Advances in Intelligent Systems and Computing 276

Álvaro Rocha
Ana Maria Correia
Felix B. Tan
Karl A. Stroetmann *Editors*

New Perspectives in Information Systems and Technologies, Volume 2

Springer

Advances in Intelligent Systems and Computing

Volume 276

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

For further volumes:
http://www.springer.com/series/11156

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Álvaro Rocha · Ana Maria Correia
Felix B. Tan · Karl A. Stroetmann
Editors

New Perspectives in Information Systems and Technologies, Volume 2

Springer

Editors
Álvaro Rocha
Universidade de Coimbra & LIACC
Rio Tinto
Portugal

Ana Maria Correia
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa
Lisboa
Portugal

Felix B. Tan
Department of Business Information Systems
Auckland University of Technology
Auckland
New Zealand

Karl A. Stroetmann
Empirica GmbH
Bonn
Germany

ISSN 2194-5357 ISSN 2194-5365 (electronic)
ISBN 978-3-319-05947-1 ISBN 978-3-319-05948-8 (eBook)
DOI 10.1007/978-3-319-05948-8
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014934570

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This book contains a selection of papers accepted for presentation and discussion at The 2014 World Conference on Information Systems and Technologies (WorldCIST'14). This Conference had the cooperative support of AISTI (Iberian Association for Information Systems and Technologies/Associação Ibérica de Sistemas e Tecnologias de Informação) and IEEE Portugal Section. It took place between the 15th and 18th of April in Funchal, Madeira, Portugal.

The World Conference on Information Systems and Technologies (WorldCIST) is a global forum for researchers and practitioners to present and discuss recent results and innovations, current trends, professional experiences and challenges of modern Information Systems and Technologies research, technological development and applications. One of its main aims is to strengthen the drive towards a holistic symbiosis between academy, society and industry. WorldCIST'14 built on the success of WorldCIST'13, held last year in Olhão, Algarve, Portugal.

The Program Committee of WorldCIST'14 was composed of a multidisciplinary group of experts and those who are intimately concerned with Information Systems and Technologies. They have had the responsibility for evaluating, in a 'blind review' process, the papers received for each of the main themes proposed for the Conference: A) Information and Knowledge Management (IKM); B) Organizational Models and Information Systems (OMIS); C) Intelligent and Decision Support Systems (IDSS); D) Software Systems, Architectures, Applications and Tools (SSAAT); E) Computer Networks, Mobility and Pervasive Systems (CNMPS); F) Human-Computer Interaction (HCI); G) Health Informatics (HIS); H) Information Technologies in Education (ITE).

WorldCIST'14 received contributions from 42 countries around the world. The papers accepted for presentation and discussion at the Conference are published by Springer (this book) and by AISTI (another e-book) and will be submitted to be indexed by ISI, EI, SCOPUS, DBLP and/or EBSCO, among others. Extended versions of best selected papers will be published in relevant journals, including SCI/SSCI and Scopus indexed journals, and in a SCI series book.

We acknowledge all those who contributed to the staging of WorldCIST14 (authors, committees and sponsors); their involvement and support is very much appreciated.

Madeira, April 2014

Álvaro Rocha
University of Coimbra

Ana Maria Correia
University Nova de Lisboa

Felix B. Tan
Auckland University of Technology, New Zealand

Karl A. Stroetmann
Empirica Communication and Technology Research, Bonn

Contents

Software Systems, Architectures, Applications and Tools

Computer Networks, Mobility and Pervasive Systems

Human-Computer Interaction

Health Informatics

Erratum

Extending Groovy's Reification and Closures

Carlos Cortinhas and Fernando Barros

Universidade de Coimbra, Departamento de Engenharia Informática, Portugal
ccort@student.dei.uc.pt, barros@dei.uc.pt

Abstract. Groovy is a dynamic language that allows the easy modification of its own behaviour. A simple mechanism to introduce changes in the core language is by using Abstract Syntax Tree(AST) transformations. In this paper we describe two new features added to Groovy: the introduction of Reification, and the enhancement of Groovy Closures. In the first extension we make it possible for the user to know the types of a generic class during runtime. In the second one, Groovy closures were improved with the addition of two methods: one to provide closures return type and another to give the closure's source code.

Keywords: Abstract Syntax Tree Transformations (AST), Groovy, Reified Types, Generics, Closures, Lambda Expressions, Attribute Oriented Programming.

1 Introduction

This paper describes some limitations of the Groovy [5][14] language and our approach to eliminate them. The Groovy language is derived from Java and it also runs on the standard Java Virtual Machine. The main difference between Groovy and Java is that Groovy is a scripting language with dynamic capabilities. These features allow, for example, the development of Domain Specific Languages (DSL's) with less effort. Another useful characteristic of Groovy is the possibility of extending the language with its own resources [7], [12]. This type of programming is commonly called attribute programming [16]. Unlike other languages like Java, C# and C++ that are hard to extend, Groovy can be easily enhanced by using annotations and AST Transformations [11], in a very similar way of what Scala Macros [8] do in the Scala language, allowing for compile-time metaprogramming and also uses AST Transformations. There is also a project with the goal of creating an easy extensible programming language, MOSTflexiPL [13]. This language allows the user to create new operators and control structures, thus extending the language.

Since Groovy is derived from Java it inherits some of the Java's limitations like, for example, the lack of Reified Generics. Groovy also added new features like Closures (known as lambda expressions) that are available in Java 8 [2][6]. Closures API already provides useful information but some important and relevant information is not available. For example, Closures return type cannot be accessed at runtime making verification and validations more difficult. Another

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2,* Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_1,

feature that can be useful is the access to Closure source code at runtime. This would provide an easy access to the contents of a Closure in a user-readable manner. This feature would obviously help the programmer during development and debugging.

In this paper we describe our approach to extend Closures with runtime access to source code and return type. These extensions were developed using Groovy's feature, AST Transformations. These allow the programmer to create annotations and the respective transformation in order to modify the default compilation behaviour.

The paper is organized as follows. The state-of-the art is reviewed in Section 2, where Reified Generics and Closures/lambda expressions are presented. Section 3 describes Reified Types, and the access provided at runtime to the Generic types of a class. Section 4 describes Closure modifications needed to support the added features and methods. Section 5 gives paper conclusion and describes future work.

2 State of the Art

In this section we provide an overview of the current state of the art on Groovy and Java annotations and what are they used for. We also explain reified generics and what languages support them and finally we'll introduce Closures/Lambda expressions.

2.1 Java and Groovy Annotations

To extend Groovy we have chosen to use annotations and Abstract Syntax Tree (AST) Transformations [11]. This type of programming is not new. Annotations are available in Java since version 5 and they are a form of metadata. In Java, annotations can be applied to classes, methods and fields. There are some projects made in Java with the objective to extend Java's annotations like @Java [9][10]. @Java provides the capability to annotate expressions. In Groovy, like in Java, we can annotate classes, methods and fields. The difference is that in Groovy it is possible to develop an annotation combined with an AST Transformation to modify an application's code, this modification enables us to change the compilation process of the language. There are already some projects in Groovy that depend on annotation programming like GContracts [3][1], for example, that implements the concept of contract oriented programming, and GPars[4] which is a framework to simplify parallel programming in Groovy.

2.2 Reified Generics

The possibility to create parameterized classes with generic types has introduced new possibilities to programming. However, many languages like Java do not support reified generic types. Java does not support runtime generic type discovery since it discards this information after the compilation [15]. This happens in order to ensure that legacy code continues to run on newer versions of the Java Virtual Machine. Other languages, like C#, type information is kept in runtime so it can be accessed.

2.3 Closures and Lambda Expressions

Lambda expressions are also known as closures in Groovy or Clojures in Java 8. Unlike methods, lambda expressions are anonymous functions that can be handled as an ordinary object. Lambda expressions have access to the environment where they are defined and they can be moved around and passed as an argument to a method, for example. However, they do possess some limitations. For example after we create a Lambda expression, we cannot know its source-code and we cannot check its return types.

Lambda expressions are also available in other languages like C# where they are known as *Predicate*'s and *Func*'s. A *Predicate* has some limitations since it can only take one argument and will always return a boolean. *Func*'s are more similar to the Groovy closures since a *Func* can take multiple arguments and we can specify a return type. Next, is an example of a *Func*.

```
class Program{
static void Main(string[] args){
    Func<string, string> myFunc = (v1) => "Hello " + v1;
    Console.WriteLine(myFunc("World!"));
  }
}
Output: Hello World!
```

Listing 1. Example of a *Func* in C# initialized with a lambda expression

In this simple example, the *Func* takes a string as an argument and returns another string.

Additionally, C# does not allow lambda expressions to be serializable by default. We consider this a very useful feature and in our project we provide a way to serialize a closure in Groovy.

3 Extending Groovy with Reified Types

Reified types is a technique used to make generic type information available during runtime. We describe in this section an extension to the Groovy language for supporting reified generics.

3.1 Motivation

As stated in previous sections, neither Java or Groovy has generics reification. This is a feature that would be very useful for both these languages. The lack of reification of generics leads to some limitations to the language. We cannot use multiple dispatch like in the following example in Listing 2

```
public void method(List<Float> list) { ... }
public void method(List<String> list) { ... }
```

Listing 2. Example of multiple dispatch methods using the generic types

In the previous example, the type information of the *List* is lost and therefore we cannot use multiple dispatch like that. Another similar example is our inability to create parameterized exceptions like the one on Listing 3:

```
class ParameterizedException<T> extends Exception
try {
  throw new ParameterizedException<Integer>(100);
} catch (ParameterizedException<Integer> pe) { ... }
```

Listing 3. Example of parameterized exception

The above example would also not be possible for the same reasons that the multiple dispatch with generic types did not work. The compiler looses the type information of generic types at the compile-time.

To solve this issue and to make the generic types available at runtime we decided to turn to Groovy and we came up with a solution, the Reified AST Transformation that we will explain in the next subsection.

3.2 The Reified AST Transformation

For supporting reified types we have developed a new AST Transformation named @Reified. This transformation saves the types that parameterize a generic class, this happens during the compilation phase of the application. To use this transformation a class needs to be annotated with the @Reified keyword, as shown in Listing 4.

```
class Book{}
@Reified
class Library <T, V> {
  static main(args) {
    Library<Book, Integer> lib = new Library<Book, Integer>()
    println lib.getReifiedTypes()
  }
}
Output: [Book, class java.lang.Integer]
```

Listing 4. Example of @Reified

Our @Reified transformation algorithm is very simple, it simply checks for parameterized declarations on every annotated source file. When such a declaration is found (line 5 of Listing 4), the generic type information is saved in the declared object and becomes accessible through the method *getReifiedTypes()*. In the example we can observe that printing the result of *getReifiedTypes()* will return the types that parameterized the object when it was instantiated.

In addition, the transformation also does type checking to guarantee that the types match. This verification occurs since Groovy, by default, does not type check these parameters automatically. In the previous example the transformation checks the parameters, which in this case are the user defined class Book and the Integer type. If the parameters don't match, an exception is thrown when the application compiles.

4 Closure Modifications

Another extension developed is the modification of the Groovy's Closures by adding some features that help the programmer to know the source code of a closure, and a method to retrieve the return type of a closure.

4.1 Motivation

Groovy Closures are executable blocks of code that can be assigned to a local variable/field. Like variables, they can be used as arguments and executed inside methods. The closures have access to the local environment where they are placed. In a distributed system we sometimes want to pass objects, in a serialized format, between applications. However, in the current Java and Groovy versions we cannot serialize a closure at one node, send it to the destination node, deserialize it and execute the closure. To solve this issue we choose to save the closure's source code in the format of a string. Now, in a distributed system, we can easily send the closure's source code to other nodes where the string will be compiled into a closure and be ready to be executed.

Closures, in Groovy, do not support runtime information about its return type, which is a problem on a dynamically typed language. In our extension closures return types are also made available through the use of a specific method.

4.2 The Closures Transformation

The current implementation of closures has several limitations. During runtime it is not possible to know the closure source code since there is no method available to access this information. Closures also lack a feature to retrieve the return type parameters.

To develop a fix for the first limitation we developed the *@Closurespp* AST transformation. This transformation will save the closure's source and make it available on runtime through the use of a new method called *getCodeToString()*.

```
1  @Closurespp
2  class TestClosures {
3    static main(args) {
4      Closure<String> a = {
5        "testing"
6      }
7      println a.getCodeToString()
8    }
9  }
10 Output:
11 {
12   "testing"
13 }
```

Listing 5. An example of the Closure.getCodeToString() method

As we can see in line 7, the method returns the source code of the closure. This helps the programmer to identify the closure during runtime.

The issue of the return type feature was addressed in a very similar manner and the *@Closurespp* transformation was also used to implement this feature. Every time a closure is detected, the return type is stored in the object and becomes available at runtime through the method *getReturnType()*.

```
class Book{}
@Closurespp
class TestClosures {
  static main(args) {
    Closure<Book> a = {
      Book b = new Book()
      return b
    }
    println a.getReturnType()
  }
}
Output:
[class Book]
```

Listing 6. Example of Closure.getReturnType() method

The output of the example in Listing 6 is the expected, as it returns the user made object *Book* specified in the closure's declaration. This feature also helps the programmer to correctly identify a closure and know its return type during runtime.

5 Conclusion and Future Work

We consider that our extensions to Groovy bring useful features. The Reified transformation allows us to retrieve the generic types during the runtime that was not possible in Groovy, or in native Java. Closure modifications introduced two new methods: one to retrieve a closure's return type, and another to return closure's source code. For future work, we will develop our AST Transformations to be inheritable since they currently are not. A current limitation is that multiple dispatch on methods is not yet available, but we plan on developing our @Reified transformation in order for it to support multiple dispatch on Groovy methods with generic arguments. Additionally we plan to develop a Global AST Transformation to apply our two transformations, the benefits of this approach is that there would be no need to annotate the classes anymore as the global transformation would be applied to all source files compiled.

References

1. DZone GContracts 1.0, `http://groovy.dzone.com/announcements/gcontracts-10-released`
2. First Lambda Expression, `http://diarmuidmoloney.wordpress.com/2011/12/10/my-first-java-8-lambda-expression/`
3. GContracts A Programming by Contract Extension for Groovy!, `https://github.com/andresteingress/gcontracts/wiki`
4. GPars Groovy Parallel Systems, `http://gpars.codehaus.org/`
5. Groovy, `http://groovy.codehaus.org/`
6. Java 8 - Closures, Lambda Expressions Demystified, `http://frankhinkel.blogspot.pt/2012/11/java-8-closures-lambda-expressions.html`
7. Axelsen, E.W., Krogdahl, S.: Groovy package templates: supporting reuse and runtime adaption of class hierarchies. ACM SIGPLAN Notices 44(12), 15–26 (2009)
8. Burmako, E.: Scala macros: Let our powers combine. In: Proceedings of the 4th Annual Scala Workshop (2013)
9. Cazzola, W., Vacchi, E.: Fine-grained annotations for pointcuts with a finer granularity. In: Shin, S.Y., Maldonado, J.C. (eds.) SAC, pp. 1706–1711. ACM (2013)
10. Cazzola, W., Vacchi, E.: @java: annotations in freedom. In: Shin, S.Y., Maldonado, J.C. (eds.) SAC, pp. 1688–1693. ACM (2013)
11. Cortinhas, C., Barros, F.: Extending the groovy language using ast transformations to monitor variables and methods. In: Rocha, Á., Correia, A.M., Wilson, T., Stroetmann, K.A. (eds.) Advances in Information Systems and Technologies. AISC, vol. 206, pp. 745–752. Springer, Heidelberg (2013)
12. Dinkelaker, T., Mezini, M.: Dynamically linked domain-specific extensions for advice languages. In: Proceedings of the 2008 AOSD Workshop on Domain-Specific Aspect Languages, 3 p. ACM (2008)
13. Heinlein, C.: Most-flexipl: modular, statically typed, flexibly extensible programming language. In: Proceedings of the ACM International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software, pp. 159–178. ACM (2012)
14. König, D., Paul, P.K., Laforge, G., Jon, J.S.: Groovy in Action, 2nd edn. Manning (2011)
15. Naftalin, M., Wadler, P.: Java generics and collections. O'Reilly (2009)
16. Wada, H., Suzuki, J.: Modeling turnpike frontend system: A model-driven development framework leveraging uml metamodeling and attribute-oriented programming. In: Briand, L.C., Williams, C. (eds.) MoDELS 2005. LNCS, vol. 3713, pp. 584–600. Springer, Heidelberg (2005)

The Connective Scheme between the Enterprise Architecture and Organisational Structure

Tiko Iyamu[1], Leshoto Mphahlele[2], and Suama Hamunyela[1]

[1] Namibia University of Science and Technology, Windhoek, Namibia
[2] Tshwane University of Technology, Pretoria, South Africa

Abstract. The demand for better services by customers and citizens keeps increasing at a rapid rate. Different methods and techniques are sought by various organizations to enable and support processes and activities to improve services, as well as to give the organization leverage towards competitive advantage. The enterprise architecture (EA) has merged as a possible solution for addressing organizational challenges and for competitive advantage. The EA deployment involves agents, which are both human and non-human. The agents, based on their interest, influence and determine how the EA is deployed. During the deployment of EA, agents transform themselves in accordance to their interest at the time and space.

Understanding of agents' interests is significant if the challenges that they pose are to be managed for successful deployment of EA. Based on this challenge, a theory, Structuration which focuses on human and non-human agents was selected to underpin the study. The study investigated the impact of agents on the deployment of EA in organizations, using the lens of structuration theory in the analysis of the data.

Keywords: enterprise architecture, information technology, structuration theory, organization.

1 Introduction

Many organizations, including government agencies, have realized the significance of the EA. There is a growing interest in organizations to adopt the EA, to assist mitigating the challenges of constant and complex changes [1]. Some of the factors that pose challenges to competitive advantage of organizations are adaptiveness, uniformity and scalability. EA is a tool that enhances the organization in terms of the objectives pursued at the time. This is argued to be done through its capability to improve the business processes, applications, data and infrastructure through standardization [2].

The EA consists of different domains, which include business, information, application, technology, infrastructure and service oriented architecture. According to [3], the focus and deliverables of the EA domains manifests in the relationship and

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_2, © Springer International Publishing Switzerland 2014

connectivity among the domains during the implementation phase of the EA. Thus, the completion of each domain contributes towards the success of developing and implementing the EA in the organization. However, organizations continue to fail to define the individual domains and their connectivity, thereby increasing complexity in the deployment of EA [4].

Each of the domains of EA has its distinct roles and boundaries. According to Hafner and Winter[5], the business architecture represents the enterprise model, which includes service exchanges and financial flows in value networks and strategies. The domain of information architecture covers data and information of all types, and their usage, interrelationships and demographics, as well as their definitions, ownership, distribution, and composition [6]. The application domain provides a foundation and guidance to the organization, on how to develop and implement application for business competitive advantage [7]. According to [8], the "Technical dimension covers the technologies and technological structures used to build the information and communication systems in the enterprise".

EA is developed and implemented sequentially, in accordance to the domains: from business to technical architectures. This is to enable logical flow and manage the dependence of the domains. During the development and implementation interactions happen among technical and non-technical agents (factors). Some of the factors include people, process and technology. According to [9] the development of the EA means adopting approaches that incorporate the people, processes, business and technology to achieve the organizational goals and objectives. The interaction that takes place during development and implementation makes the processes and technologies to be understood and interpreted differently by individuals and groups. This is often to assert individual power. Iyamu [10] argued that during the implementation, power is exercised to protect individual's interests, which shapes the outcome of information technology strategies in the organization.

The primary aim of the study was to understand the impact of social dynamics as embedded into organizational structure, on the deployment of the EA in organizations. Based on this goal, interpretive case study was employed, and structuration theory (ST) was selected as the theory that underpinned the study.

Structuration theory involves and focuses on processes and interactions that take place between actors and structural features within social context. The key tenets of structuration theory are agent (or agency) and structure. Agents are both technical and non-technical. To be categorized as an agent, individuals or technical artifacts need to have the capability to make a difference [11].

2 Literature Review

The scope of enterprise architecture (EA) includes various domains such as information, business, and technical within an organization. The domains constitute both technical and non-technical artifacts, such as people, process and technology. The scope of EA emphasizes processes, people, information and technology, and their relationship [12]. The EA is intended for various objectives, such as transformation, from one organization to another. As organization transform, both human and non-human agents also transform. According to [13], the Enterprise Architecture Management is intended to support and coordinate enterprise transformation. [14] argued the significance of the EA in organizations, and that it can assist in the development of strategies which can be incorporated into different areas such as: (1) existing processes, (2) ICT infrastructure and (3) other modules of the organizations.

Despite the increasing interest and the growing need for EA, there remain many challenging factors, which include process, people and technology. According to [15] there exist many challenges and problems continue in the development and implementation of EA. These challenges are rarely technical, but arise from process and people, which are often connected to political. According to [16], some of the critical elements to consider when adopting the EA are processes, communication, organization politics, and ownership.

3 Research Methodology

The case study research method was adopted. The strengths of the case study approach is in its flexibility to allow the researcher to freely use his or her ability to attain a variety of information within context [17]. A South African company, Bakgaga Network Services (BNS) was used as the case study. The company selection was based on two main factors, accessibility and evidence of EA deployment. According to [17] the most commonly used techniques for collecting data in a case study includes interviews.

The semi-structured interview technique was employed, which enable the research to probe (in depth) why things happened in the way that they did in the environment. [18] emphasises the critique of knowing why, where and how to collect data. According to Gray [14], the ability of probing further questions in the semi-structured interview allows the exploration of paths that have not yet been considered. A total of 14 employees were interviewed. The employees were from both business and IT departments, and at different levels of the organizational structure. The data was interpretively analyzed, using the duality of structure of the structuration theory.

As shown in Figure 1 below, the duality of structure covers structure, modality and interaction. In structuration theory, human actions are enabled and constrained by structures, yet that these structures are the result of previous actions [19].

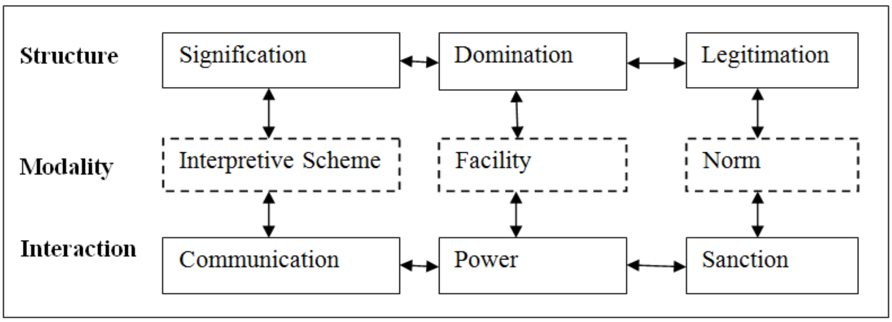


Fig. 1. Duality of Structure (Giddens, 1984)

Interaction happens between agents in a recursive characteristic manner within structure which they are a part. As depict in Figure 1, interaction with structure is through modality, which is categorized into interpretative scheme, facility and norms. The dependency of agents and structure on each other to produce and reproduce actions overtime is referred to as duality of structure.

3.1 Data Analysis: Structuration View

The objective of the study is to investigate and understand the social dynamics embedded within the organizational structure, and their role and impact on the deployment (development and implementation) of the EA. The ST was adopted to facilitate the investigation and analyse the data. In the study, data was collected from one organization, using the interview technique. The analysis of the data is presented as follows:

Agent

In structuration, agents are both human and non-human. The deployment (development and implementation) of the EA in the organization involves employees (human agents) such as the Chief Information Officer (CIO), IT architects, business analysts, IT managers, and project managers.

The development of the EA was carried out by the IT architects, under the leadership of the Chief Architect in conjunction with the CIO. The employees that were involved in or responsible for the deployment were not necessarily the ones who were involved in the implementation of the EA in the organization. The implementers of the EA include employees such as IT managers, IT architects, and other technologists. According to one of the employees, *"The offices of the CIO in conjunction with the Chief Architect are responsible for the development, as well as the implementation of the EA in the organization"*.

In the deployment of the EA, the employees are confined and act within structure (rules and resources). They were perceived to have some form of degree of influence during the deployment of the EA in the organization.

Structure

The deployment of the EA in the organization was carried out through adherence to rules and regulations, using available resources. The rules and regulations, including governance and standards were formulated within the organizational requirements as mandated by the organization.

The resources required in the deployment were vitally critical in determining the deployment of the EA in the organization. The resources included human (skills), finance and technology (software and hardware). One of the interviewees stated: *"the resources used in the deployment of the EA are driven by the organizational strategy and requirements"*.

Agents and structures depend on each other. Agents operate with structures, and structures exist or are created by agents. In the process, they both produce and reproduce actions in duality.

Duality of Structure

The duality of Structure enacts the interaction between human and the structural properties (rules and resources) during the deployment (development and implementation) of the EA in the organization. The structures and interactions that took place during the deployment of the EA were of socio-technical nature. In Table 1 below, the structuration theory analysis of the deployment of the EA is summarised. The discussion that follows after the table should be read together with the table to get a full appreciation of the analysis with duality of structure in the deployment of the EA in the organisation.

Table 1. Enterprise Architecture Deployment

Significance	**Domination**	**Legitimation**
The EA was essentially important in the organisation. Hence its accountability was with the CIO, at the highest level in accordance with the organisational hierarchy.	Under the auspices of the CIO, the IT architects and managers were responsible for the deployment of the EA in the organisation.	The EA is deployed within rules and regulations, as mandated by the organisation.
Interpretative scheme	**Facility**	**Norms**
The EA documents the current state and defines the strategic direction of technology, as well as business processes in the organisation.	The organisation uses different technologies, individuals' roles, responsibilities and skills during the deployment of the EA.	The EA was deployed in accordance with organisational definition of business processes, standards, rules and regulations. The deployment was carried out along on EA domains' perspectives.
Communication	**Power**	**Sanction**
There was interaction between the CIO and the Chief Architect. There was little communication between the CIO and the domain architects who carry out the actual deployment of the EA.	The CIO uses the organisational mandate to instruct the deployment of the EA. IT architects and managers used authority bestowed on them to protect their individual interests during the deployment of the EA.	The deployment of the EA was approved by the CIO, chief architect, the domain Architects and other stakeholders.

