasoning. This graphical representation is made of nodes where the most relevant nodes are specifically identified for a decision-making system. Fuzzy cognitive maps have their origin in a fuse of fuzzy logic and neural networks [36].

C. Genetic Algorithms

Genetic Algorithms try to simulate the evolutionary natural process and were originally proposed by Holland [37].

They are easy to apply so they can be fused with other heuristic methods creating ad-hoc solutions. However it is difficult to apply them to large, complex, difficult-to-solve problems [38].

D. Bayesian Model

Bayesian networks are described as a representation of a joint probability distribution. It is one of the most common methods for data classification in different categories [39].

The Bayesian model allow us to answer questions such as what is the probability of X being in state x_1 if $Y = y_1$ and $Z=z_1$. In other words, links the probability of A given B with the probability of B given A.

E. Evolutionary Fuzzy Neural Inference Model – EFNIM

EFNIM fuses genetic algorithms, fuzzy logic, and neural networks and has been traditionally used for civil engineering problem solving. The combination of these three algorithms makes the strengths of one cover the weaknesses of the other. So genetic algorithms are used for optimization purposes, fuzzy logic deals with uncertainty and neural networks for mapping inputs and outputs.[38]

F. Evolutionary Fuzzy Hybrid Neural Network – EFHNN

The model EFHNN includes four algorithms of artificial intelligence:

1. Neural Network
2. High Order Neural Network
3. Fuzzy Logic
4. Genetic Algorithm

Neural Networks and High Order Neural Networks, named together as Hybrid Neural Network (HNN), manage the inference engine while Fuzzy Logic deals with the fuzzy layer. Genetic algorithms optimize the final model.

The difference with EFNIM is that this model is able to manage problems more deeply thanks to the large number of HNN models. [40].

G. Support Vector Machine

This is a new way of learning, which is more powerful than traditional learning tools. It is able to solve data categorization problems and regression problems as well.

Just as neural networks do, SVM requires training and testing with a training dataset. SVM's characteristics allow it to deal better with unknown data and, generally speaking, they present some advantages over neural networks. They are being applied successfully to cost estimation in the construction industry.

H. Fast Messy Genetic Algorithm

The Fast Messy Genetic Algorithm can identify optimal solutions in an efficient way to problems with a large number of permutations. It is known because of its flexibility and because it can be fused with other methodologies to get better results [41].

The difference between it and other genetic algorithms is based on the possibility of modifying building blocks to better identify partial solutions so as to focus on a global solution faster.

I. K-Means Clustering

K-Means is an easy approach for creating data cluster from random data. It is commonly used for image pattern detection as well as for many other applications. Its main problem is that it cannot ensure an optimal convergence, but is widely used due to its simplicity.

J. Bootstrap aggregating neural networks

Bootstrap aggregating neural networks are a combination of multiple artificial neural network classifiers. They use more than one classifier based on ANNs so the final decision is taken from each classifier by a voting system [42].

K. Adaptive boosting neural networks

The main difference with *Bootstrap aggregating neural networks* is that adaptive boosting neural networks use weights that are readjusted on every iteration giving less importance to those solutions that have not been classified correctly. As a result, classifiers focus on more complex samples obtaining a faster solution each time [42].

V. Literature review

As a field of research, the application of artificial intelligence algorithms to the prediction of project success brought up a wide selection of authors' objectives. For a better comprehension we will divide them into two groups, those that try to predict project success and those that try to identify critical success factors.

A. Determining Critical Success Factors

Within these groups we find those that apply artificial intelligence algorithms to critical success factors identification for measuring project success [36], [39], [43]–[46].

Multiple studies affirm that project success is not only a matter of complying with the already known Triple Constraint, but also depends on the perception of success by the stakeholder, and that what could be satisfactory for one could be unsatisfactory for the other. Because of this, it is essential to define project success criteria [47]. Furthermore, critical success factors may vary during the project life cycle so it is important to identify them throughout the project.

From this starting point, these algorithms have been identified in literature reviews in order to detect critical success factors of projects (CSF's):

1. Neural Networks
2. Fuzzy Cognitive Maps
3. Genetic Algorithms
4. The Bayesian Model

The first paper is focussed on the construction industry, and its objective is to detect CSFs using neural networks. It identifies eight key factors for project management success [43]:

1. The number of organization levels between the project manager and project staff.
2. How detailed the project design is before the construction phase.
3. The number of control meetings during the construction phase.
4. The number of times that the budget has been updated.
5. The setting up of a constructability system.
6. Team rotation.
7. The amount of money spent on project management.
8. The technical expertise of the project manager.

The author uses data collected for his thesis [48] and analyses them with neural networks in order to get a mapping between managerial elements of project management and project management performance. The final model could be applied as a prediction tool for new project management strategies in the early stages of a project. In conclusion, this model could be also used for project budget performance prediction.

Focusing on defence projects, we find a comparative between neural networks and regression analysis tools for identifying managerial project management criteria oriented to project success in high technology defence projects. In this comparative some factors are identified as less important by regression models while they become more important with the use of neural network algorithms. Neural network algorithms are significantly more accurate when working with unknown data. The author performs his study based on data collected from 89 defence projects developed in Israel between the 80s and 90s [44].

Centred on the IT sector, classified by the authors as different from other sectors due to its complexity and high possibilities of failure, a model for mapping the project's success and CSF's perception, and the link between them is proposed. To perform this mapping, it uses *Fuzzy Cognitive Maps* (FCMs). The authors defend that the success concept in IT projects is a complex concept, not structured, and so FCM is more appropriate for dealing with this kind of ambiguity. FCM is better suited to IT projects. The authors validate their model with a real project case. The model has a weakness related to the SRMS matrix that shows some wrong data so it needs reviewing by an expert to analyse and correct results [45].