Duality of Structure: Signification and Communication

In the organization, the deployment of the EA was critical for both IT and business strategies. In the organization, the EA was generally considered important mainly for competitive advantage, through its technological and business processes. A senior manager expressed his view as follows: *"the EA plays a very critical role, mainly because it bridges business strategies with information technology capabilities in achieving the organizational strategies"*.

The EA was carried out by different individuals and groups (units) within the organizational structure. The individuals' roles and responsibilities were critical in the deployment of the EA, and as such, needed to be clearly defined. The deployment of the EA requires collaboration among all stakeholders. As such, employees who were involved in the process needed to fully understand their roles, responsibilities and accountabilities. Understanding was vested in the communication among the stakeholders and their interpretation of the events, processes and activities. Many of the employees in the organization seemed to understand the significance of their individual roles, as one of the interviewees put it: *"If the individual's roles and responsibilities are not clearly defined, you can delay the deployment of the EA in the organization, and this could put the company at risk"*.

The deployment of the EA was based on the organizational strategy, technology and business requirements. The IT architects and business analysts gathered the requirements from the business managers and users. The requirements were interpreted and translated into architectural requirements, thereafter, documented. The tasks of coordinating the gathering of both technology and business requirements and translation to architectural requirements were performed by the IT architects.

The organization had a procedural approach, which was used to allocate tasks for the deployment of the EA in the organization. The tasks were allocated and communicated to teams and units based on their accountability, specialties and area of focus. Both the IT and business units shared the tasks depending on the domain. The tasks were drilled down from units to individual level. One of the managers explained in detail: *"Normally tasks are assigned by the project owner. These tasks correspond with individual skills, experience and responsibilities, for example, a business analyst will be responsible for gathering and analyzing the business requirements"*.

Unfortunately, the communication channel as defined by the organizational structure was not strictly or often followed. Employees seem to be divided into groups and networks of friends as opposed to the organizational structures and channels. The employees associated themselves in accordance with where and with whom they find comfort. According to one of the employees, a senior manager, *"we are one team in IT, but there are different groups within the team. The reason for the formation of the groups is because some people want to shine and share information among their groups"*.

Due to the lack of structured communication among the stakeholders, the processes and efforts in the deployment of the EA was a challenge and became difficult to manage in the organization. Implications of the lack of structured communication included lack of collaboration, which often derailed the deployment of the EA through certain factors such as personalization of events and reluctance to cooperate. It also manifests to lack of synergy between business and IT stakeholders in the deployment of EA in the organization.

The deployment of the EA was influenced by certain factors, which were both human and non-human. These factors led to some individuals and groups being dominant and powerful during the deployment of the EA in the organization.

Duality of Structure: Domination and Power

Under the auspices of the CIO and sectional heads of departments, the IT unit was instructed and made accountable for the deployment of the EA in the organization. One of the senior managers explained:*"…the CIO and senior personnel in the organization within IT are obliged to ensure that EA is deployed to deliver on the business objectives"*.

All technology including the EA matters were made the sole responsibility and accountability of the IT unit. However, the participation of the business was critical, because EA was driven by the organizational strategies and objectives. Unfortunately, the business division participation was a challenge. This challenge was a derailing factor in the deployment of the EA. The challenge was a derailing factor mainly because the IT unit depended on the business for funding of its activities, processes and events.

As has been established so far in this analysis, the deployment of the EA involved many people (employees) from different units of the organization, which includes varieties of specialized skills and based on requirements as defined by the organizational structure. Some of these employees were responsible for enabling and managing the processes and activities during the deployment of the EA.

However, the organization didn't seem to have sufficient skills and knowledgeable employees to facilitate and carry out the activities that were involved in the deployment of the EA. The deployment of the EA was done through its domains, which included information, business, application and technology architectures. The insufficient skills posed a serious challenge during the deployment of the EA in the organization. Some of the interviewees as echoed by their colleagues commented: *"I used to work with the EA team, and I realised that most of the people where wearing big shoes "they are not fit to hold the positions as Architects"*.

The roles and responsibilities played a vital role during the EA deployment. The roles and responsibilities of the individuals were influenced by the relationship many of them had with their colleagues. The relationships were built on working together over the years and between employees with good (acclaimed better) knowledge of the subject. Some of the employees used the relationship that they had with their colleagues to influence decisions during the deployment of the EA. One of the interviewee's observed: *"Sometimes when you propose something relating to enterprise architecture, it feels like you stepped on someone's domain. This is because some people believe they own certain things in the organization, so they become negative towards change and they could dictate the activities of the EA"*.

Influences as controlled and managed by certain individuals were caused mainly because there was no proper performance contract and appraiser system in place. Managers could use their prerogative and the mandate bestowed on them to make decisions as they wished.

Unfortunately, some employees used their knowledge and expertise of the technologies to pursue reasons other than the organization's interest in EA deployment. These employees used their knowledge as a source of control and manipulation in the decisions that were made during the deployment of the EA in the organization. These employees used their expertise and stock of knowledge of the technologies as mechanism to control and demonstrate authority over their colleagues when it came to carrying out technological tasks during EA deployment. Such practices had an impact on the deployment of the EA in the organization. An interviewee explained: *"People guide their domains, guide their intellectual property (IP), and don't want to move or change. It is about job security"*.

Some employees in the organization overlooked the importance of the EA deployment, instead focused on personal gains or interests. These groups of employees resisted and reluctantly accepted tasks that were allocated to them, whenever they felt that events and activities in the deployment of the EA were not in their interest. Such practices had an impact on the deployment of the EA as it discouraged some employees from participating in activities and processes of the EA. One of the senior managers explained that: *"You find that there are people who run their own regimes, and if the EA is not going to benefit them and it's of no use to participate in the process"*.

The organization formulated a mandate to ensure the deployment of the EA. Different employees carried out the mandate using various technologies at various times. The employees' roles and responsibilities, including the stock of knowledge they acquired over time were instrumental during the deployment of the EA. The manifestation of individual's interests and agendas was a huge challenge that threatened the efforts of the organization in achieving the benefits of the EA.

Duality of Structure: Legitimation and Sanction

The deployment of the EA followed and adhered to the organizational rules and regulations; thereby making it required approval from various authorities along the organizational structure. The rules and regulations including available resources ensured that the deliverables of the EA fulfill the business requirements and needs for competitive advantage. A manager expressed her view as: *"The business rules are what drive the EA deployment. If there is a rule that specifies what to do or not to, this decision is critical to assist the configuration of the different domains of the EA"*.

The managers and team leaders were responsible for ensuring that the deployment of the EA was carried out within the rules and regulations of the organization. They also enforced and ensured that their staff adhered to those rules and regulations in carrying out their individual and group tasks.

Even though the managers did all they could to ensure that employees carried out their tasks as allocated and adhered to the rules and regulations, some of the employees were reluctant to abide. The rules and regulations were difficult to enforce mainly because the culture of the organization allowed individual independence and there was no performance contract and appraiser within which employees' activities could be managed. One of the team leaders explained: *"In our case, corporate governance is a set of rules, regulations and guidelines. We expect you to adhere to them, but if you don't adhere to them, you won't be penalized"*.

As a result of such a culture of individual independence and flexibility, many standards and principles existed in the environment. Some of these standards and principles were not documented, they were individualized. This resulted to duplications of processes and activities in the deployment of the EA in the organization. This implication defeated some of the benefits and intended deliverables of the EA. One of the interviewees explained: "*Because we followed different standards, the work was not consistent, because we were all from IT but when we went to business we provided different things*".

The executive committee in the organization approved the deployment of the EA. The decision was primarily based on the available resources including the organizational rules and regulations. In summary, rules, regulations and resources were vital and critical; hence it was significant how they were applied in the deployment of the EA. This makes the organizational structure very important, mainly because different degrees of power are associated with and assigned to various levels in the organizational structure.

4 Findings and Discussion

From the analysis, using the duality of structure from the perspective of the structuration, some factors were found to be critical in the deployment of the EA in the organization. The factors as depicted in Figure 2 below were viewed as critical, because they could play a multifunctional role, as enablers and with the same constraints. The factors include communication channel, task allocation, and awareness, stock of knowledge, skills transfer, and value measurement. These factors are discussed below:

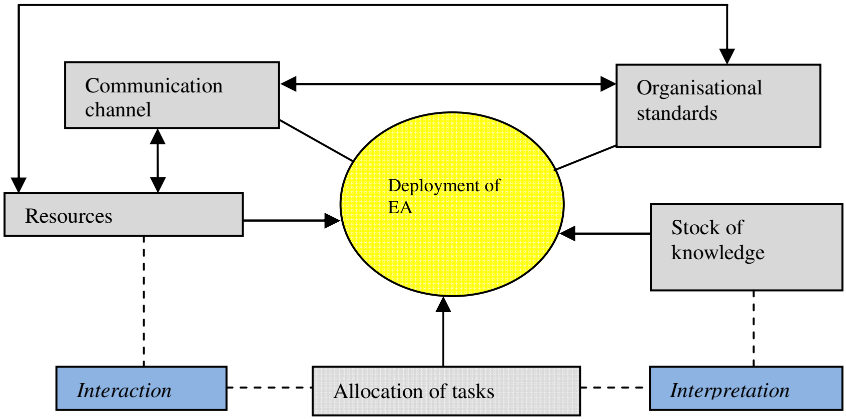


Fig. 2. Factors influencing EA deployment

Communication Channel

Communication is critical. Through it, awareness is gained; knowledge is transferred; understanding is achieved; and much more. It is even more important that communication is carried out through the channel; otherwise, the purposes and intentions as highlighted above would be defeated.

A communication channel was critical to the EA deployment in the organization. It was intended to enable synergy between the EA stakeholders. This was not the case in terms of the relationship between the IT and business units. Not only was the communication lacking between the IT and business units, but due to the individual and group practices, communication within the IT unit itself was hampered, resulting in factors such as personalized interests being more functional.

As revealed by the analysis of the data, the business unit had lost confidence and trust in the organization's IT unit. This was a manifestation of the lack of information and communication among the stakeholders including the business unit. The lack of information sharing and communication through structured channels continuously hamper the deployment of the EA. The business unit, which was the primary sponsor, thought that the IT unit was not delivering on its strategic mandate. As such they began to withdraw their support for the EA deployment.

Even when the information was shared and communicated, the content was perceived and interpreted differently by the EA stakeholders in the organization. Some employees relied on their individual stock of knowledge to convey and interpret the deployment of EA and the tasks they were assigned to perform.

Stock of Knowledge

In terms of stock of knowledge, two things are fundamental, how knowledge is stored, and how it is applied. Knowledge is often filtered before it is stored, otherwise, it would be of little or no value. On the other hand, if knowledge is applied appropriately, it would defeat its intended aim, as well as lose value.

It was critical for the organization to ensure that the tasks allocated to individuals, complemented and were aligned to their skills, knowledge and experience. This is mainly because the EA deployment relies intensively on best practises and skills. This has proved to be a challenge as most of the participating stakeholders, particularly, newly appointed IT architects seemed to lack the necessary skills and knowledge required for EA deployment. Those who had experience and knowledge of the EA held on to it, and used it as their sources of power to dominate and negotiate things of personal interests such as promotion and remuneration increases.

The analysis of the data also revealed that the inexperience and the lack of stock of knowledge had implications and an impact on the deliverables of the EA in the organization. This limitation provided some people with leverage over others. The majority of the people used it to command the space they occupied. The chief architect had the authority to dictate the approaches and methods through which the EA could be deployed. Certainly, the knowledge and skills for the EA deployment was lacking in the organization. The deployment of the EA was challenged because a lack of knowledge hampered people's capabilities to perform tasks and utilize EA technologies, approaches and methods.

Task Allocation

The tasks were an important aspect of the deployment of the EA. They represented the work that needed to be done. It was clear that the organization needed to clearly define its tasks, ensuring that employees were aware of the accountability and expected contributions of the EA deployment.

Unfortunately, many of the employees found fault on how tasks were allocated in the deployment of the EA in the organization. As a result, some employees did not, and other reluctantly accepted the tasks that were allocated to them. The fault was related to the claim that some of the tasks allocated to them did not align with their skills and experiences, and that some of their colleagues were favored.

Due to the lack of satisfaction among the employees, some of them did not take full responsibility for their lack of participation during the deployment of EA. Also, the allocation of tasks enabled the organization to assess the required technology, its usage and distribution to different teams. This was not the case with Bakgaga Networking Services, because evidence showed that people who have been in service longer had the authority to choose who they wanted to work with, and also how resources were to be utilized in the organization.

Organizational Standards

The organization mandated the deployment of the EA to be carried out within specified requirements and standards. The standards were employed to ensure that the EA delivered on its mandate to serve, enable and support the business requirements, processes and activities, using available resources. The standards enabled and at the same time constrained employees' actions during the deployment of the EA. It was important to establish such control measures, because employees' actions tend to be driven by individual beliefs and interests as opposed to organizational objectives. The standards were sometimes constraining factors, as some of them (standards) were not compatible or easy to integrate. But it was clear that some employees in the organization didn't adhere to the organizational standards. This had impact on the EA deployment. For example, some technologies were duplicated.

The analysis revealed that the standards were often not enforced by the employees. This was attributed to the fact that employees' actions were not monitored during the deployment of the EA. The culture was that some employees defined their own standards that had an impact on the EA, causing their actions to derail processes and activities.

Some of the employees explained that although there were organizational standards in place, some of the employees did not adhere to them as a form of protest. This resulted from a claim that some senior personnel promoted individuals to the architect positions based on their relationship, instead of their competency to perform the work. This culture is a legitimised practice in the organization, based on the power and mandate bestowed on the managers, that they could apply their prerogative to make final decisions. As a result, the process of advancing employees to positions as defined by the organization was often not adhered to.

Resources

Resources that included technology, people and processes were crucial during the deployment of the EA in the organization. The different technologies, which were necessary to facilitate the EA deployment, were always at the disposal of certain individuals and teams. The use and distribution of these technologies was intended to be managed through organizational standards. Unfortunately, these technologies were used as mediums for championing personal agendas, accumulating control and authority. Some of the employees in the organization withheld what was bestowed to them to nourish their self-interests.

The analysis highlighted that the significance of resources to them confirms their importance, authority and commanding capabilities; hence they constantly protect the resources because without them, they would be perceived as redundant and powerless in the organization. As a result of such human practices, the EA deployment could suffer setbacks. It is important that the use and distribution of such resources is enforced and instructed by the organizational standards at all times.

The findings from the analysis help to understand some of the factors that influence the deployment (development and implementation) of the EA in organizations. The factors as presented above were further examined (interpreted). The interpretations of the findings are discussed in the section that follows.

5 Conclusion

Without EA, organisations could fail to have a holistic view, monitor and manage its business processes, information flow, technology selection and deployment. Business and IT units' alignment and organisational structure often have an impact on and influence the deployment of the EA. How these factors impact and influence the deployment of the EA in organisations are not known. The outcome of the study helps managers of EA to gain a better understanding of the critical factors that have an impact on and influence the deployment of the EA in their respective organisations. Some of the factors were not known to them, and those that they seem to be aware of were not empirically proven. The empirical nature of the study will give the managers more confidence and prove in their actions and responses. The study contributes to existing academic literature. Currently, there are not many academic studies on the topic of EA. Also, there are no subjects and courses on EA in institutions of higher learning in South Africa and many other countries.

The outcome of the study helps managers of EA to gain a better understanding of the critical factors that have an impact on and influence the deployment of the EA in their respective organizations. Some of the factors were not known to them, and those that they seem to be aware of were not empirically proven. The empirical nature of the study will give the managers more confidence and prove in their actions and responses.

The study contributes to existing academic literature. Currently, there are not many academic studies on the topic of EA. Also, there are no subjects and courses on EA in institutions of higher learning in South Africa and many other countries.

References

1. Yu, E., Strohmaier, M., Deng, X.: Exploring intentional modelling and analysis for Enterprise Architecture. Paper Presented at the Enterprise Distributed Object Computing Conference Workshop, 10th IEEE International EDOC 2006 (2006)
2. van der Raadt, B., van Vliet, H.: Designing the enterprise architecture function. In: Becker, S., Plasil, F., Reussner, R. (eds.) QoSA 2008. LNCS, vol. 5281, pp. 103–118. Springer, Heidelberg (2008)
3. Iyamu, T.: Theoretical analysis of Strategic implementation of Enterprise Architecture. International Journal of Actor-Network theory and Technological Innovation 2(3), 17–32 (2010)
4. Glissmann, S.M., Sanz, J.: An Approach to Building Effective Enterprise Architectures. Paper Presented at the Proceeding of the 44th Hawaii International Conference on System Sciences (2011)
5. Hafner, M., Winter, R.: Process for Enterprise Application Architecture management. Paper Presented at the Proceedings of the 41st Annual Hawaii International Conference on System Sciences (2008)
6. Iyer, B., Gottlieb, R.: The Four-Domain Architecture: An approach to support enterprise architecture design. IBM Systems Journal 43(3), 587–597 (2004)
7. Pereira, C.M., Sousa, P.: A method to define an enterprise architecture using the Zachman framework. Paper Presented at the Proceedings of the ACM Symposium on Applied Computing (2004)
8. Pulkkinen, M.: Systemic Management of Architectural Decisions in Enterprise Architecture Planning. Four Dimensions and Three Abstraction Levels. Paper Presented at the Proceedings of the 39th Annual Hawaii International Conference on System Science (2006)
9. Wang, X., Zhao, Y.: An Enterprise Architecture development method in Chinese Manufacturing Industry. IEEE Journals 3, 226–230 (2009)
10. Iyamu, T.: The Factors Affecting Institutionalisation of Enterprise Architecture in the Organisation. Paper Presented at the Proceeding of the 11th IEEE Conference on Commerce and Enterprise Computing, Viena (2009)
11. Giddens, A.: The constitution of society: Outline of the theory of structuration, Cambridge, UK (1984)
12. Orlikowski, W.J.: The duality of Technology: Rethinking the concept of technology in organizations. Organizational sciences 3(3), 398–427 (1992)
13. Aier, S., Weiss, S.: An institutional Framework for the Analyzing Organisational responses to the Establishment of Architectural Transformation. In: Proceedings of the European Conference on Information Systems (2002)
14. Jonkers, H., Lankhorst, M., Ter Doest, T., Arbab, F., Bosma, H., Wieringa, R.: Enterprise architecture: Management tool and blueprint for the organization. Information Systems Frontiers 8(2), 63–66 (2006)
15. Kaisler, S., Armour, F., Valivullah, M.: Enterprise Architecting: Critical Problems. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences (2005)
16. Chaung, C.J., Lobberenberg, J.: Challenges facing enterprise architects: a South African perspective. Paper Presented at the Proceeding of the 43rd Hawaii International Conference on System Science (2010)
17. Yin, R.K.: Case study research, design and methods. Sage Publications, California (2009)
18. Iyamu, T.: Institutionalisation of the enterprise architecture: The actor-network perspective. International Journal of Actor-Network Theory and Technological Innovation 3(1), 1–27 (2011)
19. Gray, D.E.: Doing Research in the real world. Sage Publication Ltd. (2009)

Towards a Flexible and Transparent Database Evolution

Rui Humberto Pereira[1] and J. Baltasar García Perez-Schofield[2]

[1] Instituto Politécnico do Porto, ISCAP, Portugal
rhp@iscap.ipp.pt
[2] Universidad de Vigo, Departamento de Informática, España
jbgarcia@uvigo.es

Abstract. Applications refactorings that imply the schema evolution are common activities in programming practices. Although modern object-oriented databases provide transparent schema evolution mechanisms, those refactorings continue to be time consuming tasks for programmers.

In this paper we address this problem with a novel approach based on aspect-oriented programming and orthogonal persistence paradigms, as well as our meta-model.

An overview of our framework is presented. This framework, a prototype based on that approach, provides applications with aspects of persistence and database evolution. It also provides a new pointcut/advice language that enables the modularization of the instance adaptation crosscutting concern of classes, which were subject to a schema evolution.

We also present an application that relies on our framework. This application was developed without any concern regarding persistence and database evolution. However, its data is recovered in each execution, as well as objects, in previous schema versions, remain available, transparently, by means of our framework.

Keywords: orthogonal persistence, schema evolution, aspect-oriented programming.

1 Introduction

Applications refactoring is a common activity in programming practices. These practices target the applications' maintenance in order to correct design or implementation mistakes, or, yet, prolong its life cycle introducing new functionalities. Many of them imply updates in the data model of persistent data, requiring a double intervention in the application source code and database. Modern object-oriented databases enable application's objects to be made persistent without a previous schema definition, alleviating that problem. The db4o (http://www.db4o.com), Versant (http://developer.versant.com) and ObjectDB (http://www.objectdb.com) systems provide transparent schema evolution mechanisms. In these systems, some types of updates in object's structure are propagated to database without the intervention of the programmer. However, in these

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_3,

systems, complex schema updates such as class movements in the inheritance hierarchy, field renaming and semantic changes in the content of the field, require the programmer's intervention in order to convert the data from the old structure to the new one. Many of these complex schema updates require helper conversion applications. Thus, the schema updates continue to be time-consuming tasks that programmers are faced with, compromising the programmers' productivity

The aspect-oriented programming (AOP) techniques aim at the modularization of applications' crosscutting concerns, such as logging, auditing and persistence. Earlier research [1][2] works have demonstrated that AOP also enables flexible and pluggable approaches to supporting database crosscutting concerns, such as schema evolution and instance adaptation. However, these works were not totally concerned about the transparency of the whole evolution process from the point of view of the programmer. We argue that the combination of orthogonal persistence and AOP paradigms provide good means to improve programmers productivity, as well as the ease of modification of the application, and therefore its lifetime.

In this paper we provide a short overview of our prototype, the AOF4OOP framework [3][4][5]. This framework supports orthogonal persistence and database evolution following an aspect-oriented paradigm. Applications that rely on our framework are practically oblivious regarding its persistence, as well as the database evolution when changes are applied to its schema. Programmers by means of our pointcut/advice expressions can introduce additional behaviour into the system in order to deal with database evolution. These aspects of persistence are introduced in applications by means of AOP techniques.

We also discuss a case study that demonstrate how our framework prototype can improve the programmer's work in terms of quality and productivity. It is based on a geographical application that, initially, was developed without any concern regarding persistence. In a second development phase, we provide this application with the aspect of orthogonal persistence applying minor changes to its source code. Finally, we evolved its data model in order to accommodate new functionalities. Using our pointcut/advice expressions, the database was adapted, the entire process being transparent to the application in its new database schema version, as well as to the old application, in the initial schema version.

In next section we briefly discuss our framework and pointcut/advice expressions. Section 3 presents a case study of an application, whose persistence aspects were modularized by means of our framework. In Section 4 we present some related studies. Finally, we present our conclusions and address some future work in our research.

2 Framework Overview

The persistence and database evolution are, respectively, aspects of applications and databases. Regarding persistence modularization, in earlier works [3][4][5], we have presented our meta-model and prototype. In these works we discuss how the orthogonal persistence aspect is modularized in applications that rely on our framework. Our meta-model supports database multi-version schemas by means of class versioning based on a class renaming strategy. Thus, several versions of the same application can share one logical schema database. Due to space restrictions, we cannot perform a complete discussion of our meta-model and the implementation of the framework.

Our framework[1] follows the orthogonal persistence paradigm[6]. The paradigm's principles advocate that objects must be handled in the same manner, despite of its state persistence (persistent or transient), and reachable transitively. We argue that these principles, due to data access transparency and orthogonality, provide means for applications to be *oblivious* [7] regarding persistence aspects, when applying AOP techniques. Furthermore, orthogonal persistence also enables application's schema to be incrementally [6] propagated to database. Our multi-version database schema approach, enabled by our meta-model, is able to accommodate each new class version[5]. Thus, applications that rely on our framework can be developed without major concerns for persistence, schema evolution and instance adaptation. In the next sections we discuss our join point model, as well as how our weaving process enables such *obliviousness* [7]. In Section 3 we present an application, whose persistence aspects are provided by our framework.

2.1 Aspects

In AOP paradigm, pointcuts *quantify* [7] a set of join points where the code in a base program should be affected. By applying pointcut expressions in these join points additional behaviour can be activated through an *advice* that could take place before, after or around the join point. A set of pointcuts and their corresponding advices are collected in a module called *aspect* [8]. In our framework, two kind of aspects were modularized: application and framework/database. The application's aspects were: (1) persistence and (2) data integrity. And, in the framework/database scope, the modularized aspects were: (1) schema evolution, (2) instance adaptation, (3) database data integrity, (4) object storing, (5) system statistics and (6) system debugging. These aspects were constructed by means of the AspectJ [9] programming language.

Regarding the instance adaptation, this is a crosscutting concern of the classes subject to a schema evolution. Due to the requirements posed by our meta-model, that require a persistent definition of the instance adaptation aspect, we developed a database weaver and a new XML based pointcut/advice language. Using this XML based language, programmers can write the instance adaptation aspect, which is weaved at runtime by our dynamic weaver. Conceptually, the default instance adaptation aspect, provided by framework's main module [5], is extended by means of these pointcut/advice expressions. Furthermore, this aspect extension is reified as a collection UBMO meta-objects [5].

[1] `http://www.iscap.ipp.pt/~rhp/aof4oop/`

2.2 Pointcut/Advice Expressions

We developed a new kind of pointcut/advice expressions that follows the *quantification* definition posited by Filman and Friedman[7].

> *"AOP is thus the desire to make programming statements of the form In programs P, whenever condition C arises, perform action A."*

The conditions to trigger the action are specified through the following matching parameters:

`matchClassName` - Class canonical name of the advised classes. Emulated objects whose classes match this parameter are advised. This parameter allows the * wildcard. If the class name is followed by `[]`, the advised object must be an array of this class.

`matchSuperClassName` - Super class canonical name of the advised classes. Matching through a super class is meant to reduce the number of user definitions. If many target classes that share a common super class exist, all of them are advised. This parameter allows the * wildcard.

`matchOldClassVersion` - The class version of the persistent object at database being emulated/converted. This parameter allows the * wildcard.