The same authors perform a benchmarking between three emerging methodologies oriented to CSF identification. The three benchmarked methodologies are *Critical Success Chains* (CSC) [49], *Analytic Hierarchy Process* (AHP) [50]one central issue is the study of critical success factors (CSF and *Fuzzy Cognitive Maps* (FCM) [45], all focused on IT projects. The authors conclude by listing advantages, disadvantages and limitations of each of them. As *Fuzzy Cognitive Maps* is the only one related to Artificial Intelligence, we should

remark that it is the one most similar to human perception, but even so it requires an expert for its interpretation, and this situation introduces subjectivity into the model as commented in their previous paper.

Again related to IT projects, more specifically software development, the authors make a proposal of a model for identifying the CSF that could impact on project outcome to a greater extent. The objective is to provide a tool for the project manager that allows him to control those identified risks that threaten the project outcome. The authors focus on resources assignation to solve these factors. The authors attempt to identify the most important risks, and suggest the most efficient resource investment based on those risks. They define efficiency as the rate between success probability and cost. This efficiency definition is used as an aptitude function for genetic algorithms. To perform this research, the authors use a dataset of previous software development projects developed both *in-house* and outsourced within the Chilean industry. The authors applied genetic algorithms to obtain this optimization of resources, and the prediction of success. The model also suggests a cost effective investment proposal [46] .

Finally, and using an expert system based on a Bayesian model and centred on IT projects, again we find a paper whose objective is to know in advance the impact that decisions have on the project's outcome. The authors have a double aim. Firstly to create an expert model, based on a Bayesian model, which allows the project manager to analyse the impact of a decision on project outcome. They use success criteria definitions found in literature. Secondly, to make a recompilation of IT project-related knowledge. The main conclusions of this study are related to the importance of stakeholder engagement, support of senior management, goals and objectives definition and their association with project success [39].

B. Determining project success

Furthermore we find papers that try to forecast project success for the duration of the project life cycle in its early stage, or at any other time point of the project [38], [40], [42], [51]–[54]the project management, in which the final status of project is estimated, must be incorporated.In this paper, we consider estimation of the final status(that is, successful or unsuccessfulDuring the literature review, these algorithms applied to project success prediction have been found:

1. Bayesian Model
2. Evolutionary Fuzzy Neural Inference Model - EFNIM
3. Neural Networks
4. Support Vector Machine
5. Fast Messy Genetic Algorithm
6. K-Means Clustering
7. Bootstrap aggregating neural networks
8. Adaptive boosting neural networks

Artificial Intelligence application for project success predicting is relatively recent, since the first reference is from 2006. This model estimates project final state applying a Bayesian classifier to different metrics collected from a project. The aim of this research is to make this estimation in the early stages of the project. Metrics selection can be performed by experts or using statistical methods, which are more accurate. The research is focused on IT software development and the authors consider a project as successful if it has been developed on schedule, on budget and to a satisfactory quality. The study is supported by data collected from 28 software development in-house projects. Results show that an accurate success prediction can be made, but having the right metrics is a key issue for getting accurate results [51].

The next paper, focused on the construction industry, suggests a model for a dynamic project success estimation. The model, named EPSPM, fuses several artificial intelligence algorithms: genetic algorithms (GAs) for optimization, fuzzy logic (FL) for reasoning and neural networks (NN) for mapping inputs and outputs. EPSPM is integrated with the *Continuous Assessment of Project Success* tool [55], which allows us to create a real time decision-making system. The authors define project success in the construction industry as that which conforms to the budget, the schedule, the performance and the project safety, in addition to other subjective criteria. The project outcome could be influenced by many different factors along the project cycle, so it is interesting to rely on a tool that allows us to predict project success in any given project. This, somehow, could be done by human beings based on experience, as has traditionally been done. Accordingly, artificial intelligence trying to simulate the human brain could be very helpful. The EPSPM model allows estimating project success at any time selecting critical success factors in every project life cycle phase. It is also supported by a historical project database that allows pattern identification for analysis. Research results show that the suggested model is a valid tool for project managers which allows them to make project success estimations in real time [52]which requires a continuous monitoring and control procedure. To dynamically predict project success, this research proposes an evolutionary project success prediction model (EPSPM .

Centred on the sectors of construction and industry, and with a slightly different approach from the rest of the papers, we find a study based on the importance of pre-planning before the project begins.

Project schedule and budget are identified as main success key elements. A wrong project scope definition may lead to a hike in a project's budget and schedule. The aim of this paper is to create a model that permits us to predict this cost and Schedule increments based on data collected from 62 industrial projects and 78 construction projects. Their intention is to relate the planning prior to the project with its success, with the cost and schedule as principle indicators. The authors use a tool called The *Project Definition Rating Index* (PDRI) to evaluate how well the project scope is defined before the project begins. Research has been performed from data collected with this tool. The authors used two models, the first based on a statistical approach, and the second on neural networks. Even though both models confirm the link between pre-planning and Project success, neural networks are more accurate. In addition, the model based on neural networks can predict costs and time increments based on PDRI's project punctuation. [53] .

Once more focusing on the construction industry, we find a paper suggesting a hybrid model fusing several artificial intelligence algorithms. It uses an inference model named *Evolutionary Support Vector Machine Inference Model* (ESIM) for dynamically predicting project success. The model fuses *Support Vector Machine* (SVM) for learning and *Fast Messy Genetic Algorithm* (fmGA) for optimization. This hybrid model is integrated with CAPP, as previous papers have done, for identifying critical success factors and for doing a real time project success prediction. Research results are that ESIM can predict project success with remarkable accuracy. The model was trained and tested with datasets from 46 CAPP projects. To obtain better results, the authors used *K-means* to select projects with similar characteristics [54]while fmGA deals primarily with optimization. Furthermore, the model integrates the process of Continuous Assessment of Project Performance (CAPP.

With exactly the same aim of predicting Project success dynamically, there is another paper that also fuses several artificial intelligence tools. The algorithms fused are *K-means clustering*, genetic algorithms (GA), fuzzy logic (FL) and neural networks (NN). Once again CAPP is used for dynamically identifying the project's critical success factors. As above, *K-means* is used in order to get similar datasets. FL is used for dealing with uncertainty, NN for datamining and GA for

optimization. The result of this research is a new developed model named *Evolutionary Fuzzy Neural Inference Model* (EFNIM), which is able to accurately estimate project success [38].

In the same line of previous papers, not in vain shares one of the authors, we find a paper that fuses more artificial intelligence tools to create an evolutionary model. This time the selected tools are neural networks fused with high order neural networks fused with fuzzy logic and genetics algorithms creating a model named Evolutionary Fuzzy Hybrid Neural Network (EFHNN), integrated again with CAPP for dynamically identifying critical success factors. The main difference with EFNIM is the combined use of neural networks and h*igh order neural networks*, which allow greater flexibility and let us see how mapped inputs and outputs of the model really are [40].

The most recent paper is based on project planning in the early stages to predict project success in costs and schedule terms. It relies on *PDRI* for determining a rate of project definition before the project starts. To make this prediction, it uses two models based on neural networks and Support Vector Machine. As cost and schedule indicators have important differences, the authors have developed two different models, one for each indicator. In the case of costs, the authors' conclusion is that the best is *SVM* with an accuracy of 92%, followed by *Adaptive Boosting* and finally *Bootstrap Aggregating*. In the case of the schedule, results are slightly worse. With a rate of 80% corresponding to *Adaptive Boosting* followed by *SVMs* and *Bootstrap Aggregating*. As demonstrated in other papers the project's pre-planning is a critical success factor. [53].

VI. Conclusions

The possibility of project success prediction or identifying critical success factors in advance is a field of research where researchers have been working intensively for project management purposes.

Initial approaches have been based on statistical models that had not been able to answer to project Management needs. In artificial intelligence, researchers have found algorithms and tools that deal better with project uncertainty and complex environments where projects are normally developed. Several algorithms deal with specific goals.

Critical success factors identification:

1. Neural Networks
2. Fuzzy Cognitive Maps
3. Genetic Algorithms
4. Bayesian Model

Project success prediction:

1. Bayesian Model
2. Evolutionary Fuzzy Neural Inference Model - EFNIM
3. Neural Networks
4. Support Vector Machine
5. Fast Messy Genetic Algorithm
6. K-Means Clustering
7. Bootstrap aggregating neural networks
8. Adaptive boosting neural networks

The main conclusions obtained from the reviewed papers are that artificial intelligence tools are more accurate than traditional tools, but are still complementary to traditional tools. Artificial Intelligence tools are really helpful for the project manager to control and monitor the project.

However some of the reviewed models have weaknesses and limitations that indicate project managers should still use expert judgement and compare artificial intelligence results with traditional tools before making a decision, so they can adjust them if necessary.

Trending is fusing different artificial intelligence tools so they can take advantage of the strengths of a tool and cover the weaknesses of the rest. Best results are obtained when fusing artificial intelligence tools with specific project tools like CAPP, which permits real-time analysis, and PDRI, which allows the rating of how a project has been defined in its very early stages, before a project begins.

References

[1] J. K. Pinto and J. E. Prescott, "Variations In Critical Success Factors Over The Stages In T," *J. Manage.*, vol. 14, no. 1, 1988.

[2] Project Management Institute, *Project Management Body of Knowledge (PMBOK)*, 5th editio. Project Management Institute, Inc., 2013.

[3] A. Murray, N. Bennett, and C. Bentley, *Managing successful projects with PRINCE2, 2009 edition manual*. London: The Stationery Office, 2009.

[4] "Directrices para la gestión de proyectos," *UNE-ISO 21500*. AENOR, 2012.

[5] The Standish Group, "CHAOS MANIFESTO 2013." The standish Group, 2013.

[6] R. L. Glass, "IT Failure Rates - 70% or 10-15%?," *IEEE Softw.*, vol. 22, no. 3, pp. 112, 110–111, May 2005.

[7] D. Baccarini, "The logical framework method for defining project success," *Proj. Manag. J.*, vol. 30, no. 4, pp. 25–32, 1999.

[8] H. Kerzner, "The Future of Project Management By The Importance of Recognizing Change." pp. 1–127, 2013.

[9] M. Saynisch, "Beyond frontiers of traditional project management: An approach to evolutionary, self-organizational principles and the complexity theory-results of the research program," *Proj. Manag. J.*, vol. 41, no. 2, pp. 21–37, Apr. 2010.

[10] K. Davis, "ScienceDirect Different stakeholder groups and their perceptions of project success," *JPMA*, vol. 32, no. 2, pp. 189–201, 2014.

[11] A. J. Shenhar, O. Levy, and D. Dvir, "Mapping the Dimensions of Project Success.," *Proj. Manag. J.*, vol. 28, no. 2, pp. 5–13, 1997.

[12] G. Indelicato, "Project Management Metrics, KPIs, and Dashboards: A Guide to Measuring and Monitoring Project Performance," *Proj. Manag. J.*, vol. 43, no. 2, pp. 102–102, Apr. 2012.

[13] X. Wang and J. Huang, "The relationships between key stakeholders' project performance and project success: Perceptions of Chinese construction supervising engineers," *Int. J. Proj. Manag.*, vol. 24, pp. 253–260, 2006.

[14] F. N. D. Piraquive, R. G. Crespo, and V. H. M. García, "Analysis and Improvement of the Management of IT Projects," *IEEE Lat. Am. Trans.*, vol. 13, no. 7, pp. 2366–2371, 2015.

[15] L. A. Ika, "Project success as a topic in project management journals," *Proj. Manag. J.*, vol. 40, no. 4, pp. 6–19, Dec. 2009.

[16] J. G. Geraldi, L. Lee-Kelley, and E. Kutsch, "The Titanic sunk, so what? Project manager response to unexpected events," *... J. Proj. Manag.*, vol. 28, no. 6, pp. 547–558, 2010.