`matchCurrentClassVersion` - Defines the class version identifier of the running application. This parameter is required when many versions of the class already exist. In these cases, we must grant a correct target application version. This parameter allows the * wildcard.

`matchParentClassName` - This contextual parameter enables advising objects which at runtime are pointed as from a certain class. In our application (see Section 3) this parameter advises the `Coordinate` objects which are associated to a `Area` class. This parameter allows the * wildcard and `[]`.

`matchParentMember` - This is another contextual parameter that enables advising objects which at runtime are pointed as from a certain object member. In our application this parameter advises `Coordinate` objects whose reference is in the `bounds` member.

The action is defined in native programming language. Additionally, some parameters can complement the conversion process definition.

`applyDefault` - This parameter, when true, applies the default conversion behaviour, alleviating the programmers' effort in order to write this action (the body's advice). Note that, in many cases just some members of a class require user-defined code.

`outputClassName` - This parameter specifies the type of return value. This is very useful when we intend to apply an expression to several classes that share a common superclass. In these cases, the *advice* can return that superclass.

`conversionClassName` - This is an optional parameter that forces the name of the weaved class which implements the *advice*. It is useful just for debugging purposes. Thus, knowing that name, in case of any runtime error, the programmer can understand the localization of that error.

Figures 4 and 5 present examples of these expressions. They will be discussed in Section 3.4.

Our approach, in order to reify the instance adaptation aspect inside the database, is supported by an XML language. This approach presents two advantages: (1) Enables an easier edition of this aspect - Using a special graphical tool, as well as using a simple text editor. (2) The language extensibility provided by XML - New features can be added to our language without compromising the existing ones.

We also note that our approach does not require another programming language like Vegal [1]. Thus, the *advice* can be written in the same base program's programming language.

2.3 Conversion and Database Dynamic Weaving

The weaver is the central element of any aspect-oriented tool. According the pointcut/advice expressions content, it inserts new behaviour or replaces the one existing in the base program. Our database dynamic weaver takes this approach, extending the framework's default behaviour.

The framework's default instance adaptation behaviour already is capable of performing conversion, in many cases using our *direct mapping* algorithm. This mechanism supports all schema updates that can be autonomously inferred. In our prototype, the *direct mapping* deals with field changes for compatible types (e.g. `int` to `long`, `float` to `String`), movement of fields across the class inheritance structure, field removing and new fields initialized at zero or `null`. When semantic or structural updates occur in the schema, a *user-defined conversion* is required. This instance adaptation aspect, which is stored at the database as a collection of UBMO meta-objects, is weaved dynamically at runtime by our database weaver extending the default one. Inside the framework's main module, when required, the user-defined conversion code is called through the `doConversion()` system call. At runtime, our weaver compiles the *advice* source code (inside the `<Conversion/>` node) and loads a special class that implements that system call. The following example presents this function instantiated for a `Product` object in version "A" being converted to the actual one.

```
Product doConversion(Product$A oldObj,Category$A oldObjParent,Product newObj,
                     Category newObjParent)
{
     newObj.setWeight(2.2046*oldObj.getWeight());
     return newObj;
}
```

This function receives four arguments. Their names are reserved words, which are references to the old object (in version `"A"`), the object being converted (in the current schema version) and their parent objects (in the context of the conversion). The former provides data about the object in its old form, in version

`"A"`. This object pertains to that class version renamed with a suffix `$A`, having exactly the same structure (see [5]). The `newObj` provides a pre-initialized new object, or a pre-converted by means of the *direct mapping* mechanism. This option is defined in the parameter `applyDefault`. The other two provide non-local data, which improves the richness of the user-defined conversion functions. In this example, the products' weight data at the database (in version `"A"`) is represented as pounds. However, the running application expects this data in kilograms. This conversion function receives a `Product` object pre-converted through its `newObj` argument. Thus, just the `weight` member needs to be converted, alleviating the programmer's effort. In this example, the product's category information is not needed for this conversion. We highlight the expressiveness of conversion code that is enabled by our approach: (1) full programming language features, (2) access to non-local data and (3) application classes behaviour can be reused in these functions.

3 Geographical Application

In order to demonstrate the benefits of this framework, in terms of productivity and quality, we developed a real application. This proof of concept application uses data obtained from the online OpenStreetMap[2] geographical database. This online tool allows exporting a user-defined area through its coordinates. The OSM export files, in XML format[3], contains all data related with that geographical area. Each contains the coordinates of the boundaries and it is structured as a set of objects such as *Nodes*, *Ways*, *Relations* and *Tags*.

Regarding to hardware and software, application and framework were tested in a personal computer with 4GB of RAM running Linux OS and Java 1.6.

3.1 User Interface

Figure 1 depicts two screenshots of the application's user interface (UI): one presenting Madeira Island and another choosing an OSM file to be imported to the local database. This UI is organized as three panels: a menu on top, object browser on the left and maps on the right. The object browser enables the selective hiding of map objects such as points of interest and roads. The map and its objects are presented on the right side of the UI.

3.2 Data Model

The application's data model follows the same structure as the OSM files, which is presented in Figure 2. An `Area` object corresponds to an imported area through an OSM file. First the file is imported as an `OSMFile` object. Then, its contents are copied to an `Area` object. Just `Area` objects are made persistent. Thus, each

[2] http://www.openstreetmap.org
[3] http://wiki.openstreetmap.org/wiki/OSM_XML

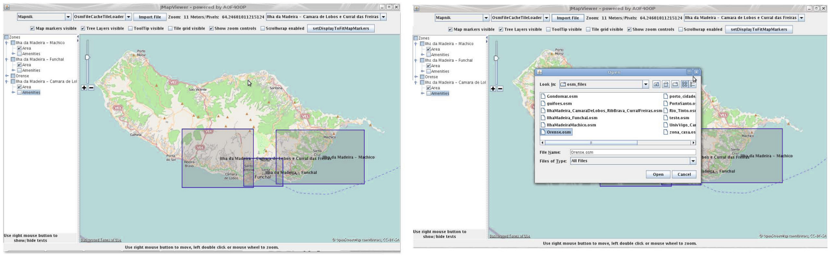


Fig. 1. User interface

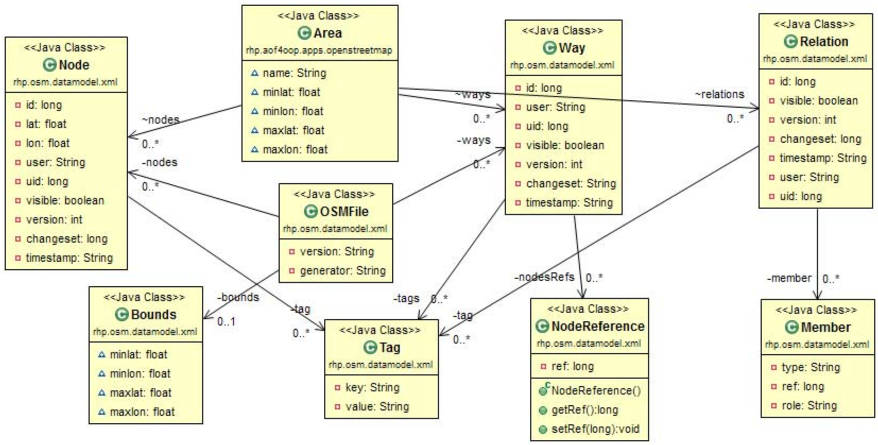


Fig. 2. Application's data model with class `Area` at version `"A"`

`Area` contains sets of `Node`, `Way` and `Relation` as arrays. Since a `Node` object can pertain to more than an `Area`, `Way` or `Relation`, they are shared, occupying just a unique slot in the database, and having a unique object identity.

3.3 Modularizing Data Persistence

At the initial phase of the development of the application, we just imported map areas to memory without any concern regarding data persistence. Map areas were imported from an OSM file and then put in a `Hashtable` collection. Thus, during the application's execution time all the imported data remained available, but is lost in any further application's execution after its restart. The following Java code listing illustrates that first stage of development.

```
Hashtable<String,Area> areas=new Hashtable<String,Area>();     /*Non persistent application*/
Area importedArea=importFile(name,file.getAbsolutePath());
areas.put(area.calcKey(),area);
```

The next listing shows how our framework alleviates the programmer's effort in order to provide applications with the persistence concern.

```
CPersistentRoot psRoot=new CPersistentRoot();             /*Persistent application*/
Area importedArea=importFile(name,file.getAbsolutePath());
psRoot.setRootObject(area.calcKey(),area);
```

We note the simplicity of the entire process. The framework's API provides a collection of persistent root objects following the same approach as the `Hashtable` collection. The `Area` objects, and its related objects, are made transitively persistent after being put in that persistent root. After that, any update in these objects, and any other that is reachable, are transparently reflected on the database.

3.4 Updating Application's Schema

In this section we discuss an update applied to application's data model. Before the update, each imported `Area` represents a rectangle in the map. After the update, the data model enables merging of areas forming a polygonal area. When the user imports an area in the map, it can be merged if it overlaps other existing areas. Thus, the class `Area` evolved from version `"A"` to `"B"`. Figures 2 and 3 depict the class `Area` in versions `"A"` and `"B"`, respectively. In version `"A"` class `Area` is a rectangle, just with a pair of coordinates, while in `"B"` it is an irregular polygon with several edges, requiring a collection of `Coordinate`.

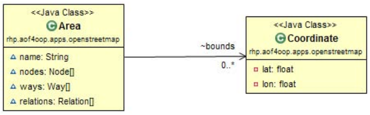


Fig. 3. Class `Area` in version `"B"`

Database Schema Evolution - As referred in Section 2, an incremental schema evolution is enabled [5] [6]. Thus, the entire process is transparent from the programmer's point of view. More information can be found in [5].

Database Instance Adaptation - The presented case, due to structural changes in class `Area`, requires user-definitions in order to enable the conversion of its instances in two directions: version `"A"` to `"B"` and `"B"` to `"A"`. Although the AOF4OOP framework enables bidirectional application compatibility, an `Area` defined as polygon in version `"B"` must be presented as a rectangle in version `"A"`. Thus, to enable that compatibility direction, the conversion procedure should produce a rectangle that includes the entire polygon or the major rectangle should fit inside the polygon. In version `"A"`, the `Area` edges coordinates are calculated from the polygon edges in version `"B"`. Figure 5 presents the required aspect definition.

In the opposite conversion direction, an array of coordinates that define the four edges of the polygon (rectangle) are required. In Figure 4 this interesting case is presented. The member `bounds` does not exist in version `"A"` of the `Area` class. The `matchParentMember` parameter enables the matching when this object member is accessed in version `"B"`. Inside the body's *advice*, the user-defined Java code produces the array of `Coordinate` using non-local information from `Area` object in version `"A"`.

```
<ubmo matchParentClassName="rhp.openstreetmap.apps.datamodel.Area"
    matchParentClassVersion="B"
    matchParentMember="bounds"
    matchClassName="rhp.openstreetmap.apps.datamodel.Coordinate[]">
  <convertion  applyDefault="false"  conversionClassName="ConvAreaCoord$A_to_B"
      outputClassName="rhp.openstreetmap.apps.datamodel.Coordinate[]">
      newObj[0]=new Coordinate(oldObjParent.getMaxlat(),oldObjParent.getMinlon());
      newObj[1]=new Coordinate(oldObjParent.getMaxlat(),oldObjParent.getMaxlon());
      newObj[2]=new Coordinate(oldObjParent.getMinlat(),oldObjParent.getMaxlon());
      newObj[3]=new Coordinate(oldObjParent.getMinlat(),oldObjParent.getMinlon());
      return newObj;
  </convertion>
</ubmo>
```

Fig. 4. Conversion from A to B

```
<ubmo matchOldClassVersion="B"
    matchCurrentClassVersion="A"
    matchClassName="rhp.openstreetmap.apps.datamodel.Area">
  <convertion applyDefault="true"
      conversionClassName="ConvArea$B_to_A"
      outputClassName="rhp.openstreetmap.apps.datamodel.Area">
      newObj.setMaxLat(oldObj.calcMaxLat());
      newObj.setMaxlon(oldObj.calcMaxLon());
      newObj.setMinlat(oldObj.calcMinLat());
      newObj.setMinlon(oldObj.calcMinlon());
      return newObj;
  </convertion>
</ubmo>
```

Fig. 5. Conversion from B to A

In both cases, as discussed in Section 2.3, at runtime, the user-defined conversion Java code is weaved as a `doConversion()` function. By means of our aspect-oriented approach, the right *advice* is triggered when a class needs to be converted. That takes place lazily only when objects are accessed. Then, that function is called at conversion-time, returning the converted object as applications expect. Thus, `Area` objects inside the database, in any version, are transparently accessed by applications regardless of the schema version.

4 Related Work

Rashid and Leidenfrost [1] in the AspOEv system, an evolution framework, made experiences in order to customize the evolution in the commercially available Jasmine object-oriented database. Their framework supports flexible adaptation and introduction of evolution mechanisms enabling a dynamic evolution of the schema and adaptation of existing instances, governing an object database by means of AOP techniques. The framework's aspects support such flexibility by capturing crosscutting hot spots (customization points in the framework) and establishing their causality relationships with the custom evolution approaches.

The AspOEv framework employs its own application programming language, Vejal [1][10], an object-oriented language with a versioned type system. Vejal has a two-level versioning identifier system: `C<1>` indicates version 1 of class `C` and `C<s=1>` implies the class version of `C` that occurs in schema version 1. In Vejal, one version of a class may be present in multiple schema versions.

Kuppuswami *et al.* [2] have also explored AOP techniques proposing a flexible instance adaptation approach. In this work, the authors have developed a system that supports instance adaptation with two *aspects*: update/backdate and selective lazy conversion *aspects*. The authors also highlight the flexibility provided by the AOP techniques to support database evolution

These earlier works proved that concerns encapsulation as *aspects* enables the easy replacement of the adaptation strategy and code, contrasting with other existing systems that introduce code directly into the class versions. They were inspiring for us. On the other hand, we argue that our approach provides programmers with transparency along with that flexibility, due to orthogonal persistence and our aspect-oriented language for database evolution. Since it does not requires the knowledge of a new programming language and enables a total modularization of database evolution.

5 Conclusions and Future Work

We presented a short overview of our framework. Applications that rely on this framework can be developed without major concerns for persistence aspects. Applying AOP techniques, the framework provides applications with that additional behaviour. Regarding database evolution, the framework, due to its orthogonality and multi-version schema approach, also enables a transparent evolution of schema and object instances. Such transparency is achieved by means of the orthogonal persistence principles and AOP techniques.

We also discussed our aspect-oriented language that extends the framework's default instance adaption aspect. This language captures the application crosscutting hot spots and establishes a relationship with user-defined conversion functions. Inside these functions, objects of a class in distinct versions can be handled through the same programming language used in the program's base.

The presented geographical application served as proof of concept, showing the gains in terms of productivity for programmers' work. Regarding this case study, we argue that our framework overcomes much of the existing limitations found in modern systems, in terms of transparency and flexibility.

Despite the benefits for programmers, our framework just provides a single and simple transaction model. Furthermore, classes that belong to the Java packages cannot be persistent due to JVM restrictions. These two issues are the subject of our current research.

References

1. Rashid, A., Leidenfrost, N.A.: Supporting flexible object database evolution with aspects. In: Karsai, G., Visser, E. (eds.) GPCE 2004. LNCS, vol. 3286, pp. 75–94. Springer, Heidelberg (2004)
2. Kusspuswami, S., Palanivel, K., Amouda, V.: Applying aspect-oriented approach for instance adaptation for object-oriented databases. In: Proceedings of the 15th International Conference on Advanced Computing and Communications, pp. 35–40. IEEE Computer Society, Washington, DC (2007)
3. Pereira, R.H., Perez-Schofield, J.: An aspect-oriented framework for orthogonal persistence. In: 2010 5th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6 (2010)
4. Pereira, R.H., Perez-Schofield, J.B.G.: Orthogonal persistence in java supported by aspect-oriented programming and reflection. In: 2011 6th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6 (June 2011)
5. Pereira, R.H., Perez-Schofield, J.B.G.: Database evolution on an orthogonal persistent programming system - a semi-transparent approach. In: 2012 7th Iberian Conference Information Systems and Technologies (CISTI), pp. 1–6 (June 2012)

6. Atkinson, M., Morrison, R.: Orthogonally persistent object systems. The VLDB Journal 4(3), 319–402 (1995)
7. Filman, R.E., Friedman, D.P.: Aspect-oriented programming is quantification and obliviousness. Technical report (2000)
8. Kiczales, G., Lamping, J., Mendhekar, A., Maeda, C., Lopes, C., Loingtier, J.M., Irwin, J.: Aspect-oriented programming. In: Akşit, M., Matsuoka, S. (eds.) ECOOP 1997. LNCS, vol. 1241, pp. 220–242. Springer, Heidelberg (1997)
9. Kiczales, G., Hilsdale, E., Hugunin, J., Kersten, M., Palm, J., Griswold, W.G.: An overview of aspectj. In: Lindskov Knudsen, J. (ed.) ECOOP 2001. LNCS, vol. 2072, pp. 327–353. Springer, Heidelberg (2001)
10. Rashid, A., Leidenfrost, N.A.: Vejal: An aspect language for versioned type evolution in object databases. In: Workshop on Linking Aspect Technology and Evolution (Held in Conjunction with AOSD) (2006)

Virtual Desktop Infrastructure (VDI) Technology: *FI4VDI Project*

C. Redondo Gil[1,2,3,4], P. Vega Prieto[1,4], M. Silva[3,4], and A.M. Teixeira[3,4]

[1] Castile and Leon Technological Center for Supercomputing (FCSCL), León, Spain
{carlos.redondo,paula.vega}@fcsc.es
[2] Innovative Business Association for Network Security and Information Systems, León, Spain
[3] Inova-Ria - Associação de Empresas para uma Rede de Inovação em Aveiro, Aveiro, Portugal
{msilva,ateixeira}@inova-ria.pt
[4] FI4VDI Project,
Inova-Ria - Associação de Empresas para uma Rede de Inovação em Aveiro, Aveiro, Portugal
http://fi4vdi-sudoe.eu

Abstract. This paper presents an analysis of the FI4VDI project, the goal of which is to develop an innovative model of service provision using cloud computing in order to create an infrastructure suite aims at large companies and SMEs alike, educational institutions, universities and/or research centres that would contribute to competitiveness and provide an efficient solution to reducing the carbon footprint derived from the use of information technology.

Keywords: VDI, Smart Data Centre, SMART IT Infrastructure, Energy Efficiency, Carbon Footprint.

1 Introduction

The aim of the *FI4VDI* project is: *to develop a federated infrastructure network for the creation of virtual desktop services*. Firstly, by evaluating the position and perception of public and private organisations in the SUDOE Space as regards the desirability of the virtualising IT operating environments, and secondly, by promoting the spread of cloud computing as a means to achieve savings, efficiency and simplicity with the primary objective of ensuring improved productivity.

The provision of cloud computing services by supercomputing centres has a positive effect on the ecological footprint; dependence on physical ICT infrastructures is reduced when these are replaced by virtual ones, and this in turn produces a marked reduction in energy consumption in these institutions.

With a federated cloud computing model, desktops can be connected to dedicated servers with high rates of effective utilisation, greatly increasing energy efficiency and reducing the carbon footprint associated with the service.

The goal of the project is to develop a federated infrastructure network for the creation and management of energy efficient ICT services.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_4,

Table 1. Organizations involved in the project FI4VDI

▲ **Organisations that participated actively as infrastructure providers included**
The Supercomputing Centre in Castile and Leon - Castile and Leon Region ES41 (Spain)
The Computing and Advanced Technologies Centre in Extremadura - Extremadura Region ES43 (Spain)
The University of Lerida Faculty of Arts Computer Centre in Ponent - Catalonia Region ES51 (Spain)
The University of Montpellier 2 Sciences et Techniques - Languedoc-Roussillon Region FR81 (France)

● **Organisations that participated actively as business associations included**
The Innovative Business Association for Network Security and Information Systems (Spain)
Association of Companies for an Innovation Network in Aveiro - Central Region PT16 (Portugal)
The Science and Technology Park Agri-Food Consortium in Lerida - Catalonia Region ES51 (Spain)

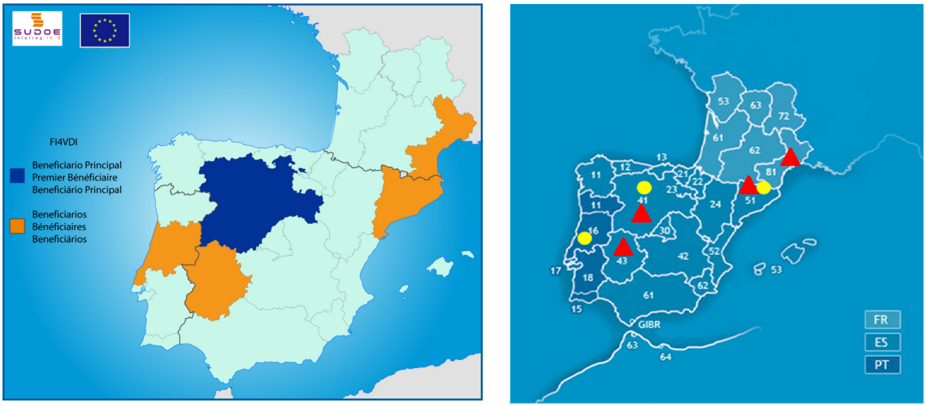


Fig. 1. Scope of the FI4VDI project in the SUDOE Space http://www.sudoe.eu Setting of the FI4VDI project: profile of the partners participating in the project.

2 FI4VDI-SUDOE Project

2.1 Objectives

The aim of the FI4VDI project was to develop a federated infrastructure network for the generation of virtual desktop services, and to promote sustainable development by leveraging the benefits deriving from transnational cooperation.

In brief, the project proposed the creation of a private cloud infrastructure using the resources available in various supercomputing centres located in different SUDOE regions, with the goal of ensuring protection of users' data and compliance with regulations pertaining to information security and the service-level agreements established. Implementation of this service would entail improved competitiveness and cost savings in the sectors targeted, where energy savings and efficiency are a distinguishing feature.

The problem addressed by the project was the need to determine the position and perception of public and private entities located in the SUDOE Space as regards the desirability of virtualisation of IT operating environments, and to promote the spread of cloud computing as a means to achieve savings, efficiency and simplicity with the primary objective of ensuring improved productivity.

The origin of the project: The provision of cloud computing services by supercomputing centres has a positive effect on ecological footprints; dependence on physical ICT infrastructures is reduced when these are replaced by virtual ones, and this in turn produces a marked reduction in energy consumption in these institutions.

Project objectives and results: With a federated cloud computing model, desktops can be connected to dedicated servers with high rates of effective utilisation, greatly increasing energy efficiency and reducing the carbon footprint associated with the service.

Project strategy and structure:

Task group 0. Preparation

Task group 1. Project coordination and management

Task group 2. Technical infrastructure development

Task group 3. Adapting applications to the cloud environment

Task group 4. Integration. Prototypes.

Task group 5. Project monitoring and evaluation

Task group 6. Publicity and information. Market capitalisation

2.2 Virtual Desktop Infrastructure - VDI as a Technology Proposal [1] [2]

The term *desktop virtualization* was introduced in the 1990s to describe the process of separation between the desktop, which encompasses the data and programmes that users employ for their work, and the physical machine. A "virtual" desktop is stored remotely on a central server rather than on the hard disk of an individual personal computer. This means that when users work on their desktop from their laptop or PC, all their programmes, applications, processes and data are stored and run centrally, allowing users to access their desktops from any device that can connect remotely to the central server, such as laptops, PCs, smartphones, or thin clients.

Through desktop virtualisation, the entire environment of an information system or the environment itself is encapsulated and delivered to a remote device. This device can be based on a completely different hardware architecture from the one used by the projected desktop environment. It can also be based on a completely different operating system.

Desktop virtualisation consists of the use of virtual machines to enable multiple network users to maintain their individual desktops on a single server or host computer. The central computer may be located in a residence, in the company or in a data centre. Users may be geographically dispersed and connected to the central computer via a local area network (LAN), wide area network (WAN) or via the Internet.

Desktop virtualisation offers advantages over the traditional model, in which each computer functions as a complete and independent unit with its own operating system, peripherals and applications. Energy costs are reduced because resources can be shared and allocated to users according to their needs, and the integrity of user information is enhanced because the data centre stores and safeguards all data and backups. Furthermore, software conflicts are minimised by reducing the total number of programmes stored on computers, and although the resources are distributed, all users can personalise and customise their desktops to meet their specific needs. Thus, desktop virtualisation provides greater flexibility than the client/server paradigm.

The limitations of desktop virtualisation include the possibility of security risks if the network is not well managed, the loss of user autonomy and privacy, the challenges involved in creating and maintaining drivers for printers and other peripherals, difficulties in managing complex multimedia applications and problems in maintaining the addresses of virtual machine users consistent with those held by the data centre.

Table 2. Benefits and advantages of VDI technology

Benefits	Advantages
Like any other technology, desktop virtualisation provides a number of key benefits that render this technology the first choice for a large number of users: • Enhanced security and reduced desktop support costs [3] • Reduced general hardware costs [4] • Ensured business continuity [5] • An eco-friendly alternative [6] • Improved data security	The main advantages are [7] as follows: • Instant implementation of new desktops and use of applications • Virtually zero downtime in the case of hardware failure • Significant reduction in the cost of new deployments • Sound capacity for managing the desktop image • The PC replacement cycle is extended from 2-3 years to 5-6 years or more • Existing desktops include multiple monitors, bidirectional audio/video, video streaming, USB port supports, etc. • Company employees can access their virtual desktops from any PC, including a PC in the employee's home • Resources tailored to desktop needs • Multiple desktops on demand • Free provision of desktop computers (controlled by the policies of each corporation)

2.3 Description of the Technical Task Groups of FI4VDI Project

Technical Infrastructure Development

Design of a federated cloud infrastructure capable of providing the selected applications.
Actions focused on definition, implementation and deployment of the system architecture, considering the hardware and software of the different cloud servers as well as the most suitable middleware to interrelate it all.