[17] B. Sauser, R. Reilly, and A. Shenhar, "Why projects fail? How contingency theory can provide new insights–A comparative analysis of NASA's Mars Climate Orbiter loss," *Int. J. Proj. ...*, vol. 27, no. 7, pp. 665–679, Oct. 2009.

[18] J. G. Geraldi and G. Adlbrecht, "On faith, fact and interaction in projects," *Proj. Manag. J.*, vol. 38, no. 1, pp. 32–44, 2007.

[19] B. Shore, "Systematic biases and culture in project failures," *Proj. Manag. J.*, vol. 39, no. 4, pp. 5–16, Nov. 2008.

[20] R. G. C. Flor Nancy Díaz Piraquive, Víctor Hugo Medina García, "Knowledge Management Model for Project Management," in *Knowledge Management in Organizations*, Springer International Publishing, 2015, pp. 235–247.

[21] K. Jugdev, F. Business, S. Winston, C. Avenue, and S. Albert, "Learning from Lessons Learned : Project Management Research Program," *Am. J. Econ. Bus. Adm. 4 13-22, 2012 ISSN 1945-5488*, vol. 4, no. 1, pp. 13–22, 2012.

[22] J. Julian, "How project management office leaders facilitate cross-project learning and continuous improvement," *Proj. Manag. J.*, vol. 39, no. 3, pp. 43–58, Sep. 2008.

[23] F. F. T. Anbari, E. G. Carayannis, and R. J. Voetsch, "Post-project reviews as a key project management competence," *Technovation*, vol. 28, no. 10, pp. 633–643, 2008.

[24] J. M. Castillo, C. Cortes, J. Gonzalez, and A. Benito, "Prospecting The Future with AI," *Int. J. Artif. Intell. Interact. Multimed.*, vol. 1, no. 2, pp. 1–53.

[25] J. M. Castillo, "A crystal ball made of agents," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 1, no. 3, p. 13, 2010.

[26] J. K. Pinto and D. P. Slevin, "Critical success factors across the project life cycle," *Proj. Manag. J.*, vol. 19, pp. 67–75, 1988.

[27] T. Cooke-Davies, "The 'real' success factors on projects," *Int. J. Proj. Manag.*, vol. 20, no. 3, pp. 185–190, Apr. 2002.

[28] J. K. Pinto and S. J. Mantel, "The causes of project failure," *IEEE Trans. Eng. Manag.*, vol. 37, no. 4, pp. 269–276, 1990.

[29] A. de Wit, "Measurement of project success," *Int. J. Proj. Manag.*, vol. 6, no. 3, pp. 164–170, Aug. 1988.

[30] I. Hyvari, "Success of projects in different organizational conditions," *Proj. Manag. J.*, vol. 37, no. 4, pp. 31–42, 2006.

[31] A. J. Shenhar, D. Dvir, O. Levy, and A. C. Maltz, "Project Success: A Multidimensional Strategic Concept," *Long Range Plann.*, vol. 34, no. 6, pp. 699–725, Dec. 2001.

[32] L. a. Kappelman, L. Zhang, and R. McKeeman, "Early Warning Signs of it Project Failure: The Dominant Dozen," *Inf. Syst. Manag.*, vol. 23, no. 4, pp. 31–36, Sep. 2006.

[33] F. Hartman and R. Ashrafi, "Project management in the information systems and information technologies industries," *Proj. Manag. J.*, 2002.

[34] K. L. Priddy and P. E. Keller, *Artificial Neural Networks*. 1000 20th Street, Bellingham, WA 98227-0010 USA: SPIE, 2005.

[35] J. K. Pinto and D. P. Slevin, "Critical factors in successful project implementation," *IEEE Trans. Eng. Manag.*, vol. EM-34, no. 1, pp. 22–27, 1987.

[36] L. Rodriguez-Repiso, R. Setchi, and J. L. Salmeron, "Modelling IT projects success: Emerging methodologies reviewed," *Technovation*, vol. 27, no. 10, pp. 582–594, Oct. 2007.

[37] J. H. Holland, *Adaptation in Natural and Artificial Systems*, vol. Ann Arbor. 1975.

[38] M.-Y. Cheng, L. Li-Cuan, H.-C. Tsai, and C. Pi-Hung, "Artificial Intelligence Approaches to Dynamic Project Success Assessment Taxonomic," *Life Sci. J.*, 2012.

[39] L. Gingnell, U. Franke, R. Lagerström, E. Ericsson, and J. Lilliesköld, "Quantifying Success Factors for IT Projects—An Expert-Based Bayesian Model," *Inf. Syst. Manag.*, vol. 31, no. 1, pp. 21–36, 2014.

[40] M.-Y. Cheng, H.-C. Tsai, and E. Sudjono, "Evolutionary fuzzy hybrid neural network for dynamic project success assessment in construction industry," *Autom. Constr.*, vol. 21, pp. 46–51, Jan. 2012.

[41] D. Goldberg, B. Korb, and K. Deb, "Messy genetic algorithms: Motivation, analysis, and first results," *Complex Syst.*, vol. 3, pp. 493–530, 1989.

[42] Y.-R. Wang, C.-Y. Yu, and H.-H. Chan, "Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models," *Int. J. Proj. Manag.*, vol. 30, no. 4, pp. 470–478, 2012.

[43] D. K. H. Chua, P. K. Loh, Y. C. Kog, and E. J. Jaselskis, "Neural networks for construction project success," *Expert Syst. Appl.*, vol. 13, no. 4, pp. 317–328, Nov. 1997.

[44] D. Dvir, A. Ben-David, A. Sadeh, and A. J. Shenhar, "Critical managerial factors affecting defense projects success: A comparison between neural network and regression analysis," *Eng. Appl. Artif. Intell.*, vol. 19, no. 5, pp. 535–543, 2006.

[45] L. Rodriguez-Repiso, R. Setchi, and J. L. Salmeron, "Modelling IT projects success with Fuzzy Cognitive Maps," *Expert Syst. Appl.*, vol. 32, no. 2, pp. 543–559, Feb. 2007.