Adapting Applications to the Cloud Environment

Selection of different applications and environments.
Implementation of a federated infrastructure, optimising the resources and efficiency of the processes involved.
Adaptation focused not solely on computing implementation and functionality, but also on modelling functional paradigms and services that respond to the growing needs of these environments.

Integration. Prototypes. Each of these involved the following actions: Prototype design - Integration of prototypes into the federated infrastructure - Functional implementation - Validation - Battery of functional tests - Stress tests - User training Dissemination and value enhancement

2.4 Adoption of Cloud Computing

Table 3. Adoption of cloud computing: software solutions

Software solutions (not necessarily in the cloud)		Rate of use (%)
Storage	Technological resources offered by providers, the function of which is to store client data in their databases	75.5
Email	A computer programme that allows clients to manage electronic mail (writing, sending, receiving, storing, organising, etc.)	74.1
Computing	Computing resources offered by providers (computing capacity on provider's servers)	65.4
Backup	Technological resources that allow temporary storage of client information, for recovery in the event of loss	59.8
Office applications	Collection of applications that enable clients to create, modify, organise, scan, print, etc., files and documents	46.4
Virtual desktops	Technology that enables clients to work on a computer using their desktop, from a terminal located elsewhere	40.4
Database management systems (DBMS)	Systems used for storage and organisation of public authority data	38.1
Collaborative tools	Systems that provide access to certain services that enable clients to communicate and work together whether or not they are located in the same physical location	34.2
Content Creation and Control (CCC)	Computer systems that enable the creation and management of content (text, images, video, etc.) and provide the option of sharing this content among members of a team	33.6
Managed File Transfer (MFT)	Software that facilitates the secure transfer of data from one computer to another over a network	32.1
Customer Relationship Management (CRM):	Software that facilitates the secure transfer of data from one computer to another over a network	28.3
Enterprise Resource Planning (ERP)	Set of tools which provide integrated management of processes and information corresponding to different business areas	25.5
Application Lifecycle Management (ALM)	Process of managing the lifecycle of a software application through governance, development and maintenance of the same	18.2
Project Portfolio Management (PPM)	A process which incorporates analysis and collective management of a group of ongoing, envisaged or imposed projects, and Supply	
Chain Management (SCM)	Application of tools for improving and automating supply by controlling stock, delivery times and other aspects	7.6

Base: Public authority users of cloud computing (n=152)[1].

[1] *Study on cloud computing in the public sector in Spain* Spanish National Institute of Communication Technologies (INTECO) July 2012. Study funded as part of the SERPLAGO project (cloud platform services for e-government and e-administration processes), funded by the 2011 INNPACTO sub-programme of the 2008-2011 National Plan of the Spanish Ministry of Economy and Competitiveness, and co-financed by the EU ERDF programme.

2.5 Indicators of FI4VDI Project

Table 4. Indicators of Performance, Outcome and impact of FI4VDI project

Performance indicators	Outcome indicators	Impact indicators
R & D projects with added value from an environmental perspective (num.) Companies and SMEs that have been part of innovation partnerships funded (num.) Projects of telecommunications networks and for encouraging the application of ICTs (num.) Deployment Type: Private/Public/Hybrid/Community Participation in Conferences, Workshops and Working day diffusion	New technologies developed (num.) Transnational cooperation in innovative networks created Tools (applications and services) for technological transfer between companies and technological centers and SMEs adopted in SUDOE countries / regions (num.) Companies and SMEs that have benefited from project results conducted (num.) SUDOE area with improved access to ICTs (Km2) Publications resulting from the project	New technologies transferred to enterprises, SMEs and / or managing entities (num.) Permanent networks of cooperation established (num.) Agents (institutions, enterprises, SMEs, etc.) connected to telecommunication networks created (num.) Rate of adoption of Cloud computing model Rate of knowledge of Cloud computing and virtualization technology

2.6 Results

Creation of a Platform as a Service (PaaS) for mass deployment of virtual desktops.

Federation of the participant supercomputing centres infrastructures.

Creation of an innovative cloud computing service aimed at users in the public and private sectors.

Enhanced competitiveness and cost savings in the sectors targeted by the service created.

Establishment of strategic recommendations: definition of reliable models for cloud computing (service levels, system capacity, system restoration, interoperability through shared service infrastructures, migration models), identification of service areas and evaluate and promotion of cloud computing as a tool for achieving cost savings and technological optimisation.

Establishment of technological recommendations: identification of existing solutions and possibilities, assessment of the real capacity of suppliers, selection of an applications and systems map, initiation of a cloud computing strategy by adopting private clouds and infrastructure and platform services, establishment of system migration plans and identification of cloud computing as a model that fosters other technologies which are either emerging or in the process of expansion, such as energy sustainability in the area of IT or open source solutions.

Establishment of management recommendations: definition of systems for assessing investment returns, analysis of organisational impact and identification of change management models, development of new contracting models and practices, standardisation and organisation of common services and definition of risk analysis models.

3 Conclusions

The spread of cloud computing in the Spanish public sector is still limited, and is more common among local authorities than among regional and state institutions.

In general, the most common mode of deployment is private.

Savings, efficiency and simplicity are the reasons that have prompted public authority institutions to contract cloud computing services.

Those public institutions that decided to adopt cloud computing did so after carrying out a legal analysis focused primarily on data protection legislation.

According to the public authority bodies that have adopted cloud computing, the principle benefits of this model are the savings in costs and time it represents, whilst integrity of services and data was identified as the main difficulty.

The Spanish public sector perceived the cloud as a technological and, above all, operational advantage which has met their initial expectations regarding cloud computing.

Among the public authority bodies already using cloud computing, the future prospects are very bright: they intended to continue working in the cloud, they would recommend this technology to other institutions and they expected to continue to obtain future benefits from using the cloud.

However, those public institutions that are not yet using the cloud were more wary: few intended to incorporate technological solutions, and only a minority of these would consider cloud computing.

Acknowledgments. Project FI4VDI was funded by the INTERREG IV B SUDOE *Territorial Cooperation Programme*.

Ref. SOE4/P3/E804-FI4VDI

Project FI4VDI-SUDOE, Corporate Logo

References

1. Introduction to Desktop Virtualization, `http://www.escritorios-virtuales.net`
2. Technical definition desktop virtualization, `http://www.escritorios-virtuales.net`

3. Permanent availability of the applications, http://www.escritorios-virtuales.net
4. How to save and perform at their best thanks to remote access, http://www.escritorios-virtuales.net
5. Virtualization Management, one of the fastest growing areas in 2009, http://www.techweek.es
6. virtualization, efficiency and environment, http://www.escritorios-virtuales.net
7. Principal Benefits of Desktop Virtualization, http://www.escritorios-virtuales.net

Exploring WebRTC Technology for Enhanced Real-Time Services

Vasco Amaral[1,2], Solange Rito Lima[1], Telma Mota[2], and Paulo Chainho[2]

[1] Centro Algoritmi, Department of Informatics, University of Minho, Portugal
[2] Portugal Telecom Inovação, SA, 3810-106 Aveiro, Portugal

Abstract. WebRTC is a standard technology which allows real-time communications between browsers, without installing additional plug-ins. In this way, for each device (computers, smart phones, etc.) with an installed browser, it is possible to perform peer-to-peer real-time communications natively, for instance, video and voice calls, chatting or instant messaging, file sharing and screen sharing. This recent technology has grown exponentially both in implemented solutions and in browsers compatibility. WebRTC is therefore an evolutionary technology with a strong growth, where more solutions Over-The-Top (OTT) could appear and where the telecommunications operators could invest creating their own service solutions. Facing the lack of standards regarding the communication between WebRTC endpoints, this paper studies in depth the WebRTC technology in order to identify its potential and to assess in which way it could impact on the telecommunications world. This paper also presents a framework that helps developing WebRTC applications and services at a higher level. As proof-of-concept a WebRTC client is developed to allow testing the services implemented in the framework.

1 Introduction

Telecommunications are present in our daily routine. Services such as audio calls, video calls and instant messaging are nowadays cheaper services indispensable to support our personal life and business relations. The Web has also grown exponentially and turned into an indispensable resource where a large amount of information flows every day. Single-Page Applications (SPAs) [1] started to be created to allow the users to redesign one part of the user interface without requesting the server, making possible to perform load balancing between the user and the server. With this evolution new technologies started to emerge allowing to create innovative service solutions, which is the case of WebRTC technology. This technology starts a new world in the developing of web solutions.

WebRTC is a recent and growing technology which aims to bring important challenges and opportunities to the Web business world, regarding service offering. This motivates a deep study of this technology potential and the implementation of new solutions that ease the development of added-value services (e.g., presence, contact list, audio and video calls, etc.). This is the main focus of this paper, which presents an encompassing WebRTC study carried out to determine

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2,* Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_5,

the maturity of this technology and its evolution. By reviewing and articulating the involved standards, the existing solutions in the topic, and by identifying the open issues in the communication between WebRTC endpoints, we expect to foster the development of new WebRTC services. Additionally, the paper discusses the main aspects involved in the development of a new framework that will allow users and external entities to develop new WebRTC solutions at a higher and easier level. As proof-of-concept, a WebRTC client is developed resorting to this framework, allowing to test the involved services.

This paper is organized as follows: Section 2 presents the concepts associated to the WebRTC technology, including its architecture and Application Programming Interfaces (APIs); Section 3 discusses the related work on WebRTC, such as the existing solutions in terms of technology and signalling; Section 4 presents the developed framework and the architectural solution to support the defined services; Section 5 include the principal conclusions of the developed work.

2 Concepts

This section presents the main concepts inherent to the WebRTC technology, such as its architecture and APIs, and describes the WebSocket protocol.

2.1 WebRTC Architecture

The WebRTC architecture is split in two distinct levels: WebRTC C++ API and Web API, as shown in Figure 1. The WebRTC C++ API is normally used by Web Browser programmers, to allow the creation of Web Browsers that support WebRTC features [2]. For the developers of Web applications, the use of the Web API is very important, as it allows to take advantage of the WebRTC features supported by the Web Browsers, using the Javascript programming language [2]. The WebRTC API is developed in three main concepts, *PeerConnection*, *MediaStreams* and *DataChannels* [2], discussed below.

PeerConnection. *PeerConnection* allows the direct communication between users (peer-to-peer communication), browser-to-browser. To establish a connection and have a signalling negotiation, it is necessary a signalling channel. This is provided through a script implemented in a Web server, using WebSockets or XMLHttpRequest. It uses the Interactive Connectivity Establishment (ICE) Protocol with Session Traversal Utilities for NAT (STUN) and Traversal Using Relays around NAT (TURN) to allow media streams to cross through Network Address Translation (NAT) and firewalls [3], [4], [5].

MediaStreams. *MediaStream* is an abstract representation of an audio and video data stream. This type of applications could be used to show, save and send its content to the remote peer. Two types of streams are available: Local MediaStream and Remote MediaStream. LocalStream is a captured stream

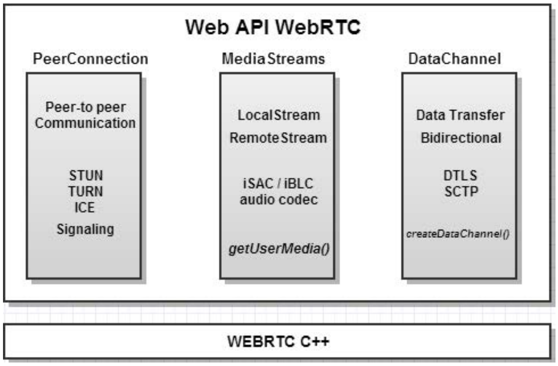


Fig. 1. WebRTC Architecture

in the same system (webcam and microphone) and Remote MediaStream is a stream that is received from the remote peer. To access the media of the terminal components it is necessary to execute the *getUserMedia()* function [3], [4], [5].

DataChannel. *DataChannel* is a bidirectional data channel in peer-to-peer connections. It allows to transfer non media data resorting to another protocols such as Stream Control Transmission Protocol (SCTP) encapsulated in Datagram Transport Layer Security (DTLS). In this way, it is possible to have a solution for NAT with confidentiality, source authentication and data integrity, for the data requiring transmission. For the creation of one data channel it is necessary to execute the *CreateDataChannel()* function in one PeerConnection's instance [3], [4], [5].

2.2 WebSockets

Websocket is a protocol which provides a bidirectional communication based on sockets, and uses Hypertext Transfer Protocol (HTTP) as transport layer. It is possible to use Websockets with existing web solutions, working over HTTP ports such as 80 and 443, which facilitates the integration of this protocol with existing solutions. Besides, this protocol is not limited to HTTP protocol, allowing a simpler handshake over a dedicated port, without reinventing the entire protocol [6], [7].

2.3 Signalling

To establish a session between users, it is necessary to have a negotiation of session parameters, described via Session Description Protocol (SDP). JavaScript Session Establishment Protocol (JSEP) defines how the Javascript applications interact with Real Time Communication (RTC) functions. This protocol is responsible for defining how clients retract the information about the type of supported media and the codecs to use, which are defined in a SDP RTCSessionDescription object. JSEP provides two mechanisms to create *offers* and *answers*, defining how they could be applied in a session [8]. This protocol does not define how the information is exchanged between users, which needs to be determined at application layer.

3 Analysed Technologies

For developing this project it was necessary to research and analyse the technologies that could be useful to develop a solution with WebRTC services. WebSockets is a protocol that is not implemented in every browser. Some platforms allow the development of WebSockets solutions in a server, being *Node.js* and *Vertx.io* the most relevant for this work.

3.1 *Node.js* and *Vertx.io*

Node.js is a server side event-oriented platform, which allows the creation of scalable web applications. The applications developed with this platform are written in JavaScript and are asynchronous, allowing to decrease overhead and increase scalability [9]. Node.js contains a HTTP server library to run web applications without using traditional HTTP servers (e.g., Apache).

Vertx.io is a next generation platform, which runs over a Java Virtual Machine (JVM), is asynchronous and event-oriented. It implements an asynchronous concurrency module, which allows the creation of scalable applications. The programming languages that can be used are: Java, Javascript, Groovy, Ruby, Python and Scala. *Vertx.io* is a plataform, which allows the creation of server and client side applications. It permits to organise the code by *Modules*, where each module can execute a specific functionality. It implements an *Eventbus* where modules and clients can be registered, allowing the communication between them [11], [10].

3.2 *WebRTC-Experiment and RCMultiConnection.js*

WebRTC-Experiment is a Javascript library repository that implements WebRTC demos, aiming at simplifying the development of this type of applications [12]. One of these libraries is called RTCMultiConnection.js [14]. With this library it is possible to develop advanced services in a simpler way, namely [15]: audio/video conferences; sharing data and files; sharing screen; renegotiate multiple stream; remove individual streams; mute and unmute streams; ban users; detect presence (detect if a user enters or leaves). It is possible to implement a solution using this library with socket.io or WebSockets [13].

4 WebRTC Framework

This section presents the main decisions and architectural issues involved in the development of the WebRTC framework. The main goal of this framework is to give to users and external entities an easier way to create services and WebRTC applications at a higher level.

4.1 Architecture

To design the architecture for the framework it is necessary to define the target services to be developed. This aspect has been decided as follows:

– audio and video calls/conferences;
– presence and contacts list;
– chat and instant messaging;
– files sharing;
– audio and video recording.

Vert.x allows to divide the code into modules, enabling the development of a modular architecture where each module has a distinct function, splitting the service's functionality.

Figure 2 illustrates the devised framework architecture and how the client interacts with the server. As illustrated, there exists a client and a server side. The defined modules are registered in the server side. In this way, the client can communicate with each module, by the Eventbus, to execute the necessary functions. The functionality of each module is addressed below.

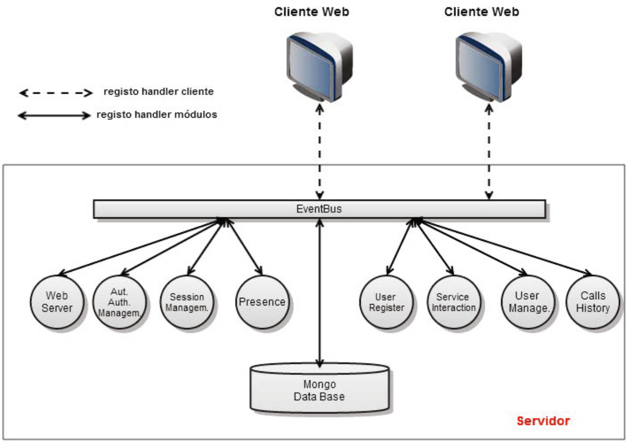


Fig. 2. Framework Architecture

Web Server Module. The Web Server module allows the creation of a web server. This server will provide all the necessary services to users so that they are able to use all the application features. It provides a complete configuration environment, being possible to configure fields such as host, port and address.

Mongo Data Base Module. Mongo DB, a non relational database, has been selected to support the database functionality. This selection intends to take advantage of a module provided by the Vert.x community that implements a simpler and dynamic way of managing the database. This module allows executing the following operations:

- ***save***: saves documents in the existing collections;
- ***update***: updates the information relative to a document and to a collection;
- ***find***: finds one or more documents in a collection;
- ***find one***: finds one document in a collection;
- ***count***: counts the number of documents in a collection;
- ***delete***: deletes / add the associated documents.

All these operations are only executed by the modules in the server, i.e., every activity executed by clients is first sent to modules, such as: Presence, Notification, Calls History, etc. All these modules are responsible for updating all the information in the database.

Presence Module. The presence module implements the presence service and the contact list. This module is responsible for adding and removing persons from a contact list and for updating the presence state of each person (Online, Offline, Not Here, Busy). Each client, besides the user profile (where it is possible to find information such as: username, e-mail, password and call history), also has a user profile group. This group allows the user to keep the information related to the subscription requests of other users. These requests are stored in a *presence_authorization_pendent* list, containing the subscription requests that are still not answered. The rejected requests are stored in a *denied_presence* list and the accepted ones are stored in the *authorised_presence* list. This latter list stores the user's contacts (see Figure 3).

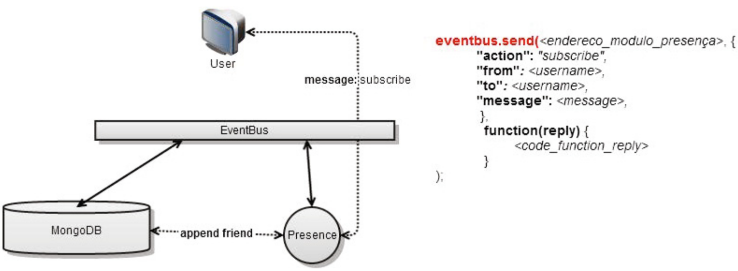


Fig. 3. Add a client to the contact list

When a client executes the system authentication, a *find* message will be sent to the presence module that is in charge of searching and loading all the contacts for that user, including the presence state of each one. Besides, it is sent another message for the same module, for updating the information of the user session. Each user has a registered address for each person in his contact list, i.e., each client has a specific registered address (*presence.* $< contact\_name >$) for every contact in his list. Every time a client update his presence state, a *publish* message will be sent to all of his subscribers, as shown in Figure 4. In this way, it is possible to keep the information about presence state always updated.

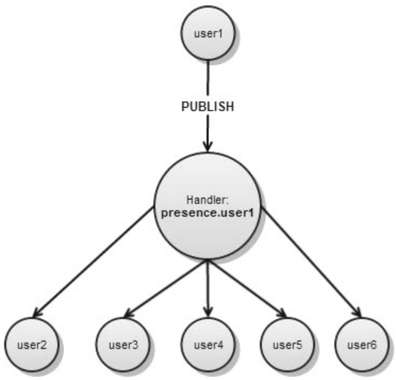


Fig. 4. Change state notification

To add a person to a contact list, a *subscribe* message is sent to this module, and the client will be notified that someone wants to add him to its contact list. From the moment the client accepts the invitation, this module will be responsible for updating the information from the two involved users, putting them in the *contactauthorised_presence* list of each other. If the request is rejected, both will be posted in the *denied_presence* list, notifying the client of that fact.

This module allows developers to use simple functions such as *updateState*($< state >$), *addFriend*($< user >$) and *removeFriend*($< user >$). Using this approach they can create applications in a simple way, having the core of the framework executing the corresponding functions.

Notification Module. The notification module deals with the notification messages related to calls. Every time a user is invited to join a call / conference session a notification is sent to that user, which can accept or reject the session. Note that several notifications are associated and handled by other modules. The notifications from the presence module are an example, for instance, when a user is added to a contact list. The same occurs when a client changes his state presence, which is handled by the presence module.

With this module it is possible to use functions such as *inviteFriends*($< friends >, < room >, < type\_of\_session >$), *acceptInvite*($< username >$)

and *denyInvite(< username >)*. These functions are related with some event handlers, allowing to know who invited, who accepted and who rejected the invitation.

Calls History Module. The call history module stores the information related to performed calls / conferences. Every time a call/conference session is started, the module stores the starting hour, date and who started the call. If someone joins the call / conference, the information is always updated, in this way, it is possible to know who were involved in that session. When a session ends, a *registcall* message is sent to this module to inform that the session can be stored in the database. Every time a user authenticates himself in the system it is possible to visualize the calls history, detailing every call made and received. The information stored in the database is the following:

– ***initiator*** - person who started the call/conference;
– ***hour*** - initial hour of a call/conference;
– ***date*** - call and conference date;
– ***duration*** - call/conference duration;
– ***participants*** - clients that have participated in a call/conference.

Figure 5 illustrates the interaction of this module with the architecture.

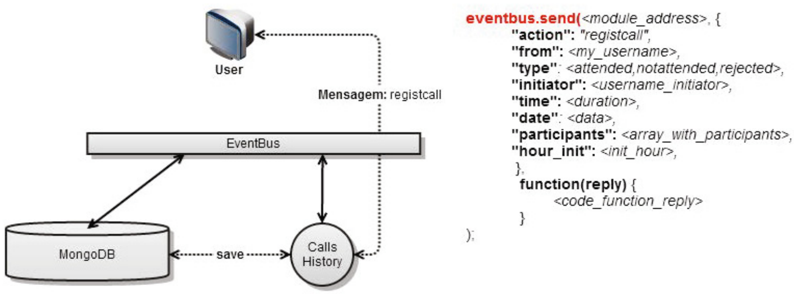


Fig. 5. Register Call

With this module it is possible to use functions such as *registCall(< type >, < message >, < username >)*, *unregistCall(< id >, < message >, < username >)* and *listCalls(< username >)*. These functions allow to register, unregister and list calls in the database.

Session Management Module. The session management module allows a complete management of each session. There are two types of sessions: users sessions and calls sessions. Users sessions are initiated every time a client authenticates himself in the system. The data stored include: username, state, session identifier and the correspondent services addresses. When a client leaves the system, his associated sessions are deleted, as shown in Figure 6.

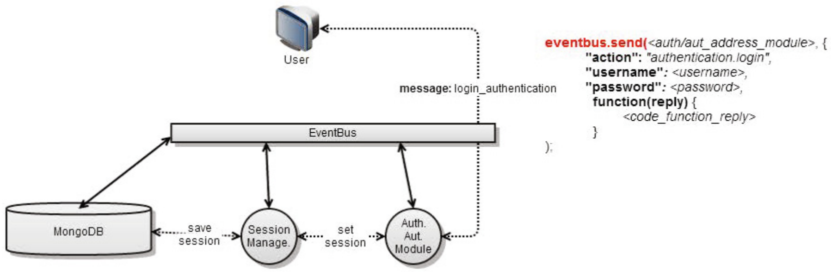


Fig. 6. Session Register

4.2 Client Application

As mentioned before, a client application was developed with the goal of demonstrating the framework functionality. This application was developed in HTML5, Javascript and WebSockets, which allows the development of single page applications with the WebRTC Web API. The use cases defined and implemented to provide the proof-of-concept include: audio and video calls / conferences; presence and contact list; chat and instating messages; file sharing; screen sharing and calls history. The signalling solution resorts to socket.io technology, included in the *Node.js* technology, which allows the creation of namespaces between peers, guaranteeing that the signalling negotiation is only made between two peers. Services such as presence, contact list and calls history, use Vert.x to exchange messages between peers and the server. Those modules interact more frequently with the server, so it is necessary to maintain all the information updated in runtime. The development of this application explores the advantages of the WebRTC-Experiment technology integrated with *Vert.x* implementation. The framework validation process was continuous, verifying and validating each module functionality, in order to proceed progressively to the development of other modules. In a final validation, all the implemented functionalities of the client application where successfully verified.

5 Conclusions and Future Work

WebRTC technology and its community are growing steadily, with a continuous increase in the number of applications and libraries. Web browsers are also evolving and the support of WebRTC APIs is now more solid and mature.

This paper is focused on clarifying and articulating the main concepts and technologies involved in the design of WebRTC solutions in order to foster the development of enhanced real-time services. As a contribution in this direction, a framework has been developed allowing the creation of services at a higher level. Using this framework it is possible to develop easily services such as: presence, contact lists, audio and video calls / conferences, chat, file sharing and screen sharing. As proof-of-concept a client supporting all these services has been implemented using Vert.x, WebRTC technologies and WebSockets protocol.

As future work, this framework will be enriched with new services such as Voice Activity Detection (VAD) and an interactive board, and improved regarding interoperability between browsers.

Acknowledgments. This work has been supported by FCT – Fundação para a Ciência e Tecnologia in the scope of the project: PEst-OE/EEI/UI0319/2014.