[46] F. Reyes, N. Cerpa, A. Candia-Véjar, and M. Bardeen, "The optimization of success probability for software projects using genetic algorithms," *J. Syst. Softw.*, vol. 84, no. 5, pp. 775–785, May 2011.

[47] G. Thomas and W. Fernández, "Success in IT projects: A matter of definition?," *Int. J. Proj. Manag.*, vol. 26, no. 7, pp. 733–742, Oct. 2008.

[48] E. J. Jaselskis, "Achieving construction project success through predictive discrete choice models," University of Texas, Austin, 1988.

[49] K. Peffers, C. E. Gengler, and T. Tuunanen, "Extending Critical Success Factors Methodology to Facilitate Broadly Participative Information Systems Planning," *J. Manag. Inf. Syst.*, vol. 20, no. 1, pp. 51–85, 2003.

[50] J. L. Salmeron and I. Herrero, "An AHP-based methodology to rank critical success factors of executive information systems," *Comput. Stand. Interfaces*, vol. 28, pp. 1–12, 2005.

[51] S. Abe, O. Mizuno, T. Kikuno, N. Kikuchi, and M. Hirayama, "Estimation of project success using Bayesian classifier," *Proceeding 28th Int. Conf. Softw. Eng. - ICSE '06*, no. 4, p. 600, 2006.

[52] C.-H. Ko and M.-Y. Cheng, "Dynamic Prediction of Project Success Using Artificial Intelligence," *Journal of Construction Engineering and Management*, vol. 133. pp. 316–324, 2007.

[53] Y.-R. Wang and G. Edward Gibson Jr., "A study of preproject planning and project success using ANNs and regression models," *Autom. Constr.*, vol. 19, no. 3, pp. 341–346, 2010.

[54] M.-Y. Cheng, Y.-W. Wu, and C.-F. Wu, "Project success prediction using an evolutionary support vector machine inference model," *Autom. Constr.*, vol. 19, no. 3, pp. 302–307, May 2010.

[55] J. S. Russell, E. J. Jaselskis, and S. P. Lawrence, "Continous assessment of project performance," *J. Constr. Eng. Manag.*, vol. 123, no. 1, pp. 64–71, 1997.

Daniel Magaña Martínez obtained his degree as an Engineer in Computer Science, and MS in Project Management at the University Antonio de Nebrija, Madrid, Spain, in 2000 and 2012, respectively. Currently he is a lecturer in Project Management at the University Antonio de Nebrija, Madrid, Spain. His research interests include Early Warning Systems in Project management. Currently he is working as CIO at University Antonio de Nebrija, Madrid, Spain.

Dr. Juan Carlos Fernández-Rodríguez. PhD in Psychology (Universidad Complutense of Madrid ,Spain), Senior Technician in Risk Prevention and Diploma in Personnel Management. Currently he is Postgraduate Coordinator at University Antonio de Nebrija and Director of Prevention Programs. He has been involved in several research projects related to the Educational Technologies and Knowledge Economy and Globalization. He has published five books, several book chapters and various papers on matters of e-Learning and Educational Technology and Psychosocial maladjustment and its influence in education.

Legal Effects of Link Sharing in Social Networks

Eugenio Gil López[1], Andrés G. Castillo Sanz[2]

[1]*Department of Computer Science, Pontifical University of Salamanca, Salamanca, Spain*
[2]*ESIT, International University of La Rioja, Logroño, Spain*

***Abstract* — Knowledge sharing among individuals has changed deeply with the advent of social networks in the environment of Web 2.0. Every user has the possibility of publishing what he or she deems of interest for their audience, regardless of the origin or authorship of the piece of knowledge. It is generally accepted that as the user is sharing a link to a document or video, for example, without getting paid for it, there is no point in worrying about the rights of the original author. It seems that the concepts of authorship and originality is about to disappear as promised the structuralists fifty years ago. Nevertheless the legal system has not changed, nor have the economic interests concerned. This paper explores the last developments of the legal system concerning these issues**

***Keywords*— Links In Social Networks, Rights Of Knowledge Sharing, Web 2.0.**

I. Introduction

THE year 2014 has come to resolve an important issue in relation to the regime for the protection of copyright in the Internet, since two decisions of the Court of Justice of the European Union (in later CJEU) have been published, which have tried to clarify as much as possible whether there is any violation of the right of public communication in the case of links that lead to works of intellectual property without authorization.

These are both the sentences of February 13 2014 of the Fourth Chamber of the European Court of Justice in the Svensson[1] case and the sentences of October 21 2014 of the Ninth Chamber in the Bestwater [2] case. Both resolutions should be related to the recent reform of Spanish intellectual property law by means of the law 21/2014 of 4 November, with effects from year 2015.

In order to place all in the context we are situated, we must bear in mind that unto the author of a work with intellectual property the Revised Text of the Law of Intellectual Property (in later TRLPI) gives two types of rights, on the one hand the moral rights (article 14 TRLPI) and the other hand rights of property or exploitation (articles 17 et seq. of our revised text). The existence of these moral rights is emerging as one of the essential differences to Anglo-Saxon copyright law, where the absence of these rights results in that their defense may only happen in a contractual manner, and, in the absence of clause containing them, these moral rights from article 14 cannot be alleged.

This right, subject of controversy in the two aforementioned sentences, is integrated into the exploitation or heritage rights and it was already envisaged in the Berne Convention of 1886 for the protection of literary and artistic works in its articles 11, 11Bis and 11Ter. Interestingly, these articles do not use the designation of public communication (although they mention the right of transformation or translation), but this expression is used when a reference in article 11Bis to the "... public communication by means of loudspeaker..." is made.

Perhaps as heir to this agreement, the Directive 2001/29 of the European Parliament and of the Council of 22 May 2001 on the harmonization of certain aspects of copyright and other rights related to copyright in the information society, does not define in paragraph 1 of article 3[1], what there must be understood by a right to public communication, and therefore many problems, which we are faced with nowadays, come from there.

II. Public communication in legal environment

This precept tries to differentiate between 2 assumptions included in the field of public communication. On the one hand that which is generally called communication to the public, and on the other hand a concept more reduced and included within the first such as the right of making available.

Both issues are also regulated in our TRLPI, albeit with the novelty that a definition of public communication is given here. Article 20 states "that by public communication it would be understood any action whereby a plurality of people could have access to the work, without prior distribution of copies to each of them".

On the basis of these considerations it had been raised in the aforementioned two resolutions the question if the act of putting on a page a link on which you can click, and which takes us back to another page containing a work which is subject to copyright, does violate or not the right of public communication of the author. In both cases there is a prejudicial question posed to the European Court of Justice, in the first case by a Swedish court (case Svensson), and in the second (Bestwater) by a German court.

The issue is important because the use of including links to other resources is one of the bases of Internet and if the Court finds that there is infringement there, it would have hampered extremely that activity of adding links, because anybody can see the practical difficulties associated with the need to obtain the permission of the owners whenever a link should be included in a web page [3].

The first of the cases refers to journalistic publications communicated through the website of a newspaper in which two holders of rights on them work, and the second one arises from a promotional video of a company that is posted on youtube without authorization from their authors and that it is later linked on the website of competitors using the technique of framing or transclusion. Regardless now of the differences between the 2 cases, we are going to analyze quickly the conclusions of the European Court of Justice.

The Court differentiates 2 concepts in the right to public communication, on the one hand the communication and on the other its character of being public. Regarding the first issue, the European Court of Justice defends that an act of communication takes place in the form of making available through the links, and the Court expresses the same terms in the Bestwater case referring to what is affirmed in this sentence.

DOI: 10.9781/ijimai.2015.3511

Therefore, as it can be derived from article 3, paragraph 1, of Directive 2001/29, in order that there exists an Act of communication, it is sufficient that the work would be made available to the public, in such a way that their members could have access to it, being not decisive that such persons make use or not of that possibility.

On the second question, that communication would become public; the European Court of Justice says that "the protected work must be effectively communicated to a «public». For the purposes of article 3, paragraph 1, of Directive 2001/29, the term public refers to an indeterminate number of potential recipients and, moreover, involves a considerable number of persons".

This is what happens in the case of web links, and therefore we would be faced with a communication of the work to the public. However the Court points out that an authorization by the author is not required since in order for their asking so, the recipients of the work should conform a new public which should be different from the one who could have accessed the work potentially in its original location. Since the potential audience of the first communication is logically all Internet users, the audience to which the links are directed, are in essence the same, since on the first page there was no restrictive measure for the access by the users. The first obtained conclusion is clear; the linking with the indicated requirements is not a copyright infringement.

III. The Spanish Intellectual Property Law Reform

From January 1st of this year 2015 has entered into force the reform of our intellectual property law through the law 21/2014 4 November [4]. There are many issues facing this reform (especially in the rights management procedures), but we will focus essentially on two points that affect directly the issue we are dealing with.

First of all we will stop at the regulation known as rate Google or its technical name AEDE Canon. The cited rule establishes:

"The making available to the public by electronic providers of aggregation of content of non-significant fragments of content, reported in periodicals or on regularly updated websites which have an informative, a creation of public opinion or an entertainment purpose, shall not require authorization, without prejudice to the right of the editor or, in its case, to other holders of rights to receive fair compensation. This right shall be inalienable and it will be effective through intellectual property rights management bodies. "

This rule is about such services as Google News or Digg, because what they do is basically to establish links to other media news. They could continue doing it, but they should pay to the publishers the corresponding compensation.

Now, if, as we have seen above, the activity of linking with the analyzed requisites is lawful, we must ask ourselves to what extent the payment of this compensation is lawful too. We must also think that the existence of these news aggregators is beneficial to publishers in the majority of cases, since much of the traffic that gets to them is through those links, and that results in page access and ultimately in contracts of advertising.

About the regulation of this issue, it must be said that the experience shown by other neighboring countries has served very little. In 2007 a similar procedure was initiated in Belgium under the name of Copiepresse case [5], in which this Association of publishers of newspapers was facing against Google. The Belgian courts gave reason to Copiepresse considering that it violated the right of reproduction and the communication to the public, and Google was forced to cancel its linking activities in this country. The funny thing is that about two months later, the same entity felt obliged to ask Google to resume its activity in the face of the decline in visits to the portals of their affiliates. What seems clear is that what the Belgian editors wanted was not to stop appearing in the search engine, but to continue appearing and charging. Google response was very clear, all or nothing, or appear without charging or not listed. The backing down of the editors had got to give reason to the American giant.

Something similar has happened in France, although with the peculiarity that, in view of what already happened in Belgium, the precaution was taken of qualifying legally this compensation as waivable, which has allowed the newspapers to decide if it compensates them or not the activity offered by Google. In fact, the search engine will maintain full access to the content of the news editors and in return it will make a contribution of EUR 60 million to finance projects of digital media of French publishers.

These experiences should serve as a reference for future conflicts between Google and publishers from different countries, but in the Spanish case and ignoring what happened in the neighboring country, the voluntary resignation of the editors to it has been tried to be avoided by establishing legally the charging of this canon to be indispensable, which results in benefit of the corresponding rights management body. Thus once more the facts are showing that the management entities, far from being entities that comply with its obligation to promote culture (understanding within it, as it can not be otherwise, science and technology), are lobbies who currently occupy the role that in the Illustration was occupied by the Church [6].

It is one more try of these management bodies to look for compensation for the reduction of their incomes as a result of the expansion of the digital technology in the terms of copyright, and that intent was already evident in other areas such as p2p networks, where they have acted against the users of these networks [7][13].

We found the main problem about this issue in the poor drafting of the precept and the lack of clarity in what it wants to regulate. We will focus our criticism mainly on two issues.