References

1. Single page apps in depth (2012), http://singlepageappbook.com/
2. WebRTC: General Overview (2011), http://www.webrtc.org/reference/architecture
3. Loreto, S., Romano, S.P.: Real-Time Communications in the Web: Issues, Achievements, and Ongoing Standardization Efforts. IEEE Computer Society 16(5), 68–73 (2012)
4. Johnston, A., Burnett, D.: WebRTC: APIs and RTCweb Protocols of the HTML5 Real-Time Web, 1st edn. Digital Codex LLC, USA (2012)
5. HTML5 Rocks: Getting Started with WebRTC (July 23, 2012), http://www.html5rocks.com/en/tutorials/webrtc/basics/
6. The Websocket Protocol (2011), http://tools.ietf.org/html/rfc6455
7. Introducing websockets: Bringing sockets to the web (2010-2012), http://www.html5rocks.com/en/tutorials/websockets/basics/
8. Javascript Session Establishment Protocol, JSEP (2012), http://tools.ietf.org/html/draft-ietf-rtcweb-jsep-03
9. Node.js (2012), http://nodejs.org/
10. Vert.x (2012), http://vertx.io/
11. Vert.x - main manual (2012), http://vertx.io/manual.html
12. Realtime/Working WebRTC Experiments (2012), https://github.com/muaz-khan/WebRTC-Experiment
13. RTCMultiConnection Documentation (2012), https://github.com/muaz-khan/WebRTC-Experiment/tree/master/RTCMultiConnection

Integrating Public Transportation Data: Creation and Editing of GTFS Data

Mário Braga, Maribel Yasmina Santos, and Adriano Moreira

ALGORITMI Research Center, University of Minho,
Guimarães, Portugal
{mario.braga,maribel.santos,adriano.moreira}@algoritmi.uminho.pt

Abstract. The current state of standardization related to representation and exchange of data about public transportation systems is still at its infancy, which leads to severe interoperability issues in projects that depend on the data from several diverse sources. In many cases, the interoperability issues arise from the use of rudimentary information systems, or even paper-based procedures, to manage operational data such as schedules and tariffs. In these cases, exchanging data with external systems is very difficult. This paper describes the development of a web-based application aiming to simplify the creation and editing of public transportation data that could be easily exchanged in a normalized format. This description is preceded by a discussion about a data model that could ease data interoperability. Here, the GTFS reference, with some adjustments, is used as a guideline for the definition of such transportation data model.

Keywords: GTFS, GIS-T, Interoperability, Transportation.

1 Introduction

The growing migration to urban areas leads is leading to the increase of transportation demands, namely parking space, energetic consumption, ambient pollution and traffic congestion, in the transportation of people or goods. Public Transportation Systems (PTS) are often used to mitigate the impact of the aforementioned issues. Through the interaction with PTS, users may benefit from a more efficient, low cost and sustainable journey [1,2,3]. Services provided by PTS can be an added-value to the users, when enhanced with data from different public transportation authorities, and when combined with other information like weather or road congestion.

PTS for buses, subways, trains and other public vehicles are currently available in many cities and countries. These systems, usually complex, combine concepts and technologies from Geographic Information Systems (GIS) and Transportation Information Systems in what is usually known as Geographic Information Systems for Transportation (GIS-T)[4]. While the domain of GIS may be considered to be in a mature state, the Transportation Information Systems are still

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_6,

moving towards standardization. The current state in the transportation systems or in their processes slow down projects like the TICE.Mobilidade[1] where data exchange is so critical.

The TICE.Mobilidade project, a large Portuguese project integrating more than 25 partners, aims to develop a loosely-coupled PTS that provides applications and services that allow users to reduce the costs of their daily travels and promotes the use of PTS. Although this platform seems similar to others PTS, there are some characteristics that distinguishes it from other systems. The core component in the project is a platform that then aggregates data from different types of public transportation sources and provides it to the software development community so they can develop their own applications. This core platform, named One.Stop.Transport (OST), faces many challenges: it has to provide mechanisms to import data from different transportation authorities; store the fed data under an homogeneous data model and; provide mechanisms that allow the development and deployment of applications developed by third party entities. Although all the aforementioned challenges have a critical impact to the success of the OST platform, this paper will only focus on the approaches conducted to develop mechanisms that help the interoperability between the OST platform and public transportation authorities.

As will be described in the remaining of this paper, the OST platform adopted the General Transit Feed Specification (GTFS) as a fed protocol, and this decision must be complemented with the development of tools that help public authorities to manage their data. This paper presents a data model capable of storing and supporting the edition of multiple GTFS feeds, and also a web-based application that makes use of that data model to store the gradual creation of GTFS data.

This paper is organized as follows. Section 2 presents the GTFS model. Section 3 describes the OST platform and its main components. Section 4 presents the proposed extension to the GTFS model, while section 5 introduces the proposed editor application. Section 5 concludes with some remarks about the presented work.

2 General Transit Feed Specification

Many transportation authorities use legacy data models that have little or even no preoccupation with interoperability with other systems. Interoperability is now a crucial area of study in PTS and many efforts have been conducted towards the development of standards. In 2006, Google launched a transit service that merges public transit information and maps in order to provide multi-modal trip planning. Users can easily use this service by picking origin and destination locations in a map and having, as a result, a multi-modal travel itinerary including time, price and trajectory. While the service represents a trivial PTS functionality, the way how Google feeds this service soon became a *de facto* standard.

[1] `http://tice.mobilidade.ipn.pt/`

Looking towards data exchanging and interoperability, Google provides a reference format named General Transit Feed Specification (GTFS) through which public or private transportation authorities can exchanges their data with the Google services. The GTFS rapidly increased its popularity[2] leading many public authorities (which use GTFS format in the Google Transit service) to release their transit information for third-party developers [5]. As a consequence, many transit applications now support the exchange of data in GTFS format. As an example, the Travel Assistance Device, a mobile application that alerts the users when the destination stop is getting near, uses GTFS as the input format to load the stops and schedules [6]. The Graphserver, an open-source multi-modal trip planer, supports the GTFS format in order to load public transportation data to its multi-modal algorithm. The GTFS OpenStreetMap Sync [7] and the OpenTripPlanner[3], both open-source software systems that synchronize GTFS with OpenStreetMaps, allow users to plan a trip that can combine multiple means of transportation. This is also observed in other domains, for instance in intelligent mobile advertising applications or context-aware systems, which use GTFS data in order to display specific advertising [8] or even in the assessment/benchmark of public transportation systems [9].

In the initial version, GTFS only supported the exchange of static data for schedules, stops, agencies, trips, fares and zones, but in order to cover the demand for real-time information, Google launched in 2011 an extension named GTFS real-time[4] that can be used to make real-time information available to the users. With the availability of real-time data, public transportation users spend less time waiting, feel more safe and likely to use public transportation services [10]. This new feature reinforced GTFS as an exchange reference in order to become, in fact, a standard for public authorities to exchange their data.

2.1 Taxonomy

The GTFS format is composed of thirteen entries hold in Comma-separated values (CSV) files in which six of them are mandatory and the remaining are optional. The six mandatory entries hold the information about the agency name, contacts and language (Agency.csv); route short and long name, descriptions, color and type (Routes.csv); trip short name, headsign and direction (Trips.csv); stop name, latitude and longitude and other information like support for wheelchair boarding (Stops.csv); all the information related to the schedules like arrival and departure time (Stop_times.csv) and service days when the trips occur (Calendars.csv). The optional entries increase the value of the public transportation data by adding information about the feed itself (Feed_info.csv); attributes and rules to apply a fare (Fare_attributes.csv and Fare_rules.csv); data that define the trajectory of a trip (Shapes.csv); data that define connections between routes (Transfers.csv); an entry to define exception dates when a service

[2] A total of 703 feeds in 2013.
[3] `http://opentripplanner.com/`
[4] `https://developers.google.com/transit/gtfs-realtime/`

is explicitly active or inactive motivated by a certain event (Calendar_dates.csv) and; an entry (Frequencies.csv) that represents the time between trips (departure at first stop and arrival at last stop).

Although GTFS have been developed to be used as a sharing format, it can be conceptualized as a data model. All the entries of the GTFS clearly define all the relations and proprieties needed to construct an Entity-Relation (ER) diagram. As shown in Figure 1, this approach supported the development of a model with the mentioned entities and a new, named, Zones that holds the ID that links stops to a certain fare rule.

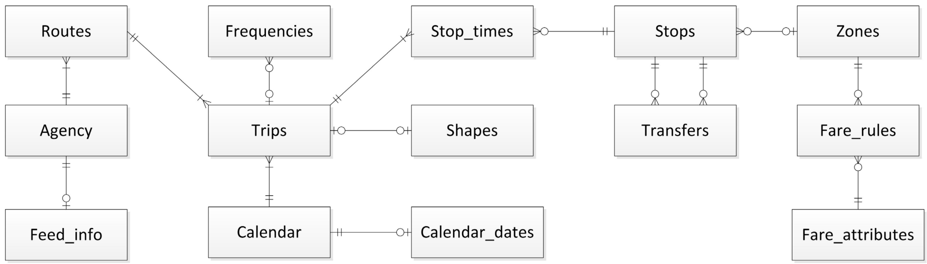


Fig. 1. GTFS Entity–Relationship model

2.2 Tools

For public transportation authorities to validate their own GTFS data, Google provides the *feedvalidator* and *schedule_viewer* applications. The *feedvalidator* inspects the syntax of the feed and ensures that the files and the content match the specifications defined in the GTFS reference. The *scheduler_viewer* application helps users to see the data in a different way providing stop locations and shapes over a map, filter trips or services that occur in a certain date, or provide access to the departures at a certain stop.

Notwithstanding, both applications make use of GTFS-ready data. In order to create new data, public transportation authorities have to make use of other non-official tools or develop their own tool. Currently, there are some alternatives in this matter, as there are some tools[5] available for a fee, and there are open-source projects[6][7] that can be deployed and personalized in such a way that public authorities can create or modify their GTFS files. Yet, there are few applications that allow the creation of GTFS files in a smooth way.

3 The OST Platform

The OST platform intends to provide a sustainable environment where users themselves may develop applications to fulfill their needs or expectations. To

[5] `http://www.transiteditor.com/`
[6] `https://github.com/OneBusAway/onebusaway-gtfs-modules/wiki`
[7] `http://code.google.com/p/transitdatafeeder/`

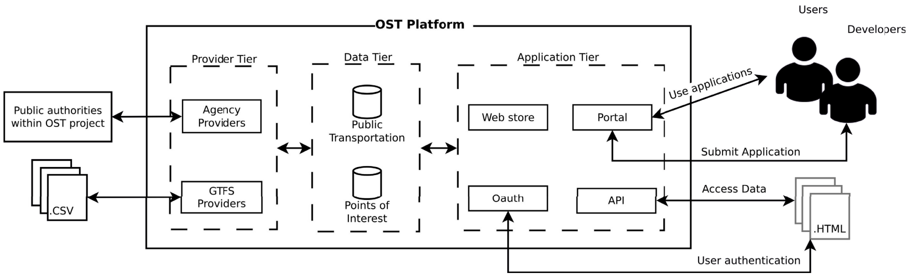


Fig. 2. OST Platform architecture

achieve so, the platform provides ways to import and aggregate data, and also provides data that users can use in their own web-based applications.

The OST platform relies on a loosely-coupled architecture composed by three tiers (see Figure 2). The Provider Tier contains services that load data into the OST platform. Here, the platform supports two different approaches. The first approach connects directly to transportation authorities that take part in the project. This approach showed to demand specific modules specifically crafted to each authority and very difficult to reuse in other authorities. In order to reach other authorities, an alternative approach was followed by using the GTFS reference as a feed protocol to import public transportation data. In this second approach, the transportation authorities are responsible for the creation and submission of the GTFS files to the OST platform.

The Data Tier holds the data models that store data for public transportation and other types of data that enhance the OST platform, such as Points of Interest. As to the last tier, named Application, it provides the interface to interact with other applications/systems or users. The Application tier is decomposable into four main components: a set of services for providing the data so that the software development community can proceed with the development of web-based applications (API); Services for authentication (Oauth); a main web application (Portal) that joins additional information in order to help the community to develop new applications and a place where users can view/access to the already deployed applications and; a Store that handle the hosting of web-based applications. The Store allows the submission of two different types of applications. In one hand, it allows the deployment of applications that will be hosted by the OST platform. On the other hand, it is possible to submit applications that are hosted outside of the OST platform. In this later case, there is no deployment under the OST platform, the only thing that has to be submitted to the Store is the URI to the application. Despite all efforts conducted by the project to improve interoperability with transportation authorities, the current state of the public transportation information systems, in particular the lack of standardization, incapacitates the authorities to provide fully compliant GTFS files to the platform. The major issues are the inconsistency in the models, the quality of the data itself and, in some cases, the absence of mandatory information. To overcome these limitations, it was necessary to create a mechanism that allows the gradual creation, editing and storing of GTFS data up to the point when data completely fulfills the platform's needs. Since the OST platform integrates a set of services for: a) promoting web applications (Store); b) providing authentication mechanisms to OST users (Oauth) and; c) feeding the transportation model with GTFS files (GTFS providers), the adopted approach was to materialize this mechanism as a web-based application connected to a data model capable of storing the gradual creation of GTFS data.

4 GTFS Extension

As was described before, GTFS can be understood as an ER data model. However, because of its primary goal, the original conception of GTFS as a data model reveals some limitations when used in an application that integrates multiple data providers.

4.1 Primary Keys and Source IDs

GTFS is used as a format to allow single public transportation authorities to exchange their data. Theoretically, in this context, each provided ID is unique. However, different authorities can use the same IDs, making it difficult to join data from different providers into the same database. In other words, the data model can't use the provided IDs as primary/foreigner keys.

The issue with the original IDs is that they have substantial importance mainly because they are used as a reference to contextualize some information, for instance, many routes or services are known by their IDs and if they are not stored and showed to users, they may not understand the data that are provided. For instance, the route "1:Linha da Cidade" is known by route 1 and not by route "Linha da Cidade". To overcome this issue, a new attribute, named source_id, was added to the model to store the original ID of each record. This approach also assumes that in case of an update, the data provider can simply search by the original ID and proceed with the modification. Otherwise, it would be necessary to load again all the data.

4.2 Data Providers

The approach mentioned before was also used to detach data from different data providers. One of the requirements for a GTFS editor is that each user (data provider) can only access to its own data. By analyzing the GTFS ER model (see Figure 1), this could be done by using the Feed_info. The Feed_info holds all the information that describes the feed authority that fed the data and is directly linked to the Agency entity. By using this entity it is possible to retrieve the agencies that the authority fed and backtrack all the remaining data that are related to that agency. Such an approach would work whenever all the data are linked and is possible to backtrack all data to the Feed_info. However, this entry is not mandatory and some public authorities may not even include the Feed_info entry in their feed. To overcome this problem, it was added a new attribute (user_id) associated to a new entity, which supports user authentication and additional information about the users.

4.3 Geographic Representation

Aiming to represent the geographic data in a more efficient way, the original representations used in GTFS to store geographic references were replaced by new formats. In its original structure, the geographic location of a stop is represented by two proprieties named latitude and longitude, and the trajectories to define a trip are represented by linking an ordered set of latitude and longitude attributes. Those attributes where replaced by the the creation of the Open Geospatial Consortium (OGC) Points and Linestrings fields, allowing the storage and processing of geographical objects. All the mentioned adaptations were applied to the previous ER and the obtained model is shown in Figure 3.

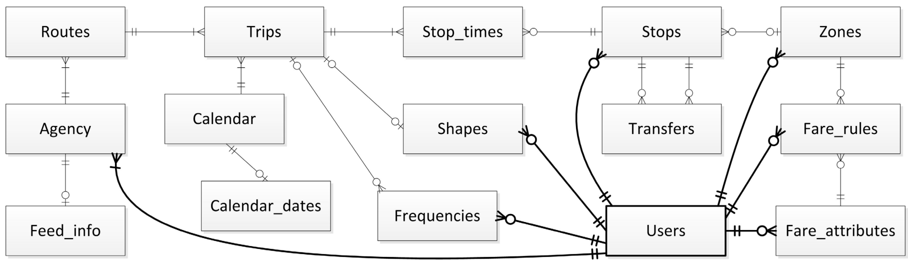


Fig. 3. Extended GTFS Entity–Relationship model

5 GTFS Editor Application

On top of the data model described in the previous section, a web-based application has been developed that enables users to access, edit and store their data in an interchangeable format. The application was conceptualized as a thin client application, in other words, it was divided in two different parts, one that resides in the server side and another that runs on the client side (see Figure 4). Although the editor emerged as an applications of the OST platform, it is not integrated in this platform. This way, users who use the Editor to store their data do not, automatically, make them available in the OST platform. The process of feeding the OST platform stays at the responsibility of the user itself who has to use the GTFS Editor in order to download their current GTFS data from the Editor server and then upload them to the OST platform through the GTFS provider interface (as illustrated in Figure 2). In order to use the GTFS Editor web-application, users can access directly[8] or use the Portal[9] website on the the

[8] `https://hera.dsi.uminho.pt/editorgtfs/web/`
[9] `https://www.ost.pt/`

OST platform. Two types of authentication coexist, one with the application itself (for non-OST users), and or another one that uses the OST authentication services (Oauth).

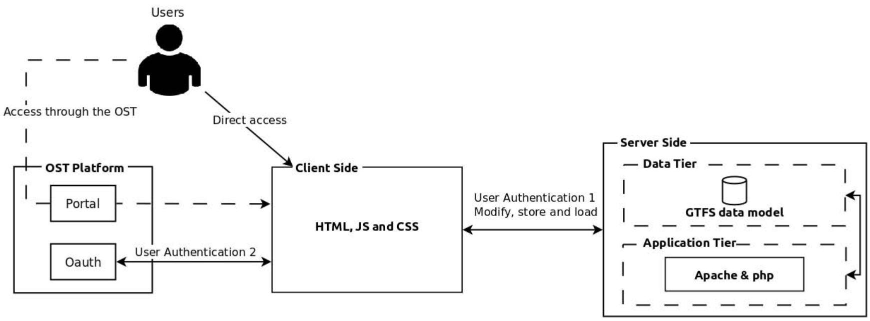


Fig. 4. Application architecture

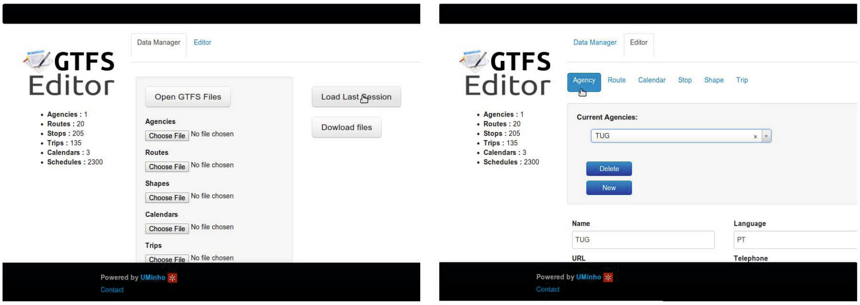


Fig. 5. Data Manager and Agency workspace tab

After login, users can start the editing of data by loading their existing GTFS files (if any) or create new data. Each modification made to the data is stored into the server's database and, at any time or from any place, users can reload their data (Load last session) and resume the edition (Figure 5). The Editor tab (see Figure 5) allows users to enter the workspace where it is possible to create, view, modify and delete data. In the current version, the workspace contains five inner tabs that are related to the following GTFS entities: Agency, Route, Calendar, Stop, Shape and Trips. Here, the inner tab Trips clusters both Trips and Stop_Times entities. One of the concerns in the development process was to create a simple and yet user friendly interface. To achieve so, some processes integrate automatic mechanisms. For instance, the creation of schedules includes smart processes that automatically fill the stops based on the route and calendar information (Figure 6) and a function to estimate the arrival and departure times for a list of sequential stops[10]. In other cases, the application makes use of digital maps to assist the drawing or edition of geographic elements (Figure 7).

[10] The estimation is made by dividing the time of a trip by the total number of stops.

Fig. 6. Fill of stops taking as input the route and calendar

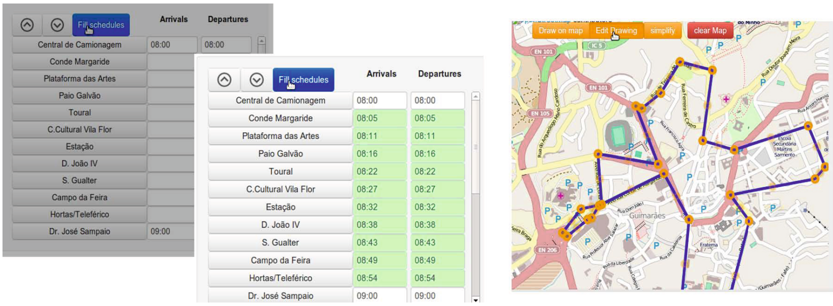


Fig. 7. Estimation of time arrivals/departures and edition of trips trajectories

6 Conclusions

With this work, the GTFS reference has been used to inspire the development of a model that supports multiple GTFS feeds where the data are attached to users; support geographic data types and; save original IDs. The proposed model can also be used as a core data model to many GIS-T applications or services that need to store public transportation information. This model can be seen as an extension of the GTFS reference, with little impact in the original GTFS format, so, the model can evolve to include GTFS real time feature without any compliance issue.

Aiming to facilitate the exchange of transportation data, a web-based editor has been developed. This application allowing users to create, store and edit their own GTFS files. In its current version, the editor supports the mandatory entries: Agency, Routes, Trips, Stop_times, Stops, Calendar and the optional entity Shapes. Future versions should support the remaining optional entities. The present work is crucial to the success of the OST platform since the majority of transportation agencies still use proprietary data models and exchange formats. Most Portuguese public transportation authorities still avoid the adoption of normalized approaches, mainly because they are stuck in legacy systems or methodologies. The developed model and web-application can help those agencies to make progresses in its mechanisms for storing data but also in their current interoperability issues.

Acknowledgment. Research group supported by FEDER Funds through the COMPETE and National Funds through FCT Fundação para a Ciência e a Tecnologia under the projects n. 13843 and PEst-OE/EEI/UI0319/2014.

References

1. Chorus, C.G., Molin, E.J., Van Wee, B.: Use and effects of advanced traveller information services (atis): a review of the literature. Transport Reviews 26(2), 127–149 (2006)
2. Chorus, C.G., Walker, J.L., Ben-Akiva, M.E.: The value of travel information: A search-theoretic approach. Journal of Intelligent Transportation Systems 14(3), 154–165 (2010)
3. Golob, T.F., Regan, A.C.: Impacts of information technology on personal travel and commercial vehicle operations: research challenges and opportunities. Transportation Res. Part C: Emerging Technologies 9(2), 87–121 (2001)
4. Thill, J.-C.: Geographic information systems for transportation in perspective. Transportation Research Part C: Emerging Technologies 8(1), 3–12 (2000)
5. Ferris, B., Watkins, K., Borning, A.: OneBusAway: a transit traveller information system. In: Phan, T., Montanari, R., Zerfos, P. (eds.) MobiCASE 2009. LNICST, vol. 35, pp. 92–106. Springer, Heidelberg (2010)
6. Barbeau, S.J., Winters, P.L., Georggi, N.L., Labrador, M.A., Perez, R.: Travel assistance device: utilising global positioning system-enabled mobile phones to aid transit riders with special needs. IET Intelligent Transport Systems 4(1), 12–23 (2010)
7. Tran, K., Barbeau, S., Hillsman, E., Labrador, M.A.: Go_synca framework to synchronize crowd-sourced mapping contributors from online communities and transit agency bus stop inventories. International Journal of Intelligent Transportation Systems Research, 1–11 (2013)
8. Evans, C., Moore, P., Thomas, A.: An intelligent mobile advertising system (iMAS): location-based advertising to individuals and business. In: 2012 Sixth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS), pp. 959–964 (July 2012)
9. Hadas, Y.: Assessing public transport systems connectivity based on google transit data. Journal of Transport Geography 33, 105–116 (2013)
10. Ferris, B., Watkins, K., Borning, A.: Onebusaway: results from providing real-time arrival information for public transit. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1807–1816 (2010)

Model and Application Architecture Indicators of Evaluation the Enterprise Architecture

Jihane Lakhrouit[1], Karim Baïna[1], and Khalid Benali[2]

[1] Alqualsadi Research Team on Enterprise Architecture, ENSIAS, Mohammed V- Souissi University, Rabat, Morrocco
[2] LORIA – INRIA – CNRS – UMR 7503, BP 239, F-54506 Vandoeuvre-lès-Nancy Cedex, France
jihane.lakhrouit@gmail.com, baina@ensias.ma, benali@loria.fr

Abstract. Enterprise Architecture EA provides a "knowledge base and support for decision making within the enterprise and it serves as the blueprint of current situation and strategy for future directions of the enterprise". The planning and modeling aspect of EA is already fairly well covered in the literature, while the evaluation and the attributes of EA quality have attracted less interest and yet enterprise architecture quality has been perceived as a prerequisite for realizing its benefits [6]. The principle goals of EA evaluation are to document the significance of the EA to decision makers, and to identify the development needs of the EA. However it is very difficult to find a complete method of evaluation in term of indicators of evaluation and which ensures the follow-up of all the stages of development of the architecture of company. This article presents a generic model of evaluation enterprise architecture. We do not want to use attributes directly because the model definition could vary in different organizations. Therefore, we introduce this artifact as an intermediate layer as not to limit the flexibility of evaluation.

Keywords: Enterprise Architecture, EA, Evaluation, Assessment, Quality indicators of evaluation.

1 Introduction

Nowadays, enterprise architecture (EA) has garnered considerable attention from both practitioners and academics in the fields of information systems and business management [1]. Enterprise architecture (EA) is an approach to managing the complexity of an organization's structures, information technology (IT) and business environment, and facilitating the integration of strategy, personnel, business and IT towards a common goal through the production and use of structural models providing a holistic view of the organization [8,9,6].

The end product is a set of artifacts that describe in varying degrees of detail what a business does, how it operates and what resources it requires. These artifacts are often graphic models [10]. To complete the implementation process of the EA, we proceed to an evaluation regarding to the needs. The results of this evaluation will be

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_7,

a useful basis for improving the system in terms to achieve the goal and vision of the organization. This is important because they help to show the strengths and weaknesses of the company and led to the capture and recognition of gaps. This can be used to guide business strategies. This document aims at presenting the possibility to evaluate EA.

The paper is structured as follows: the second section describes the existing evaluation methods of enterprise architecture, the third section presents an evaluation model of the enterprise architecture, the fourth section presents the indicators of EA, and finally the last section is to conclude and propose the perspectives.

2 Background Literature

In the following section, we present our literature review related to our research. It covers existing evaluation methods of enterprise architecture. We can classify the methods of evaluation into 3 categories according to Razavi, M [2].

2.1 Analysis Methods and Tools for Software Quality

These methods are focused on evaluating single software architecture to verify if and where there may be problems in it [2]. We can take as examples of these methods the Software Architecture Analysis Method (SAAM) [4] and the Architecture Tradeoff Analysis Method (ATAM) [5].

2.2 Software Quality Attribute Measurement Methods Based on the Multi Criterion Decision-Making (MCDM) Methods

The Multi criterion Decision-Making (MCDM) are gaining importance as potential tools for analyzing complex real problems due to their inherent ability to judge different alternatives (Choice, strategy, policy, scenario can also be used synonymously) on various criteria for possible selection of the best/suitable alternative (s). These alternatives may be further explored in-depth for their final implementation. We can take as an example of these approaches the article [6]. In this paper, we extended the idea of these methods in the EA domain.

2.3 Evaluation Methods of Enterprise Architecture

Property or quality attribute evaluation of software models is different from quality attribute evaluation of EA models. Software application is one of the four fundamental layers of EA therefore assessing an attribute in enterprise architectures includes evaluating the attributes from different points of view whereas assessing quality attributes in software is from one of these points of view [2]. Despite the importance of this evaluation domain, there is a lack on research on this domain, except some works, that we will detail below, and are divided into two categories: maturity methods and evaluation methods.

- *Maturity methods:*

The goal of these methods is the evaluation of the stage of an organization's EA and to enhance its quality. We can take as examples (U.S. Department of Commerce, 2003; Chief Information Officers Council, 1999; U.S. Government Accountability Office, 2003; Industry Advisory Council, 2005; National Association of State Chief Information Officers, 2003; Office of Management and Budget, 2005). The downside with these maturity models is the fact that they seem to be more or less domain specific, especially developed for the various areas of the public administration, and the maturity models for evaluating the EA of private sector companies, are still hard to find [2]. In the same context there are some works that have used the attributes of maturity models to create an evaluation method, we can take as example [15].

- *Evaluation methods of enterprise architecture:*

The most important evaluation methods in the EA community including [3][2][6][7][13]. A brief overview of these methods is described as below.

Razavi et al. [3] studied the concept of EA quality attributes and identified initial measures for EA maintainability in the context of EA. The advantage of this study, he provides a narrow view of EA quality, and having a fluctuating level of granularity from abstract (EA function) to detailed (source code). But the disadvantage, he studied only one quality attribute. Ylimäki, T [2] proposed a generic evaluation model based on the combination of the 12 potential critical success factors CSF for EA, the key questions assigned to each CSF and the maturity levels to evaluate the stage of each CSF. The advantage is that the method is very clear and well explained. The disadvantage it has not presented all the quality attributes of the EA to make a complete evaluation, we can take as example the CSF alignment of EA with business requirements was not explicitly identified as an EA quality factor [6]. Niemi, E [6] defined 10 quality attributes for EA products and services and sub-factors, to be further empirically validated in different contexts and organizations, and with different EA products and EA services. The advantage, this work presents a large extent of the quality criteria of EA. But the inconvenient, it only presents the non-functional characteristics of EA products, services and processes that comprise the overall quality of EA, omitting the aspect of implemented EA. Lim et al. [7] defined 14 EA quality attributes by analyzing EA frameworks. The advantage, this work presents a large extent of the quality criteria of EA. But the disadvantage it has not presented all the quality attributes of the EA to make a complete evaluation it doesn't present any model or methodology to use it for an evaluation. Khayami, R [13] attempts to introduce mentioned about determine EA qualification and its qualitative characteristics more clearly. This article can be used as a reference to investigate EA qualification and its models. Also, it can help stakeholders to explain the qualitative requirements more exactly, but the inconvenient it doesn't present any model or methodology to use it for an evaluation and has not presented all the quality attributes of the EA to make a complete evaluation.

3 Evaluation Model of Enterprise Architecture

Our evaluation model will be presented in four steps: first, section 3.1 presents the evaluation model, second section 3.2 proposes the evaluation objects, and finally section 3.3 shows how to use the model and how to calculate the Aggregation functions.

3.1 The Concepts of the Evaluation Model

Stakeholders are people involved in developing or using the EA, and therefore have a direct relation with these criteria [17]. Each stakeholder can participate to evaluate one or more indicator and each indicator is evaluated by several actors. There are two main groups of stakeholders in our model: **Direct reviewers:** These represent direct users of the indicators. As direct users of the metrics, they were involved as participants in the research and in the interpretation of the results. The attribute "task_ reviewers" specifies the task of the evaluator in the organization. **Indirect reviewers**: These represent indirect users of the criteria, as they received the criteria information as feedback on their work. These were project analysts working on application development projects, who were responsible for developing the enterprise architecture. The attribute "function_ reviewers" specifies the function of the Reviewer in the development of enterprise architecture in the organization. Each of these stakeholder groups is likely to have a different perspective on the utility of each indicator.

Metrics precise the scale measuring the attribute, the result is attributed to "value_element" specified in the element. It is important to define the indicator and the attributes as well as the belonging metric clearly mission [18]. In our approach we use 4 metrics (0: poor, 1: acceptable, 2: good, 3: very good)

Element specifies how to evaluate each indicator. It may vary from one organization to another one. We introduce this artifact as an intermediate layer as not to limit the flexibility of evaluations. Each company sets its own elements to validate the evaluation indicators. We can use 3 types of attributes: **Questions**: We define a series of questions to evaluate the indicator. The result of each element can vary from 0 to 3. **Checklists**: Checklists are the quickest and cheapest way to evaluate anything, including enterprise architectures. Each evaluator must be assigned a particular role and prepare questions relating to the position in question, for example, might evaluate the architecture or effort in terms of their role only. The affected metric is 0 or 3. **Simulations and prototypes**: This method consists in developing simulations, "What if", and scenarios to validate the hypotheses and conduct experiments about the existing architecture. Sometimes this is the only way to deal with specific issues. The result of each simulation can vary from 0 to 3. The difference between the 3 types of the elements is the type of the attribute "result_element" that precise the response of element: The "result_element" of checklist element is a boolean result. The attribute "result_element" of question and simulation elements is list of four responses and each response is corresponding to one metric from 0 to 3. After developing and measuring the elements of the indicator, we calculate the value "value_indicator" of

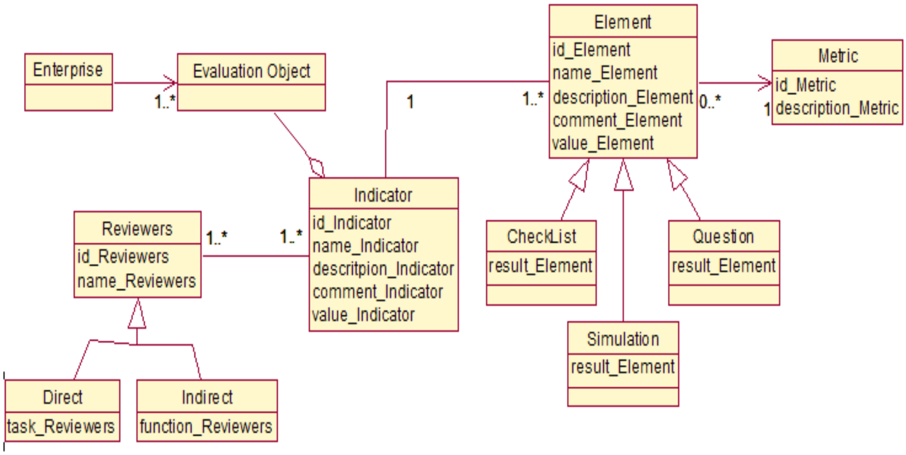


Fig. 1. The evaluation model of enterprise architecture

each indicator as follows: Value indicator = The sum of value of elements in relation with the indicator / the number of these elements.

3.2 The Evaluation Objects

Evaluation Object represents the objects that we must evaluate in the enterprise architecture. Based on the definitions of enterprise architecture and [14] we divide the evaluation object into 6 categories: (1) business architecture, (2) data architecture, (3) application architecture indicator, (4) technology architecture, (5) communication and deliverables and (6) Roadmap.

Roadmaps: Road maps defines the series of broad initiatives and timelines outlining how the enterprise will bridge the gap between the current state as-is and the desired future state to-be [19]. The transition plan designed to convert the “as-is” to “to-be.”“As-is” architecture is a set of documentations related to the functions, services, and technologies currently available at the organization. “To-be” architecture is a set of documentations about policies, strategies, and plans related to correction and development of enterprise architecture development. The transition plan is a plan that includes all activities and components necessary for the transition of “as-is” architecture to “to-be” architecture [20] [21] [22] [23] [24]. **The business architecture** defines the business strategy, governance, organization, and key business processes. **The data architecture** describes the structure of an organization's logical and physical data assets and data management resources. **The application architecture** provides a blueprint for the individual application systems to be deployed, their interactions, and their relationships to the core business processes of the organization. **The technology architecture** describes the logical software and hardware capabilities that are required to support the deployment of business, data, and application services. This includes IT infrastructure, middleware, networks, communications, processing, standards, and so on. **Communication and**

deliverables defines the awareness of the EA, the sharing of information in the organization and the pace and maintenance of the documentation and the deliverables.

3.3 Aggregation Functions

In this section we precise the method of calculating the value of the result of evaluation of EA:

EVAL_EA: the final result of evaluation. **EVAL_DATA**: the result of evaluation data architecture. **EVAL_BUSINESS**: the result of evaluation business architecture. **EVAL_TECHNOLOGY**: the result of evaluation technology architecture. **EVAL_APPLICATION**: the result of evaluation application architecture. **EVAL_COMMUNICATION**: the result of evaluation communication. **EVAL_ROADMAP:** the result of evaluation roadmap.

EVAL_EA = (EVAL_DATA + EVAL_ BUSINESS + EVAL_TECHNOLOGY + EVAL_APPLICATION + EVAL_COMMUNICATION + EVAL_ROADMAP) / 6.

EVAL_DATA = the sum of “value_indicator” of all data architecture indicators**/** the number of data architecture indicators. **EVAL_BUSINESS** = the sum of “value_indicator” of all business architecture indicators **/** the number of business architecture indicators. **EVAL_TECHNOLOGY** = the sum of “value_ indicator” of all technology architecture indicators **/** the number of technology architecture indicators **EVAL_APPLICATION** = the sum of “value_ indicator” of all application architecture indicators **/** the number of application architecture indicators. **EVAL_COMMUNICATION** = the sum of “value_ indicator“ of all communication indicators **/** the number of communication indicators. **EVAL_ROADMAD** = the sum of “value_ indicator“ of all roadmap indicators **/** the number of roadmap indicators.

4 Proposition of Quality Indicators

There are several definitions of the quality attributes in the literature but more used is the IEEE’s definition, Quality is “the totality of features and characteristics of a product or a service that bear on its ability to satisfy stated or implied needs” and the Quality attribute is the element which affects quality degree. In this section we will discuss only the Application Architecture indicators.

The quality indicators proposed are derived from indicator proposed by Moody [25], from software quality indicator defined by McCall, from Boehm quality, from enterprise architecture evaluation indicator defined by the members of Workshop on analysis and evaluation of enterprise architectures [26] and from others works [6, 13, 36, 37].

Agility which addresses an enterprise's ability to manage change, and it is an essential characteristic for the survival of enterprises that have to operate in dynamic environments where change is constant [27]. **Interoperability** means the ability of different software systems, applications, and services to communicate and exchange data in an accurate, effective, and consistent manner [28]. **Maintainability**. Regarding the rapid variations in business environment conditions and necessities, organizations must adapt their application architecture and strategies to new conditions and decisions. Enterprise architecture must to have controllability of accuracy of current functions and must adapt itself to new variables rapidly. This is called maintainability of a system [13, 34, 35]. **Reusability**: The ability of enterprise components to be used in more than one system, such as business reference model and services. Reusable modules reduce the implementation time and cost, increase the likelihood that prior testing and use has eliminated bugs and localizes code modifications when a change in implementation is required. **Reliability**: The architecture does what it is supposed to do, and what the user expects it to do, and it does so without breaking anything in the process: accuracy, error tolerance, consistency and simplicity in design **Testability**: The new architecture is easy to validate, that the software meets the requirements [29]. **Scalability** is the ability of a system, network, or process, to handle a growing amount of work in a capable manner or its ability to be enlarged to accommodate that growth [30]. **Security:** i- Stability: How stable is the system in the sense of in what rate it breaks down and becomes inaccessible or loses data [32]., ii-Validity: To what extend does the system ensure that information within it is correct (for example not tampered with) [31] and iii-Secrecy: How well does the system protect information from being seen by outsiders [33]. **Alignment with business:** The extent to which the architectural processes and deliverables are in tune with what the business wants and is capable of. **Use of an architectural method:** The extent to which a (common) architectural method is used. **Development of architecture**: The approach to architecture development, varying from isolated, autonomous projects to an interactive process of continuous facilitation. **Use of architecture**: The way architecture is used: merely as a conduit for information, as a means of governing individual projects or even as a tool for managing the entire organization.

5 Conclusion

The topic of enterprise architecture has been gaining significant attention from both academia and industry due to the inefficiencies of current IT architecture to cope with rapid changes in business environment. In this paper we have discussed a very important and timely subject which is the evaluation of the enterprise architecture.

We proposed a model that can be adjusted to the needs of different companies by adapting the evaluation objects, the indicators and the evaluation elements. The indicators defined are based on the characteristics of good enterprise architecture and its organization in a quality model. This model can be used to analysis of weak and strong points of an enterprise architecture. The next step in our researches is divided into two areas: the first is a complete example of developing the model including evaluation of all indicators and the second is a prototype of a tool to support such evaluations. This tool has a modular architecture, so that aggregation functions, indicators, metrics, and evaluation objects can be implemented as modules in order to provide extensibility.

References

1. Yalimaki, T.: Potential critical success factors for enterprise architecture 2(4), 29–40 (2006)
2. Yalimaki, T.: Towards a Generic Evaluation Model for Enterprise Architecture. Journal of Enterprise Architecture 3(3), 9–16 (2007)
3. Razavi, M., Shams Aliee, F., Badie, K.: An AHP based approach toward enterprise architecture analysis based on enterprise architecture quality attributes. Knowledge and Information Systems 28(2), 449–472 (2011)
4. Kazman, R., Abowd, G., Bass, L., Clements, P.: Scenario-Based Analysis of Software Architecture. IEEE Software, 47–55 (1996)
5. Clements, P., Kazman, R., Klein, M.: Evaluating Software Architectures: Methods and Case Studies. Addison-Wesley (2001)
6. Niemi, E., Pekkola, S.: Enterprise Architecture Quality Attributes: A Case Study. In: 46th Hawaii International Conference on System Sciences, pp. 38779–38789 (2013)
7. Lim, N., Lee, T., Park, S.: A Comparative Analysis of Enterprise Architecture Frameworks Based on EA Quality Attributes. In: Proceedings of the 10th ACIS International Conference on Software Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing, pp. 283–288 (2009)
8. Kaisler, S.H., Armour, F., Valivullah, M.: Enterprise architecting: Critical problems. In: Proceedings of the 38th Hawaii International Conference on System Sciences, 10 p. (2013)
9. Lange, M., Mendling, J., Recker, J.C.: A comprehensive EA benefit realization model – An exploratory study. In: Proceedings of the 45th Hawaii International Conference on System Sciences (2012)
10. Steenbergen, M.: Architecture Maturity Matrix (2005), http://www.swean.com/dokument/Architecture_maturity_matrix_Intro.pdf
11. IEEE Standard 1471-2000, IEEE Recommended Practice for Architectural Description of Software-Intensive Systems. IEEE (2000)
12. Yalimaki, T.: Enterprise Architecture Compliance: the Viewpoint of Evaluation. In: Proceedings of the European Conference on Information Management and Evaluation (ECIME 2007), September 20-21 (2007)
13. Khayami, R.: Qualitative characteristics of enterprise architecture (2011)
14. The OpenGroup, TOGAF version 9 (2009), http://www.opengroup.org/togaf/
15. Song, H.: Enterprise Architecture Institutionalization and Assessment. In: 9th IEEE/ACIS International Conference on Computer and Information Science, pp. 871–877 (2010)
16. Zachman, J.A.: The Framework for Enterprise Architecture. ZIFA report
17. Moody, D.: Measuring the Quality of Data Models: An Empirical Evaluation of the Use of Quality Metrics in Practice (2006), http://is2.lse.ac.uk/asp/aspecis/20030099.pdf
18. Addicks, J.: Enterprise Architecture Dependent Application Evaluations. In: 3rd IEEE International Conference on Digital Ecosystems and Technologies, pp. 594–599 (2009)
19. Javanbakht, M.: A new Method for Enterprise Architecture Assessment and Decision-making about Improvement or Redesign (2009)
20. Spwak, S.: Enterprise Architecture Planning: Developing a blueprint for Data, Applications, & Technology (1993)

21. NASCIO, Enterprise architecture development Tool-Kit v 3.0, National association of state chief information officer (2004)
22. NASA-ESDIS, Blueprint Technologies. Best practice approach to enterprise architecture (2004)
23. Schekkerman, J.: Enterprise architecture source card. Institute for enterprise architecture development (2004)
24. EA Practice team. Enterprise Architecture Transition Plan, Version 1.0, vol. 1 (2005)
25. Moody, D.: Measuring the Quality of Data Models: An Empirical Evaluation of the Use of Quality Metrics in Practice (2006)
26. Klein, J.: A Workshop on Analysis and Evaluation of Enterprise Architectures (November 2010)
27. Doucet, G., Gøtze, J., Saha, P., Bernard, S.: Coherency Management: Using Enterprise Architecture for Alignment, Agility, and Assurance. Journal of Enterprise Architecture, 1–12 (2008)
28. Lankhorst, M.: Enterprise Architecture at Work, Modeling, Communication, and Analysis. Springer, Heidelberg (2009)
29. Baina, K.: A tool for comparing and selecting workflow engines, pp. 330–337 (2007)
30. Bondi, A.B.: Characteristics of scalability and their impact on performance. In: Proceedings of the 2nd International Workshop on Software and Performance, Ottawa, Ontario, Canada, pp. 195–203 (2000)
31. Finne, A.: Blueprint for General Quality Attribute Model. In: Proceedings of IRIS 28 (2008)
32. Irani, Z.: Information systems evaluation: navigating through the problem domain (2002)
33. Palmius, J.: Criteria for measuring and comparing information systems Information & Management, vol. 40. Elsevier, scandinavia iris (2007)
34. ISO/IEC 9126-1: Software Engineering Software Product Quality, Part 1: Quality model, International
35. Organization for Standardization (ISO), Geneva, Switzerland, 2001. IEEE Std 1219-1998: Standard for Software Maintenance. Institute of Electrical and Electronics Engineers(IEEE), New York (2007)
36. Steenbergen, M.: Architecture Maturity Matrix (2005), http://www.swean.com/dokument/Architecture_maturity_matrix_Intro.pdf
37. Niu, N.: Enterprise Information Systems Architecture—Analysis and Evaluation. IEEE Transactions on Industrial Informatics 9(4) (November 2013)

Software Tools for Project Management – Focus on Collaborative Management

Maria Ester Ferreira and Anabela Pereira Tereso

CGIT Research Centre,
School of Engineering, University of Minho,
4800-058 Guimarães, Portugal
ferreira.ester01@gmail.com, anabelat@dps.uminho.pt

Abstract. Today, projects involve members from different geographical areas more than any other time in history. Thus, adequate Collaborative Project Management Software (CollabPMS) solutions are needed to enable individuals and organizations to manage, communicate and work together across time and space barriers. This article describes a set of managerial and collaborative functionalities that a CollabPMS should provide to support the complexities of a distributed project effectively. Out of hundreds software packages available in the market, four were selected, ProjectLibre, Redmine, Microsoft Project 2013 and Clarizen, to assess if they have the described functionalities. Clarizen can be considered the best software for managing distributed projects, because it provides, by default, all the managerial functions and the collaborative features that support the coordinated collaboration level. ProjectLibre was the software that less stood out in this evaluation, although it provides the majority of the outlined managerial functionalities it doesn't support any level of collaboration.

Keywords: Project Management, Distributed Projects, Collaboration, Software.

1 Introduction

The Project Management Institute, a worldwide leading non-profit professional association in the area of project management, has defined project management as the application of knowledge, skills, tools and techniques to plan activities in order to meet the project's requirements [1]. In the past of project management, projects were typically developed at a single location [2]. However, due to business globalization and technology advancements, a new type of projects called distributed projects have arisen involving project collaborators from different geographical locations and organizations. As the number of distributed projects increases and the project management shifts towards a more collaborative approach, the project management software designed to support the projects at a single location is becoming inadequate [3]. To handle projects involving geographically dispersed teams the project managers should rely on collaborative project management software. This kind of software can help people to manage projects and collaborate more easily regardless of their

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_8, © Springer International Publishing Switzerland 2014

location. In this paper we propose a set of functionalities that a collaborative project management software should have to effectively address the challenges of distributed projects. Currently, the market is loaded with hundreds of project management software solutions to assist in project management. By selecting four software packages available in the market was possible to judge their capability to handle distributed projects. For that, it was evaluated whether they have or not the managerial and collaborative functionalities that we outlined as ideal to handle distributed projects.

In this paper we review some of the problems associated with traditional project management and we explain why collaborative project management is the adequate approach to deal with distributed projects. We also bring up the concept of project management software and collaboration. This paper also proposes a framework of functionalities that a collaborative project management software should have to deal with distributed projects. In the end, a survey was carried out, to see if some software packages are suitable or not to handle efficiently distributed projects. These software packages will be also classified according to the collaboration level that they support.

1.1 Collaborative Project Management

Over the past two decades, the economic globalization and market dynamics have been increasing the need for business partnerships and projects across nations [4]. Thanks to the information technology advances, projects with collaborators from different geographical locations, organizations and cultural backgrounds became in large number effectively implemented [5, 6]. Issues like negotiation of the project goals, scheduling, task allocation, parallel working on the same task and resource sharing requires particularly high degrees of collaboration among distributed team members [2]. As collaboration has become an important part of project management, the traditional project management paradigm has been shifting toward a more Collaborative Project Management (CollabPM) paradigm [4]. The traditional project management approach focuses on a single project at a single location and some of its common mistakes include: ineffective information flow and communication between all the team members, distant project tracking, a reactive management approach and lacks of a document management system. Together all these mistakes account for the reason why many distributed projects either fail or are significantly less efficient than they could be [2–4]. The collaborative project management model is more concerned with efficient sharing of information in all directions, better communication among project contributors, close tracking of the project work, proactive management and the presence of a document management system. These are some of the important issues required to successfully implement distributed projects [6, 7]. The CollabPM paradigm has emerged as a vehicle by which the cost and duration of distributed projects can be potentially reduced while maintaining the quality and scope of the projects [3].

1.2 Project Management Software

The project management paradigm has changed not only in terms of how projects are managed but also in terms of computing environments. In the past, project management meant three things: pencil, paper and brains. However, over the last decades, the arising of personal computers (PC) and the growth of project management software (PMS) has especially contributed to change the face of project management [3, 8]. Nowadays, it is almost unthinkable that anyone would plan and manage a project without the support of PMS. Managing a project involves a considerable data and information analysis that cannot be easily handled without the aid of PMS. Even for simple projects developed at a single location, making changes and updating the entire network is something difficult to do without using a PMS [9]. Moreover, unlike local projects, in globally dispersed projects, face-to-face interactions among project contributors are often impossible. Thus, as the number of distributed projects increases and the project management shifts towards a CollabPM paradigm, information and communication technology solutions are needed for converting physical collaborative actions into virtual ones [2]. As organizations deal with projects with distributed teams, the management, communication, coordination and tracking of ongoing project work must rely on a specific kind of PMS, called Collaborative Project Management Software (CollabPMS) [7]. The CollabPMS helps people to manage projects and collaborate more easily regardless of their geographical location, in more efficient and effective ways than ever before to successfully achieve the project's goal [4, 10]. Although many companies claim they have CollabPMS packages, they are not all equal. Depending on the type of computer applications that they have, these CollabPMS can support different levels of collaboration. As people collaborate, there are at least two hierarchic levels in which they can work to achieve the project's purpose, namely coordinated level and concerted level [11] (Figure 1).

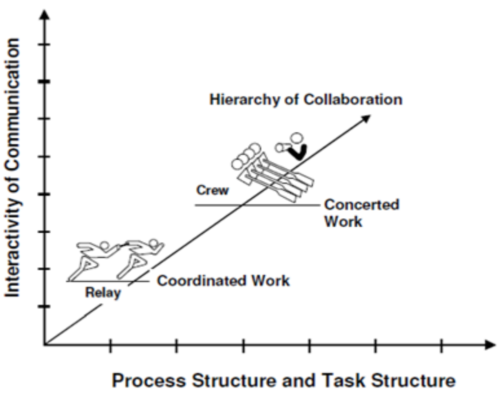


Fig. 1. The hierarchy of collaboration, adapted from [11]

Coordinated Level - In this collaboration level, team members make individual performances and the success of each team member depends on the timely receipt of deliverables from other members. Therefore, team productivity depends on the ability to coordinate efforts and is the sum of the sequential individual performances. Regarding the task structure, the project tasks follow a sequential order and there is a progressive integration of the work. Because, in this level, the coordination among project team members is required, there is a great need of interactive communication. Typical computer applications to support coordinated work include electronic mail, workflow automation and group calendaring [11].

Concerted Level - Represents a higher level of collaboration. At this level, the performance of any member may directly and immediately influence the performance of all members. Thus, the concerted level requires tight coordination among project individuals and the team productivity is the sum of all concerted team performances. Task structure is more demanding for concerted work than for coordinated work because results need to be continuously integrated and any behavior of each team member directly affects the work of others. Furthermore, at this stage of collaboration, the need for interactive communication is nearly continuous. A typical computer application to support concerted work is called Group Support System (GSS) [11].