On the one hand, the expression 'no significant fragment' could clearly be improved from the point of view of legal technique (whilst unknown in the intellectual property laws of countries of our environment).If we give the consideration of non-significant to the fragment, this limits the objective reasons for charging a compensation. It seems that it does not lead to appreciable harm that could justify charging for it. Therefore it seems illogical that if the European Court of Justice have come to recognize that the inclusion of contents of free access on the Internet in a different web site, through the technique of framing, is not a public communication and therefore it can be implemented, with more reason this activity can be made in relation to a snippet of the work. It seems clear that it is more damaging to the author to provide access to the work as a whole than to mere fragments of it.

On the other hand and in accordance with which was stated above, the qualification of this rate as fair compensation, and therefore forced collection through management bodies, excludes from itself the general and original operation of copyright: the author is the center of the intellectual property and as such he can decide both to charge for the use or exploitation of his work, and to give it away for free, without thereby limiting his creator status in any way. This second possibility is being limited deliberately through the drafting of the precept contained in our rule.

Relating all these issues with the interpretation established by the ECJ in the cited cases, there is no doubt for us that under no circumstances we would be faced with a breach of the right of public communication of the author in the activity of these portals. If there is not such a possibility in the analyzed cases, we could hardly find it in the activity from Google News. With this situation in mind we need to ask ourselves if we are violating another right of the author, such as

the right of reproduction. This is a different question, because in many cases a reproduction in the servers of the linker is indeed being made, so just as David Maeztu [80] rightly indicates, insofar as a reproduction is not played via link on the local servers, there will be no problem. It would therefore be enough to modify portals such as Google News to display directly the text without reproducing it on their servers, to make this tax to be meaningless.

An interesting issue related to this (though we will leave it to a future work) would be to analyze the possibilities that exist to get know previously if we are violating these rights on the part of the news portal, which leads us to analyze the relationship between law and artificial intelligence, or in other words, the possibility of applying to these cases criteria of artificial intelligence, which could permit us to discriminate, from a legal perspective, those offending cases of those which are not, by means of the application of reasoning and legal ontology. By the end of the sixties a series of artefacts came out which were called mostly legal expert systems, although the term was coined already in the mid-1950s in the field of Artificial Intelligence by the professor of Standford John McCarthy.

There have been several developments and works around the entailment between AI and law, which have focused not only on search management in document databases, but also in areas like legal decision taking or in the creation and development of legislative systems.

About the issue of this paper, the difficulties in the application of AI come from three aspects [9]:

- The system should recognize that the user query is formulated in legal language
- It should contain a computerized expression of the applicable legal rules
- It must bring online these rules together with the query, draw a conclusion and provide a legal response.

The difficulties in fulfilling these three requirements, which are based on the characteristics of legal language, which is interpretable and ambiguous, limit greatly the chances of success of these systems. In the same way that it is done in the field of education, the results depend on having sufficient data on carried out assessments, and the possibility of detecting anomalous behaviors in the development of new preventive and corrective actions [10]. Perhaps in this field and because of the described circumstances, the claims to carry out a reproduction of the human brain should be limited, and the focus should be put on the development of tools for the analysis of vast amounts of information, and so provide recommended guidelines, which in any case must be analyzed and contrasted by the human intellect to handle the actual postulation. It must be taken into account in any case that the definition of analytics have gone further in recent years, however, to incorporate elements of operations research, such as decision trees and strategy maps to establish predictive models and to determine probabilities for certain courses of action [11].

IV. The Regulation of Webs with Links

Another issue that the effects of the two resolutions that give cause for this work have been intended to apply on is the treatment that should be given to the pages of links that allow carrying out the downloading of protected works without counting with the appropriate rights.

We shall not dwell now in the evolution suffered by our jurisprudence in the legal qualification of these actions. Although we can say that, while in a first moment our courts defended that the activity of these sites did not constitute a crime against intellectual property, nowadays "the resolutions that consider that a offence is committed against copyright in the case of web pages with links to P2P networks, in which the owner of the website goes into the file-sharing site, extract from it a link to a particular file, film, music or other work, and incorporates it as a direct download element in his own website, without showing whatever information about the exchange type, so the user accesses the content directly from the page, have a greater legal weight "[12].

In relation with these pages the issue of the application of the criterion established by the European Court of Justice, and therefore their possible legality as an act of linking to intellectual property works, is raised. However, it seems clear as a matter of fact, that we do not assume the same course, since in these cases the link will not provide directly the work, but it is an instrument through which, and following a series of steps, we can finally access the work. For example in the case of p2p networks, the access to a work is completed by means of communication protocols and the download of specific software that allows us to get hold of the work in question, which clearly exceeds the mere activity of linking, which the CJEU makes reference to.

V. Conclusion

One of the issues that is more striking in the analyzed issue is the disconnection in some respects between the Spanish legislator and the decisions of the European Court of Justice. It is in fact a much more expensive point if (as in this case) the judgment is prior to the enactment of the concerned regulation. This has led to, that in the case that we deal on; it has even been declared that the Google tax established by the law of November 2014 is quite dead before birth.

And why is it so? It is due basically to the subjugation of our legislators (and by extension of the political power) to the guidelines and requirements of the management entities of intellectual property rights. These institutions had been marked as a result of the judicial procedures initiated against them on the basis of the bad administration of their funds in a not-too-distant time.

But it is not the first time that this situation occurs. In 2014 the CJEU had to admonish Spain because of the enforcement that was being made of the famous digital canon regulation. That enforcement was based on a single idea, such as it was the huge collection through this concept, by means of a establishing of an indiscriminate system that was contrary to the legal reality, which has recently got to be amended by imperative of the CJEU.

It seems clear that the rights management entities play a role, and that anyone who wishes to, can make use of them to carry out the management of the rights that the law grants. But we must not forget that the central figure of this system is the author and not the management entity, which often endorses the position of the first and gives the feeling that outside them there is nothing in the field of intellectual property, whilst in reality just the contrary is true: The best and most rewarding, that is the authors, their works and above all the freedom of the designer in the exercise of his rights, is outside them.

The Web 2.0 is changing our society without doubt. But not all parts of are cultural, economic and social system have the same adaptability. We are living "in interesting times", and we are bound to see how new paths will be tried for social change from individuals, communities and institutions.

If projects such as [14] and [15] will be connected with Web 2.0, the experience would also be incredible. It would stimulate to increase the adaptability, with these new paths explained in previous paragraph, in other knowledge areas.

References

[1] Sentencia del Tribunal de Justicia de la Unión Europea de 13 de febrero de 2014 ». En el asunto C-466/12. [Online].

[2] http://curia.europa.eu/juris/liste.jsf?pro=&nat=or&oqp=&dates=&lg=&language=es&jur=C%2CT%2CF&cit=none%252CC%252CCJ%252CR%

252C2008E%252C%252C%252C%252C%252C%252C%252C%252C%252C%252Ctrue%252Cfalse%252Cfalse&num=C-466%252F12&td=%3BALL&pcs=Oor&avg=&page=1&mat=or&jge=&for=&cid=331915

[3] Sentencia del Tribunal de Justicia de la Unión Europea de 21 de octubre de 2014. [Online].

[4] http://curia.europa.eu/juris/liste.jsf?pro=&lgrec=es&nat=or&oqp=&dates=&lg=&language=es&jur=C%2CT%2CF&cit=none%252CC%252CCJ%252CR%252C2008E%252C%252C%252C%252C%252C%252C%252C%252C%252C%252Ctrue%252Cfalse%252Cfalse&num=C-348%252F13&td=%3BALL&pcs=Oor&avg=&page=1&mat=or&jge=&for=&cid=331915

[5] M. Iglesia Andrés de la (2014). Comentario a la sentencia del Tribunal de Justicia de la Unión Europea en al caso Svensson: sobre la naturaleza jurídica de los enlaces a obras protegidas pro derechos de. [Online]. Actualidad Jurídica Uría Menéndez /37-2014.

[6] http://www.uria.com/documentos/publicaciones/4269/documento/e08.pdf?id=5461.

[7] Ley 21/2014 de 4 de noviembre de reforma de la Ley de Propiedad Intelectual. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2014-11404

[8] http://www.copiepresse.be/

[9] J. Cueva González-Cotera de la (2014). El nuevo canon a las universidades: tras la apropiación del canon digital para las copias privadas, la del open Access. El profesional de la información, 2014, marzo-abril, v. 23, n. 2, p. 187.

[10] E. Gil López, A. Castillo Sanz (2013). Legal Issues Concerning P2P Exchange of Educational Materials and Their Impact on E-Learning Multi-Agent Systems, in International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 2, June, pp. 73-78.

[11] D. Maeztu (2014). El canon EADE, muerto antes de nacer. Blog El derecho y las Normas [Online]. http://derechoynormas.blogspot.com.es/search?updated-min=2014-01-01T00:00:00%2B01:00&updated-max=2015-01-01T00:00:00%2B01:00&max-results=50.

[12] C. Fernández Hernández y Pierre Boulat (2015). Inteligencia Artificial y Derecho. Problemas y perspectivas. [Online].

[13] http://noticias.juridicas.com/conocimiento/articulos-doctrinales/9441-inteligencia-artificial-y-derecho-problemas-y-perspectivas/

[14] J. Campo-Ávila del, R. Conejo, F. Triguero, R. Morales-Bueno (2015). Mining Web-based Educational Systems to Predict Student Learning Achievements, in International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 3, Nº 2, pp. 49-54.

[15] A. G. Picciano (2014). Big Data and Learning Analytics in Blended Learning Environments: Benefits and Concerns, in International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 2, Nº 7, pp. 35-43

[16] Sentencia de la Audiencia Provincial de Castellón de 12 de Noviembre de 2014. Fundamento Jurídico 3º.J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility (Periodical style),” *IEEE Trans. Electron Devices*, vol. ED-11, pp. 34–39, Jan. 1959.

[17] Gil, E., and A. Castillo-Sanz, “Legal Issues Concerning P2P Exchange of Educational Materials and Their Impact on E-Learning Multi-Agent Systems”, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 2, issue Regular Issue, no. 2, pp. 73-78, 06/2013. Abstract

[18] Lorezo, W., R. Gonzalez-Crespo, and A. Castillo-Sanz, “A Prototype for linear features generalization”, International Journal of Artificial Intelligence and Interactive Multimedia, vol. 1, issue A Direct Path to Intelligent Tools, no. 3, pp. 59-65, 12/2010. Abstract

[19] Broncano, C. J., C. Pinilla, R. Gonzalez-Crespo, and A. Castillo-Sanz, “Relative Radiometric Normalization of Multitemporal images”, International Journal of Artificial Intelligence and Interactive Multimedia, vol. 1, issue A Direct Path to Intelligent Tools, no. 3, pp. 53-58, 12/2010. Abstract

Andrés Castillo-Sanz was born in Madrid in 1964. He has degrees in Physics, Sociology and Theology from the Complutense University of Madrid. The doctoral degree was obtained in 2004 at the Pontifical University of Salamanca. He teaches at International University of La Rioja and Pontifical University of Salamanca. He is researcher in in the field of applied multiagent systems, health sciences and social movements. He also participates in several research groups on different Universities in Spain and other European countries.

Eugenio Gil was born in Galicia (Spain) in 1972. He has a degree in Law from Deusto University in Spain. The doctoral degree was obtained in 2014 at the Pontifical University of Salamanca. He is researcher in Pontifical University of Salamanca (UPSA) and in International University of La Rioja (UNIR) in the field of Information Technologies Law. Actually he is the Academic Secretary at the School of Engineering and Architecture in Pontifical University of Salamanca, Madrid campus.