2 Analysis of Project Management Software Packages

A number of project management architectures have been proposed, as Dixon's (1988) model which only involves planning and scheduling the activities, resource identification and allocation and cost management. Dixon's model does not include document management system and hasn't also any collaborative aspect. Therefore it underestimates the complexity of distributed projects and the collaboration required to make them successful. Maurer (1996) presents a generic architecture that allows users analyze dependencies between information items and plan the project in terms of time and resources. The Maurer's model also supports workflow management and project monitoring. This model does allude to collaboration; however it focuses only on the coordination level, and does not address the concerted level [4].

Our research goal is to outline which important functionalities a CollabPMS should provide to support the complexities of a distributed project effectively. Therefore, the following sections will describe the managerial functionalities that a CollabPMS should have to support an efficient management of distributed projects. Additionally, because projects with team distribution requires high degrees of collaboration among them, to successfully achieve the project's goal, this research also describes which collaborative functionalities a CollabPMS should have to support the different levels of collaboration. Currently, the market is loaded with hundreds of software packages to assist in project management. From the current generation, four software packages were selected, namely, ProjectLibre, Redmine, MPP 2013 and Clarizen to assess whether they have or not the managerial and collaborative functionalities that we outlined as ideal to handle distributed projects.

2.1 Managerial Functionalities

An ideal CollabPMS should provide a set of managerial functionalities capable of answering to the challenges placed by the project team and providing, at the same time, an integration of all the project's processes over its life cycle [12]. Based on the work developed by Jaafari and Manivong (1998), we describe a series of managerial functions (table 1, first column) that a CollabPMS should offer to help perform an efficient management of distributed projects. Using a project example adapted from Hillier and Lieberman (2010) about constructing a new plant for a major manufacturer was possible to evaluate whether ProjectLibre has (✓) or hasn't (✗) the managerial functionalities outlined to perform efficient project management. Moreover, based on published information, that is, scientific papers and information sourced from the software companies' websites, it was possible to perform the same evaluation on Redmine, Microsoft Project Professional 2013 (MPP 2013) and Clarizen. The results of these evaluations can also be found in Table 1.

Table 1. Managerial functionalities matrix

Software Name	ProjectLibre	Redmine	MPP 2013	Clarizen
Type of software	Desktop	Web-based	Desktop	Web-based
License	Open source	Open source	Proprietary	Proprietary
Company	ProjectLibre	Redmine	Microsoft	Clarizen
Price	Free	Free	$1,159 per PC	$29,95 /user/month
Trial version	-	-	✓	✓
Gantt Chart	✓	✓	✓	✓
Task dependencies	✓	✓	✓	✓
Critical path	✓	✗	✓	✓
Milestones	✓	✗[1]	✓	✓
Resources Allocation	✓	✗[1]	✓	✓
Budgeting	✓	✗[1]	✓	✓
Simulation	✓	✓[2]	✓	✓
Setting Baseline	✓	✓	✓	✓
Project tracking	✓	✓	✓	✓
Portfolio Management	✗	✓	✗[3]	✓
Reporting	✓	✓[4]	✓	✓
Document Management	✗	✓	✗[5]	✓
Importing and Exporting data	XML	XML	XML, CSV, TSV	XML

[1] There is the option of setup a free plugin.
[2] Only scheduling simulations.
[3] There is the option of setup Project Server 2013 (sold separately).
[4] Reporting only regarding work in progress.
[5] There is the option of setup SharePoint 2013 (sold separately).

Project scheduling plays a crucial role in ensuring the success of a project [9]. Analyzing Table 1, it is possible to verify that all the software packages, provide the functionality Gantt chart and task dependencies [13–15]. Furthermore, ProjectLibre, MPP 2013 and Clarizen allow to calculate and highlight the critical path in the Gantt chart, while Redmine does not calculate neither displays the critical path [16, 17]. Milestones are useful for the project manager to evaluate if the schedule itself is proceeding as expected [9]. ProjectLibre, MPP 2013 and Clarizen let establish milestones along the project's schedule [18, 19]. The Redmine program does not have this feature. Yet, this issue can be solved by installing a free plugin from the Redmine plugins directory which adds milestones [20]. For the planning phase to be completed, after scheduling the project's activities, it is necessary to deal with resources and costs planning [9]. ProjectLibre, MPP 2013 and Clarizen allow to carry out the allocation of resources [21, 22]. Redmine does not allow resource allocation, however, once again, this problem can be overcome by installing a free plugin for assigning resources [23]. ProjectLibre, MPP 2013 and Clarizen allow to calculate the cost of each resource and, in the end, establish a budget for the project [21, 24]. By installing a costless plugin for budgeting, Redmine is also able of manage costs [25]. In relation to the simulation or what-if scenarios, like ProjectLibre, both MPP 2013 and Clarizen allow the project manager to go into schedule, resources and cost simulations to find the best answer to the project constraints. For Redmine to be able to simulate trade-offs like resources versus time, or costs versus time, the project manager has to install the plugins for budget and resources allocation, otherwise with Redmine it is possible only to make simulations on the project schedule. Regarding the baseline functionality all the software packages offer the option of saving the baseline plan [26–28]. All the software packages also afford tracking the project progress [21, 28, 29]. Both Clarizen and Redmine offer the Portfolio Management functionality, but like ProjectLibre, the MPP 2013 does not provide the management of a portfolio. However, it is possible to unlock this functionality by installing the Microsoft Project Server 2013, which is sold separately from MPP 2013 and is a flexible program for project portfolio management [30]. In reference to the reporting functionality, ProjectLibre, MPP 2013 and Clarizen offer a set of predefined reports to see the overall project's status. These reports can provide information related to work in progress, resources and budget. Redmine also provides a predefined report, but this report only provides information on the work in progress [21, 31]. The functionality of document management system is displayed for Clarizen and Redmine [21, 28]. ProjectLibre does not provide a document management system, like MPP 2013. The SharePoint 2013 is a paid web application platform developed by Microsoft Corporation that can interface with MPP 2013 and includes document management [32]. ProjectLibre, Redmine and Clarizen can interface with the different versions of MPP because both allow to export and import entire projects in the XML format. Additionally, to create XML files, MPP 2013 allows to import and export data in comma-separated values (CSV) file format and Tab Separated Values (TSV) file format [28, 33].

2.2 Collaborative Functionalities

Particularly for projects which have their team members spread in different places and tasks depending on different team members' performance, the collaboration between the members is crucial [34]. We used the work developed by Nunamaker et al. (2002) about the typical computer applications that are required to support the coordinative and concerted levels of collaboration, to define which collaborative functionalities a CollabPMS should provide to efficiently support the distributed teams. We define like them that and ideal CollabPMS must be designed to provide the concerted level of collaboration. It is the concerted level of collaboration that adds real value to project management, therefore an ideal CollabPMS must afford the GSS technology (see section 1.2) [4]. Moreover we use their work to classify the software packages ProjectLibre, Redmine, MPP 2013 and Clarizen according to the collaboration level that they support. Therefore, was necessary to evaluate whether these four software packages have (✓) or do not have (✕) the typical collaborative functionalities that are required to support each of the two collaboration levels. Table 2 displays the results of this evaluation.

Table 2. Matrix of collaborative functionalities

Software Name	ProjectLibre	Redmine	MPP 2013	Clarizen
Type of software	Desktop	Web-based	Desktop	Web-based
License	Open source	Open source	Proprietary	Proprietary
Company	ProjectLibre	Redmine	Microsoft	Clarizen
Price	Free	Free	$1,159 per PC	$29,95 /user/month
Trial version	-	-	✓	✓
Email	✕	✓	✕[6]	✓
Workflow Automation	✕	✓	✕[7]	✓
Group calendaring	✕	✓	✕[7]	✓
GSS	✕	✕	✕	✕

Clarizen offers its clients several collaborative functionalities as email, workflow automation and group calendar [21, 35]. Nevertheless, Clarizen has not adopted the GSS technology. The MPP 2013, commercialized by Microsoft Corporation, does not provide by default the email and the workflow functionality. But, by setting up Microsoft Lync 2013, emails between team members can be exchanged [36]. Moreover, by installing Microsoft SharePoint, it is possible to have workflow automation [37]. Regarding group calendar, MPP 2013 does not have this capability; however, by setting up SharePoint, this drawback can be solved [38]. It is important to keep in mind that both Lync and SharePoint are sold separately from MPP 2013. Like Clarizen, MPP 2013 does not include GSS. Redmine provides email, workflow

[6] There is the option of setup Lync 2013 (sold separately).
[7] There is the option of setup SharePoint (sold separately).

and group calendar functionalities but does not provide the GSS [28]. ProjectLibre does not have any collaborative functionality to support any of the two levels of collaboration.

3 Discussion

Our architecture considers more factors or functions (table 1 and 2) than the previous two models. This architecture allows establish interdependency of tasks, forming the project team, assigning resources and tasks to team members, defining milestones for the project, making project schedule, project tracking, document management system, reporting, portfolio management, import and export data. Only a software architecture with all this managerial functions is able to effectively supporte distributed teams [12]. Moreover, only a collaborative platform that allows the concerted level of collaboration can facilitate group discussion and negotiation [4].

The review of table 1 shows that Clarizen is the only software offering by default, all the managerial functionalities that ideal software should have to support an efficient project management. The MPP 2013, commercialized by Microsoft Corporation, is also able to provide the same overall managerial functionalities. However, for that, their clients must pay for complementary programs (see section 2.1). Redmine does not provide one of the essential managerial functionalities, the critical path analysis function. Capabilities like milestones, resource allocation and budgeting, which are not also provided by default, can be installed as plugins by the users for free. ProjectLibre like Redmine is an open source software with free license, but in ProjectLibre lacks two functionalities, namely, the Portfolio management and document management system (table 1). Although the four PMS packages range between open source and proprietary, desktop and web-based application, all the selected PMS somehow cover the traditional management areas, that is, scheduling, resource allocation, costs and project control, areas that are essential for project management [3, 12]. It should be noticed that Redmine was included in this observation because it offers free plugins that allow specifically to manage these traditional areas. Considering the price of each PMS, there are two distinct groups: group 1- ProjectLibre and Redmine with free licensing and group 2 - MPP 2013 and Clarizen with paid licensing. It is possible to verify that the companies of group 2, Clarizen Incorporated and Microsoft Corporation are able to provide their clients with respectively a PMS or a combined PMS solution which frames into the concept of what is an ideal PMS to manage projects efficiently [12]. The study of Liberatore & Pollack-Johnson (2003) contributes somehow to this observation. According to them, high-end packages, that is, those PMS packages that are more expensive, tend to have more features capable of handling with huge and complex projects than the less expensive ones. They also state that larger firms are those that can usually afford more expensive PMS [39]. The concerted level of collaboration truly offers effective and efficient support for managing the complexities of a distributed project [4]. At this level, all the team members must contribute in concert to the group efforts. Therefore, the performance of any team member influences the other members' performance [4, 7]. Group Support System (GSS) is the commonly used technology to support the concerted level of collaboration [11]. The GSS can be defined as an information technology specifically designed to provide structured process support for group meetings. The GSS technology allows team members collaborate and increases the effectiveness of decision making while reducing travel expenses [40]. However, none of the four software packages already discussed have invested in GSS technology, as a consequence they don't support group problem solving and decision making [3, 6]. Therefore, none of the software packages support concerted work. The Redmine and Clarizen are both web-based software packages and have all by default the functionalities to address the coordinated level of collaboration; explicitly they have email, workflow automation and group calendar capabilities. In the coordinated level, the team success depends on the ability to coordinate efforts [3]. While Redmine offers all the capabilities of the coordinated work for free, in Clarizen it is required to pay, that is, it is needed to purchase the software. The MPP 2013, plus its third party applications as SharePoint and Lync 2013, can also provide coordinated work to its users. But, for that the costumer must be willing to pay for the MPP 2013 and an extra for web-based third party applications. Moreover ProjectLibre doesn't support any of the two collaboration levels, it doesn't present any collaborative feature therefore can be considered inadequate for managing distributed projects [3]. It is possible to notice that collaboration in distributed projects needs a web-based infrastructure; this observation supports the finds of Chen et al. (2006) that software packages with web interface are the ones that can fully support collaboration among dispersed team members. The author Romano et al. (2002) developed a study about collaborative project management software and his conclusions were that most of the CollabPMS available at that time in the market just provided the lower levels of collaboration. For example, the Microsoft Project 2000 which was one of the selected software packages by Romano provided at that time coordinated work [7]. Nowadays, the version of Microsoft Project 2013 (MPP 2013) still provides the coordinated level. In the past, researchers had been proposing and developing CollabPMS prototypes to support concerted level [4, 7]. At the present, there is still the need to improve the software packages to provide concerted collaboration to companies who need it.

4 Conclusions

None of the four tested software packages have all the functionalities that we describe as ideal to support the implementation of distributed projects, because none of them provides the GSS functionality, and therefore they are not able to afford the concerted level of collaboration. As a result it is needed that these software packages and possibly the entire PMS market moves forward to provide a continuous collaboration among distributed team members. Nevertheless, in the overall scenario Clarizen can be considered the best software for managing distributed projects because it is the only one that, by default, provides all the outlined managerial functions and all the collaborative features that support the coordinated collaboration level. ProjectLibre was the software that performed the worst in this evaluation because of lack of support for any level of collaboration despite providing the majority of the outlined managerial functionalities.

References

1. Project Management Institute: A Guide to the Project Management Body of Knowledge. Project Management Institute (2008)
2. Evaristo, R., van Fenema, P.C.: A typology of project management: emergence and evolution of new forms. Int. J. Proj. Manag. 17, 275–281 (1999)
3. Chen, F., Romano, N.C., Nunamaker, J.F.: A Collaborative Project Management Approach and a Framework for its Supporting Systems. J. Int. Technol. Inf. Manag. 15, 1–16 (2006)
4. Chen, F., Nunamaker, J.F., Romano, N.C.: A Collaborative Project Management Architecture. In: Proceedings of the 36th Hawaii International Conference on System Sciences (2002)
5. Helbrough, B.: Computer assisted collaboration - the fourth dimension of project management? Int. J. Proj. Manag. 13, 329–333 (1995)
6. Nunamaker, J.F., Romano, N.C., Briggs, R.O.: A framework for collaboration and knowledge management. In: Proceedings of the 34th Annual Hawaii International Conference on System Sciences (2001)
7. Romano, N.C., Fang, C., Nunamaker, J.F.: Collaborative Project Management Software. In: Proceedings of the 35th Hawaii International Conference on System Sciences (2002)
8. Reiss, G.: Project Management Demystified: Today's Tools and Techniques. Spon Press (1995)
9. Marchewka, J.T.: Information Technology Project Management: Providing Measurable Organizational Value. Wiley Publishing (2003)
10. Nidiffer, K.E., Dolan, D.: Evolving Distributed Project Management. IEEE Softw. 22, 63–72 (2005)
11. Nunamaker, J.F., Romano, N.C., Briggs, R.O.: Increasing Intellectual Bandwidth: Generating Value from Intellectual Capital with Information Technology. Gr. Decis. Negot. 11, 69–86 (2002)
12. Jaafari, A., Manivong, K.: Towards a Smart Project Management Information System. Int. J. Proj. Manag. 16, 249–265 (1998)
13. Microsoft Corporation: Choose the right view of your project schedule, `http://office.microsoft.com/en-us/project-help/choose-the-right-view-of-your-project-schedule-HA102749464.aspx?CTT=5&origin=HA102926132`
14. Clarizen Incorporated: Clarizen Releases Version 2.5G with Time Tracking and Gantt Chart, `http://www.clarizen.com/media-center/press-releases/clarizen-partners-with-mulesoft-to-accelerate-customer-time-to-value-with-automated-application-integrations/item/clarizen-releases-version-2-5g-with-time-tracking-and-gantt-chart.html`
15. Redmine: Gantt Chart, `http://www.redmine.org/issues/14582`
16. Microsoft Corporation: Show the critical path of your project, `http://office.microsoft.com/en-us/project-help/show-the-critical-path-of-your-project-HA102748968.aspx`

17. Clarizen Incorporated: How To View the Critical Path, https://success.clarizen.com/entries/24078911-v5-How-To-View-the-Critical-Path
18. Microsoft Corporation: Add a milestone, http://office.microsoft.com/en-us/project-help/add-a-milestone-HA102748898.aspx
19. Clarizen Incorporated: What is a Milestone, https://success.clarizen.com/entries/24128683-What-is-a-Milestone-
20. Redmine: Plugins Directory - Milestones, http://www.redmine.org/plugins/redmine_milestones
21. Clarizen Incorporated: Features A - Z, http://www.clarizen.com/project-management-solutions/online-project-management-features.html
22. Microsoft Corporation: How do resource assignments drive the schedule?, http://office.microsoft.com/en-us/project-help/how-project-schedules-tasks-behind-the-scenes-HA102749287.aspx?CTT=1#_Toc312058394
23. Redmine: Plugins Directory - Allocation, http://www.redmine.org/plugins/redmine_allocation
24. Microsoft Corporation: Enter rate-based costs for people and materials, http://office.microsoft.com/en-us/project-help/enter-rate-based-costs-for-people-and-materials-HA103465589.aspx?CTT=1
25. Redmine: Budget Plugin, http://www.redmine.org/projects/redmine/wiki/pluginbudget
26. Microsoft Corporation: Set and save a baseline, http://office.microsoft.com/en-us/project-help/set-and-save-a-baseline-HA102809638.aspx
27. Clarizen Incorporated: Save as Baseline, https://success.clarizen.com/entries/24092307-v5-Save-as-Baseline
28. Redmine: Redmine Features, http://www.redmine.org/
29. Microsoft Corporation: Project Professional (2013), http://office.microsoft.com/en-us/project/professional-project-management-desktop-software-project-professional-FX103797571.aspx
30. Microsoft Corporation: Choose your Project, http://office.microsoft.com/en-us/project/project-management-and-ppm-software-comparisons-microsoft-project-FX103797367.aspx
31. Microsoft Corporation: Pick the right report, http://office.microsoft.com/en-us/project-help/pick-the-right-report-HA102843515.aspx?CTT=1
32. Microsoft Corporation: Plan document management in SharePoint (2013), http://technet.microsoft.com/en-us/library/cc263266.aspx
33. Clarizen Incorporated: Can I export my work plan?, https://success.clarizen.com/entries/24127676-Clarizen-Community-FAQ
34. Evaristo, J.R., Munkvold, B.E.: Collaborative infrastructure formation in virtual projects. J. Glob. Inf. Technol. Manag. 5, 29–47 (2002)
35. Clarizen Incorporated: Clarizen: Creating Workflow Solutions to Transcend Generations, http://technorati.com/business/article/clarizen-creating-workflow-solutions-to-transcend

36. Microsoft Corporation: Using Lync, http://office.microsoft.com/en-us/sharepoint-server-help/use-a-group-calendar-HA101810606.aspx?CTT=1
37. Microsoft Corporation: Introduction to workflows, http://office.microsoft.com/en-us/sharepoint-server-help/introduction-to-workflows-HA010154424.aspx
38. Microsoft Corporation: Use a Group Calendar, http://office.microsoft.com/en-us/sharepoint-server-help/use-a-group-calendar-HA101810606.aspx?CTT=1
39. Liberatore, M.J., Pollack-Johnson, B.: Factors influencing the usage and selection of project management software. IEEE Trans. Eng. Manag. 50, 164–174 (2003)
40. Hayen, R.L., Swaby, S.A., Huang, Z.: Use of Group Support Systems in Today's Society. Issues Inf. Syst. 7, 120–126 (2007)

An Adaptable Infrastructure to Generate Training Datasets for Decompilation Issues

Javier Escalada and Franciso Ortin

University of Oviedo, Computer Science Department,
Calvo Sotelo s/n, 33007, Oviedo, Spain
escaladajavier@uniovi.es, ortin@lsi.uniovi.es

Abstract. The conventional decompilation approach is based on a combination of heuristics and pattern matching. This approach depends on the processor architecture, the code generation templates used by the compiler, and the optimization level. In addition, there are specific scenarios where heuristics and pattern matching do not infer high-level information such as the return type of a function. Since AI has been previously used in similar scenarios, we have designed an adaptable infrastructure to facilitate the use of AI techniques for overcoming the decompilation issues detected. The proposed infrastructure is aimed at automatically generating training datasets. The architecture follows the *Pipes and Filters* architectural pattern that facilitates adapting the infrastructure to different kind of decompilation scenarios. It also makes it easier to parallelize the implementation. The generated datasets can be processed in any AI engine, training the predictive model obtained before adding it to the decompiler as a plug-in.

Keywords: decompilation, automatic pattern extraction, automatic dataset generation.

1 Introduction

Disassemblers are tools to facilitate the examination of binary code. Without them, binary code consists of a mere sequence of numeric codes. There are many scenarios where it is necessary to analyze applications for which the source code is not available. Examples are malware analysis, maintenance of programs in production, and reverse engineering of file formats and communication protocols.

A disassembler performs the relatively simple and mechanical task of translating groups of numbers to the textual representation of the associated instructions (mnemonic). However, the abstraction level of the output assembly programs remains very low. Therefore, the user must still try to guess the algorithms and data structures of the original high-level source code used to compile the binary file. For this reason, the analysis of binary code is complex, tedious and error prone, even with the use of disassemblers [1].

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_9,

Decompilers are aimed at improving this situation. These tools transform binary programs into high-level source code (e.g., C, Pascal or COBOL) that preserves the semantics of the original code. However, obtaining the exact original source code from the binary code it is an undecidable problem [2, 3] because of the high-level information discarded in the compilation process. This information is precisely the one that facilitates the understanding of the semantics of the program.

The most widespread approach used to retrieve this missing information is to apply some simple heuristics [3, 4] in combination with pattern matching [5, 6]. These patterns or idioms (caused by the compiler code generator) can be structural (the graph structure of the code), binary (assembly instructions) or behavioral (memory or character string manipulations). However, it seems that this approach has certain limitations [1] such as knowing when a function actually returns a value, inferring the types used in the original program, and detecting the size of the value returned by a function. Therefore, some authors have previously proposed solutions based on Artificial Intelligence (AI): to detect *Function Entry Points* (FEPs) [7], applying best-first search systems for control flow structuring [8], and to categorize sections of binary files between code and data sections in systems based on the Von Neumann architecture [9]. Since these previous works obtained satisfactory results, we believe that AI-based approaches may complement conventional ones to improve the existing decompiler tools.

The decompilation process shows a high influence in three factors. The first one is architecture of the processor (e.g., Intel x86, ARM or PPC), since each architecture has its own assembly language. The second one is the code generation patterns used by the compiler. Each compiler has its own templates for compiling the different high-level structures [10]. The last factor is the optimization level. With a elevated level of optimization, compilers use different generation patterns for the same high-level structures [11]. This increases the search space and hence also increases the differences between the decompiled source code and the original one. Therefore, each pattern and heuristic found by a human expert depends on these factors. If one factor changes, new patterns and heuristics need to be defined. In addition, experts must command these factors. On contrary, machine learning techniques do not require explicit knowledge of these topics. A predictive model is trained with the inherent information obtained from a dataset [12], so we just have to automatically build these datasets.

The main contribution of this paper is the development of an adaptable infrastructure to allow the automatic generation of datasets independently to the above-mentioned factors. In subsequent research works, these datasets will be used to study AI-based decompilation approaches as a complement to conventional ones.

The rest of this paper is structured as follows. Section 2 describes a motivating example, and the proposed infrastructure is presented in Section 3. Section 4 discusses related work, and Section 5 concludes and describes our future research work.

2 A Motivating Example

The most widely used decompilers, including Hex-Ray 1.5, do not detect whether the original function actually returns a value or not. For example, after compiling with MSVC++ 2012 the program in Figure 1, the function in Figure 2 is later obtained when decompiling it with Hex-Ray for Intel x86. Figure 2 shows how the decompiled function declares a return value (`int`), whereas the original function (`strset` in Figure 1) does not.

```
void strset (char *str, char value, unsigned int len) {
    char * current = str;
    if (str == NULL || len == 0)
        return;
    while (len--)
        *current++ = value;
}
```

Fig. 1. strset function code

```
01: int __cdecl sub_401000(int a1, char a2, int a3) {
02:   int result;                 // eax@1
03:   int v4;                     // ecx@4
04:   int v5;                     // [sp+0h] [bp-4h]@1
05:   result = a1;
06:   v5 = a1;
07:   if ( a1 && a3 ) {
08:     while ( 1 ) {
09:       v4 = a3--;
10:       if ( !v4 )
11:         break;
12:       result = v5;
13:       *(_BYTE *)v5++ = a2;
14:     }
15:   }
16:   return result;
17: }
```

Fig. 2. Decompiled version of strset in Fig. 1

The decompiled version returns `result` after inferring that the function uses the C standard calling convention (`__cdecl`) [13, 14]. In this calling convention, the retuned value is stored in the `EAX` register. Therefore, `result` is the C representation of the `EAX` processor register (as shown in the comment in the line 2, Figure 2). The problem is that `EAX` is not only used to return function values. The Intel ABI (Application Binary Interface) allows the compiler to use `EAX` as a general-purpose register. In line 12 of Figure 2, `EAX` gets the value of `v5` and it is later used (the assignment line 13). The decompiler does not have sufficient information to infer if the value of `EAX` is a residual value (caused by the expression of line 12) or the value returned by the function (valuable). Therefore, the second option is always chosen.

From a practical point of view, a decompiler does not detect whether the value of `EAX` is residual or valuable because it does not have a proper heuristic for it. A human expert has not found a pattern for this scenario, and a conservative approximation is used: functions are always declared with a return value, which may not be used in its invocations.

In the above example we can also see other limitations. The first one is the correct distinction between address and value variables. For instance, the first argument of `strset` in Figure 1 is declared as a `char *`, whereas it is an `int` in the decompiled version (Figure 2). Another lack is the correct inference of types. In Figure 2, the third argument is incorrectly decompiled as `int` (it is `unsigned` in Figure 1).

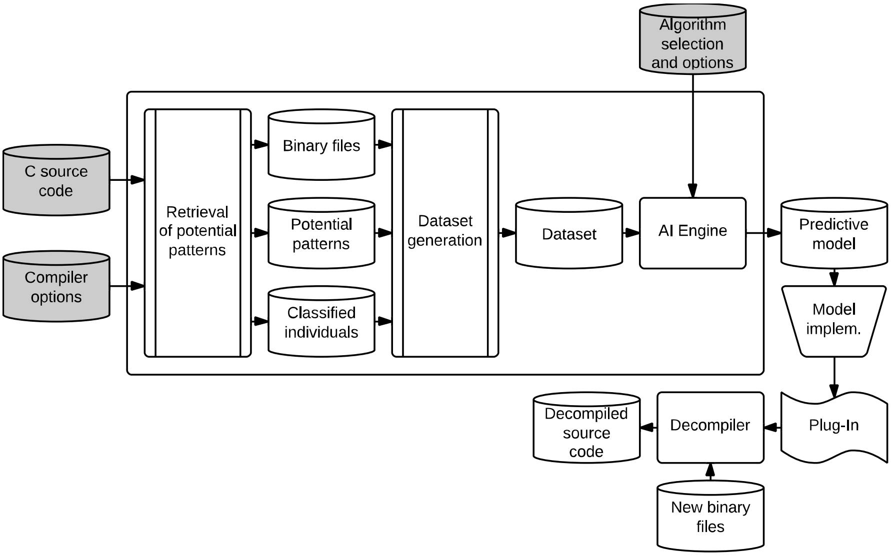


Fig. 3. The infrastructure in the global process

3 The Proposed Infrastructure

3.1 The Architecture

Figure 3 shows the proposed architecture. Gray elements identify input arguments. Our infrastructure follows the *Pipes and Filters* architectural pattern [15]. It facilitates the modification and replacement of the processing elements (*filters*), adjusting the infrastructure to the specific needs of a concrete problem (see Sections 3.2 and 3.3).

Our infrastructure receives *Compiler options*, *C source code* and *Algorithm selection and options* as inputs. The first process, *Retrieval of potential patterns*, extracts low-level *Potential patterns.* It also selects and classifies the individuals (*Classified individuals).* An individual is an element defined by a set of features and classified with respect to them. The *Binary files* are generated as a result of an internal process (Section 3.2), stored to be used as input for a later process (Section 3.3).

For each selected individual, the *Dataset generation* process checks for the occurrence of the potential patterns in the *Binary files*. The output is a *Dataset* that indicates for each high-level code fragment (functions in our example), the occurrence of potential patterns. This dataset is used by the *AI engine* to train a *Predictive model* capable of classifying new individuals under a predefined error coefficient.

At this point, the new heuristic described by the *Predictive model* could be added to a specific *Decompiler*. In order to do this, the *Predictive model* must be implemented by hand, and added to the *Decompiler* as a *Plug-in*. Therefore, the *Decompiler* will be improved, obtaining source code more similar to the original one.

In this article, we focus on the internal details of the first two processes (*Potential patterns retrieval* and *Data set generation*), which we have already implemented. In subsequent articles, we will describe how we have used the *Datasets* generated with the proposed infrastructure in different decompilation scenarios.

3.2 Retrieval of Potential Patterns

Figure 4 shows the internal structure of the *Retrieval of potential patterns* process first shown in the infrastructure architecture (Figure 3). As in the previous diagram, grey items indicate those elements to be provided by the user. Unlike the *C source code* and *Compiler options* inputs, the remaining gray colored elements (*Instrument code*, *Extract & classify individuals*, *Sentinel* and *Block to pattern*) only have to be configured in an initial adjustment phase. After performing this initial adjustment, the infrastructure simply processes a collection of *C source code* inputs.

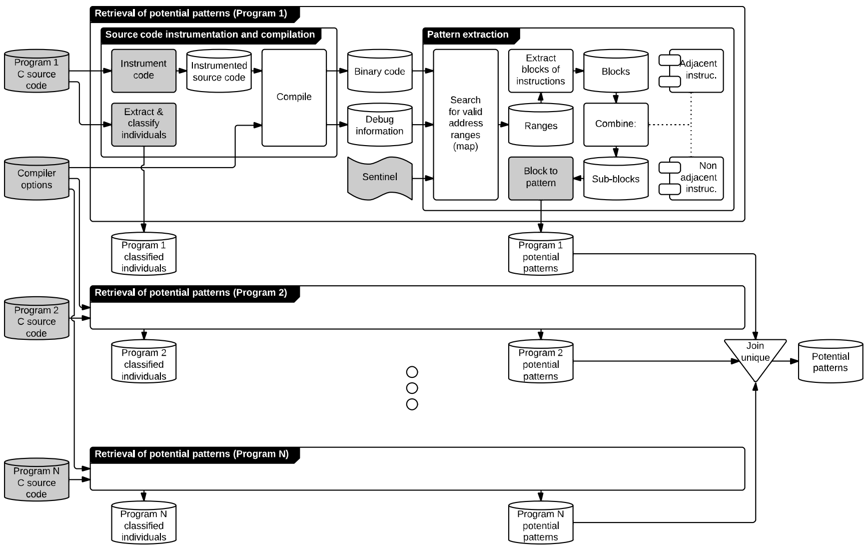


Fig. 4. Retrieval of potential patterns process

```
int foo(int x, int y) {
  if (x < 0 || y < 0)
    return 0;
  if (x > y)
    return x + bar(x);
  return y + bar(x);
}
```

```
int foo(int x, int y) {
  if (x < 0 || y < 0)
    __RETURN1__: return 0;
  if (x > y)
    __RETURN2__: return x + bar(x);
  __RETURN3__: return y + bar(x);
}
```

Fig. 5. Example function and the instrumented version

The *Instrument code* process highlights specific high-level patterns in the *C source code*. In the illustrating problem, the process highlights all the `return` statements with a C label, as shown in Figure 5. This instrumentation process is implemented with a source code instrumentation tool –we have used Clang (LLVM) [16].

In parallel, the *Extract & classify individuals* process produces a list containing the dataset individuals and its classification (*Classified individuals*). In our example, functions are the individuals to be classified as functions returning a value or not. We have also used Clang [16] in the implementation of this process.

Then, the *Instrumented source code* is compiled with debugging information. In our implementation, we use the Microsoft PDB format. This information is extremely important because it allows us to link high-level language structures with the corresponding instruction blocks in the binary files (*Search for valid address ranges*). In our example, the debugging information is used to link the C labels with concrete addresses in the binary file. A *Sentinel* function is also used in this process. This function is aimed at identifying the addresses associated to the high-level language structures. In the motivating example, it is a Python function checking whether a given address has a label with the following pattern: `__RETURN\d+__`. If so, it selects the *range* between that address and the last one in the current function branch.

The *Ranges* of instruction addresses are the input of the *Extract blocks of instructions* process. Not all the instructions in these blocks are relevant to classify the individuals. For example, all the *n* instructions in a block may not be relevant to detect whether the function returns a value or not. However, we do not have enough information to know which ones are the relevant. Therefore, we break these blocks into smaller potential patterns. The user can select between tow implementations of the *Combine* process. The first one combines *Adjacent instructions*. For *n* instructions, the possible combinations are: 1 block of *n* instructions, 2 blocks of *n-1* instructions and so forth, until *n* blocks of 1 instruction. The second implementation is to merge different combinations of *Non-adjacent instructions*. The user must select a specific implementation in the initial adjustment phase.

The last *Block to pattern* step exports all the blocks of combined instructions (*Sub-blocks*) in a format understandable by the following process (Figure 6). This format is problem specific. In our example, we store the instructions in reverse order (patterns are matched in this order) together with the relative offset to the last instruction in the function.

The whole *Retrieval of potential patterns* process shown in Figure 4 is applied for each *C source code* input. Once the *Potential patterns* for each source program have been extracted, they are joined in a single collection of *Potential patterns*, erasing duplicates (*Join unique*).

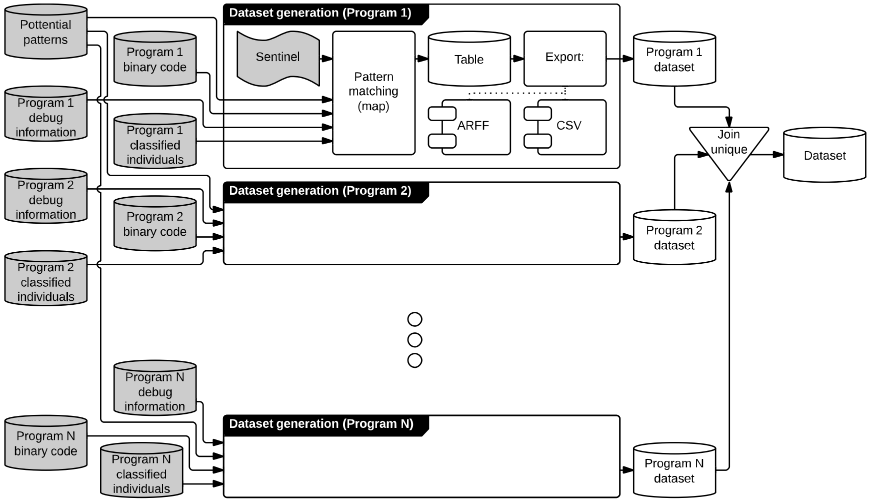


Fig. 6. Retrieval of potential patterns process

3.3 Dataset Generation

Figure 6 shows the structure of the *Dataset generation* process previously identified in Figure 3. This process takes the *Potential patterns*, the *classified individuals* and the *binary code* with *debugging information* generated in the previous process, and produces a *Dataset* indicating the occurrence of *Potential patterns* in high-level code fragments. For each individual, *Pattern matching* checks the occurrence of *Potential patterns* in the related address ranges inside the *binary code*. This check is done through the *Sentinel* function, configured in the initial adjustment task, which must understand the custom format of the *Potential patterns* (Section 3.2).

The output is a binary *Table* summarizing the occurrences. In the function return problem, this table has one row for each function in the source code and one column for each potential pattern identified. The values are whether that pattern is matched in the associated function.

Finally, the *Export* process translates this table into some input format for the *AI engine* (Figure 3). As we used Weka in our implementation, the concrete implementation of *Export* we chose was ARFF. This process can be replaced with other implementations, such as CVS and JSON. Each program generates a different dataset. *Join unique* takes all the datasets and generates one, erasing duplicated individuals.

3.4 Parallelization

For an effective predictive model, we need to compute datasets covering a large search space. This implies working with a big amount of programs, performing computation intensive tasks with each one of them (i.e., instrumentation, compilation, and binary code analysis). In order to reduce execution time of the whole process, our infrastructure has been designed as a set of independent tasks (rectangles in Figure 4 and Figure 6), following the *Pipes and Filters* architectural pattern. This pattern promotes parallelization of tasks (*filters*), executing each one when its input is available, and producing the corresponding output [15]. Besides, the *Join unique* process shown in Figure 4 and Figure 6 waits until all its inputs have been generated, and generates a single set.

Moreover, the large amount of independent data to process also facilitates data (and pipe-line) parallelization. The two *dataset generation* and retrieval of *potential patterns* are parallelized by processing different programs concurrently. Since the general objective of finding potential patterns in binary files would commonly require a high number of input programs, this data parallelization will make the most of supercomputing systems.

Therefore, any existing implementation of the MapReduce [17] approach (e.g., Hadoop, Amazon Elastic MapReduce or another cloud-based alternative) could be an effective solution for paralleling our infrastructure. In particular, we have identified BinaryPig [18], built on top of Apache Hadoop, Apache Pig and Elasticsearch, as an appropriate tool to facilitate the parallel processing of source and binary files. Currently, we have implemented a prototype Python scheduler that includes tasks to a pool of worker processes. A new task is added to the list of scheduled tasks when all the inputs it requires are available for processing.

4 Related Work

Extracting binary patterns from a big set of applications in an automatic way has already been applied in previous research works. In [7], the authors perform automatic extraction of a huge set of binary patterns, called *idioms*. They extract all the possible groups of 1, 2 or 3 instructions from the binary files. Then, using forward feature selection, these idioms are reduced to a subset. This subset is used (with other structural patterns) to train a Conditional Random Field (CRF) [19], capable of detecting Function Entry Points (FEPs). Unlike our proposed solution, they used specific (and therefore not reusable) processes.

In [20], the same authors use the same approach to classify sections of binary files depending on the compiler used. They utilize custom processes to generate the datasets. A manual processing of binary files to classify some sections compiled with different compilers is applied (statically linked system libraries).

There are other areas apart from decompilation, where it is also necessary to process binary files automatically. In these areas, there exist research works that use *ad hoc* techniques to extract features in binary files. In [21, 22], the authors perform the automatic extraction of features from Windows programs. With the extracted features, they compare some methods to detect whether an executable file is packed or not.

5 Conclusions

The automatic extraction of heuristics and low-level patterns from high-level source code can offer improvements to the existing decompilation tools. We have implemented a generic infrastructure, and adapted it to the example problem of determining whether a C function returns a value or not. Our infrastructure is able to extract the relationship between high-level language structures and its representation in an assembly language. In addition, our approach is independent of the processor architecture, the language compiler, and the optimization level. To facilitate the adaptation of our infrastructure, we have designed it following the *Pipes and Filter* architectural pattern. This pattern also makes it easier the task and data parallelization of the infrastructure, reducing the execution time of processing large volumes of data.

We are currently using this infrastructure to extract patterns for training supervised classifiers. Thus, we are analyzing the potential application of AI to some of the problems mentioned in this article. The first problem we have tackled is the example presented in this article, but the following future work is to extend it for inferring better information about variable types and the use of pointers.

Acknowledgements. This work has been partially funded by the Spanish Department of Science and Innovation under the National Program for Research, Development and Innovation: project TIN2011-25978.

References

1. Guilfanov, I.: Decompilers and beyond. Black Hat USA (2008)
2. Troshina, K., Chernov, A., Derevenets, Y.: C Decompilation: Is It Possible. In: Proceedings of International Workshop on Program Understanding, Altai Mountains, Russia, pp. 18–27 (2009)
3. Troshina, K., Chernov, A., Fokin, A.: Profile-based type reconstruction for decompilation. In: 2009 IEEE 17th International Conference on Program Comprehension, pp. 263–267. IEEE (2009)
4. Cifuentes, C.: A structuring algorithm for decompilation. In: Proceedings of the XIX Conferencia Latinoamericana de Informática, pp. 267–276 (1993)
5. Schwartz, E., Lee, J., Woo, M., Brumley, D.: Native x86 decompilation using semantics-preserving structural analysis and iterative control-flow structuring. In: USENIX Secur. Symp. (2013)
6. Fokin, A., Derevenetc, E., Chernov, A., Troshina, K.: SmartDec: Approaching C++ Decompilation. In: 2011 18th Working Conference on Reverse Engineering, pp. 347–356. IEEE (2011)
7. Rosenblum, N., Zhu, X., Miller, B., Hunt, K.: Learning to analyze binary computer code. In: Proceedings of the 23rd Conference on Artificial Intelligence, Chicago, pp. 798–804 (2008)

8. Van Emmerik, M.: Boomerang: Information for students, http://boomerang.sourceforge.net/students.php
9. Wartell, R., Zhou, Y., Hamlen, K.W., Kantarcioglu, M., Thuraisingham, B.: Differentiating code from data in x86 binaries. In: Gunopulos, D., Hofmann, T., Malerba, D., Vazirgiannis, M. (eds.) ECML PKDD 2011, Part III. LNCS, vol. 6913, pp. 522–536. Springer, Heidelberg (2011)
10. Watt, D., Brown, D., Sebesta, R.W.: Programming Language Processors in Java: Compilers and Interpreters and Concepts of Programming Languages (2007)
11. Muchnick, S.S.: Advanced compiler design and implementation (1998)
12. Alpaydin, E.: Introduction to Machine Learning. The MIT Press (2010)
13. MSDN: Calling conventions, http://msdn.microsoft.com/en-us/library/k2b2ssfy.aspx
14. Jönsson, A.: Calling conventions on the x86 platform, http://www.angelcode.com/dev/callconv/callconv.html
15. Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P., Stal, M.: Pattern-Oriented Software Architecture. A System of Patterns, vol. 1. Wiley (1996)
16. Clang: a C language family frontend for LLVM, http://clang.llvm.org/
17. Dean, J., Ghemawat, S.: MapReduce: Simplified Data Processing on Large Clusters. Commun. ACM 51, 1–13 (2008)
18. Hanif, Z., Calhoun, T., Trost, J.: BinaryPig: Scalable Static Binary Analysis Over Hadoop. Black Hat USA 2013 (2012)
19. Lafferty, J., McCallum, A., Pereira, F.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Proc. 18th International Conf. on Machine Learning 18th International Conf. on Machine Learning, pp. 282–289 (2001)
20. Rosenblum, N.E., Miller, B.P., Zhu, X.: Extracting compiler provenance from program binaries. In: Proceedings of the 9th ACM SIGPLAN-SIGSOFT Workshop on Program Analysis for Software Tools and Engineering, PASTE 2010, p. 21. ACM Press, Toronto (2010)
21. Ugarte-Pedrero, X., Santos, I., Bringas, P.G.: Structural feature based anomaly detection for packed executable identification. In: Herrero, Á., Corchado, E. (eds.) CISIS 2011. LNCS, vol. 6694, pp. 230–237. Springer, Heidelberg (2011)
22. Santos, I., Ugarte-Pedrero, X., Sanz, B., Laorden, C., Bringas, P.G.: Collective classification for packed executable identification. In: Proceedings of the 8th Annual Anti-Abuse and Spam Conference on Collaboration, Electronic Messaging, CEAS 2011, pp. 23–30. ACM Press, Perth (2011)

Reengineering of Software Requirement Specification

Abeer AlSanad and Azedine Chikh

College of Computer and Information Sciences, King Saud University, Saudi Arabia

Abstract. Software requirements engineering is an important subset of software engineering. There has been many researches study the software requirements engineering field and specially the Software Requirement Specification (SRS). In this paper, a reengineering of the software requirement specifications document, according to an updated version of the IEEE 830 standard, as a semi-structured data using XML technology, was made. It aims to enrich the semantic of the SRS. In addition, it serves the customer and the supplier to have a better understanding and experiencing with SRS. A prototype has been developed to show how the new SRS model works. Also, a case study was taken to illustrate the proposed XML schema. Finally, the performance of the proposed approach was assessed according ten main criteria required in any good SRS.

Keywords: Reengineering, Software requirement specification, XML schema.

1 Introduction

Obviously, software engineering is getting into fast development environment. Software engineering is an important field and we can notice its importance through the increase need and wide spread of using software and gradually raising of the demand for software engineers.

Software engineering can be defined as an engineering discipline that is concerned with all aspects of software production. Its generic activities in all software processes are: requirement specification which describes what the system should do and its development constraints, development: production of the actual software system, validation: checking that the software is what the customer really wants, and evolution: which is changing the software in response to changing demands [1]. This paper will mainly concern about the first activity, software requirement specification, and its documentation. Software requirement has been defined by many researchers. Ian Sommerville, defines software requirement as “The process of establishing the services that the customer requires from a system and the constraints under which it operates and is developed.” [2]. He considered software requirement as a process that mainly concerns with the services and constraints.

At the beginning of software development projects, software developers should start with a software requirements gathering process. The outcome of this process is a document which is usually referred to as a Software Requirements Specification

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_10, © Springer International Publishing Switzerland 2014

(SRS). It is considered as the first deliverable of a project. SRS is very important and could not be ignored. Its main function is to record the client's requirements and business needs shown in a written form. It turns out to be the basis for the rest of the software development process. Client and developer based their understanding of the software's requirement on the SRS [3].

Fortunately, there is a standard for writing SRS from IEEE called IEEE 830 and it is available since 1998. But this standard is old and has not been updated since the time it is delivered. Because of that, in this paper, a suggestion to update the IEEE 830 standard was proposed. This update will improve and clarify the standard more and make it easier to apply for SRS stakeholders. Usually it is difficult to deal with requirement specification documents in its written form. Also, software stakeholders cannot use them efficiently because of the general complexity of the document. So a reengineering of SRS document is needed. The reengineering of SRS document was made in this paper. The SRS document was written as semi-structured data to make it more semantic and more reasonable and easier to write, understand and manage for customers, suppliers and other stakeholders. It makes the dealing with requirement easier and more professional. This was made by writing the updated IEEE 830 standard using XML.

XML stands for Extensible Markup Language used to describe data. An XML is a very powerful language to define a data exchange format especially if there are no predefined formats and we have a lot of parameters. XML defines its syntax using either Document Type Definition (DTD) or an XML Schema. It permits the authors to define whatever tags they need and also design the structure they want [4], [5].

After presenting the updated standard using XML, ten criteria was used to evaluate our approach. These criteria are: complete, consistent, correct, modifiable, ranked, testable, traceable, unambiguous, valid, and verifiable. Then, a prototype and a case study were used to show how the proposed approach can work and be applied.

In this paper, an improvement for the IEEE 830 standard was made. In addition, the updated IEEE 830 standard was represented as an XML schema. Moreover, our approach was judged against the ten criteria that should be satisfied in any SRS to be considered a good SRS. Furthermore, a prototype was built to show how the proposed model of SRS works. Also, a case study on the Library Management system of the University of Ballarat was used to apply our model.

The following section contains some related works to the current work. Section 3 related to our contribution covers the updated IEEE 830 standard, its representation using the XML schema, the performance measures of the proposed approach, the system prototype and the case study. Finally, a conclusion of this work and some future works are presented in section 4.

2 Related Works

In 2003, Lee et. al. proposed a technique for creating an SGML/XML document from paper-based documents. This technique takes text regions with hierarchical structure as input. It based on analyzing syntactic logical structure and produces it more accurately and quickly than previous works of which the basic units are text lines. Their approach was tested and the results showed that the logical structure analysis was successful and it generated a document model automatically. Mainly, the approach analysis the structure and generates SGML/XML documents that enhances the reusability of documents [6].

In 2003, Ishitani proposed a new approach to convert printed documents to XML documents using OCR. The proposed approach used a hierarchical transformation strategy that depends on a pivot XML document which based on the XHTML DTD. By using this hierarchical strategy, there is an ability to obtain several different types of XML document from the same pivot XML document. By applying this approach to a range of documents, results showed that the approach is effective for different types of printed documents [7].

In 2007, Mehrdad Nojoumian in his research work presented in [8] aims to ease the dealing of complex document and permit navigation of structure and semantics described by the document. He worked re-engineering PDF-based documents to enhance electronic documents for example specifications, conference proceedings, technical books, etc to facilitate dealing with them by end-users. He planned to take a raw PDF version of a published specification, and convert it into a hypertext document. He expected that if the developers of specifications use the format he developed in publishing their documents, it would to great extent aid end-users of the specifications. He divided his overall approach into two distinct phases: Phase 1 dealt with extracting the document's logical structure and representing the result in XML. The result includes the content information and excludes irrelevant details of the original document's presentation. Phase 2 of the approach work on facilitating user dealing with the document, which includes browsing, navigating, and concept exploration. This can be done by creating a multi-layer HTML version of the document.

3 Reengineering the Software Requirement Specification

The first step of writing SRS is choosing the standard template to follow where many are available. A requirements documentation standard is very important to provide concise, readable, easy to digest, quality requirements. These standards do a great job in supporting the consistency of the documentation efforts of a team. When using standard in writing SRS, a document written by many team members appears that it has been written by a single person [9].

One of the most important standards is IEEE 830 standard which is described in the next section.

3.1 Updated IEEE 830 Standard

In general, IEEE Standards documents are developed in the IEEE Societies and the Standards Coordinating Committees of the IEEE Standards Association (IEEE-SA)

Standards Board. The IEEE develops its standards through a development process that approved by the American National Standards Institute [10].

Although IEEE 830 standard is a powerful addition to the SRS world, it has not been revised since 1998. This explains why it supports some old concepts. An improvement to the IEEE 830 standard has been done through this research work by changing the structure of the most important and heavy part of the SRS, which is the Specific Requirements section.

This improvement was made by changing the divisions of Specific Requirements section appears in Fig. 1 part (A) to a new division shown in Fig.1 part (B). Instead of dividing the specific requirement to external interfaces, functions, performance requirements, logical database requirements design constraints, software system attributes, organizing the specific requirements, additional comments, it is divided based on requirements' categories identified by Soren Lauesen [11] which are: data requirement, functional requirement, nonfunctional requirement, managerial requirement, also, other requirement and members was added. This division will make the SRS more modern, clearer and easier to write and understand. In this paper, we called the standard after modification the updated IEEE 830 standard.

3. Specific requirements: 3.1 External interfaces. 3.2 Functions. 3.3 Performance requirements. 3.4 Logical database requirements. 3.5 Design constraints. 3.6 Software system attributes: 3.7 Organizing the specific requirements. 3.8 Additional comments.	3. Specific requirements: 3.1 Data Requirement. 3.2 Functional Requirement. 3.3 Non Functional Requirement. 3.4 Managerial Requirement. 3.5 Other Requirement. 3.6 Members.
(A)	(**B**)

Fig. 1. Specific requirement section of IEEE 830 standard (A), specific requirement section of Updated IEEE 830 standard

3.2 XML Schema of the Updated IEEE 830

In this section, a representation of the SRS of UPATED IEEE 830 standard as semi-structured data was made to make it more semantic and more reasonable and easier to write, understand and manage for customers, suppliers and other stakeholders. This representation was made through creating an XML schema. The schema was developed by using Altova XML Spy 2011 Enterprise software. The resulted XML schema encompasses two parts: XML Tree and Meta Data. The XML Tree is consisting of the elements in the Schema. Meta Data is represented by the attributes and some elements in the schema. A list of attributes and their significance are described next.

The SRS has many attributes, as shown below:

- Project_ID : attribute of type ID
- Project Name: attribute of type string. It contains the name of the project that SRS is written for.
- Domain: is an attribute of type string. It determines the domain in which the software fall in. Example of domain could be finance or human resource (HR).
- Project Type: is an optional attribute of type string. It determines the type of the project. Example of project type could be In House development, Tender and Sub-contracting.
- Description: is an optional attribute of type string. It contains a brief description of the project.

In addition, the SRS has many elements and some of which have attributes. A list of elements which have attributes and their description are shown below:

1. History: is an element that has one attributes and a set of unbounded elements Version:
   a. SRS_Date: is an attribute of type Date. It shows the first day date of writing SRS.
   b. Version: is an element with two attributes (Version_ID of type ID and Version_Date attribute of type Date that it shows the date of publishing the corresponding version).
2. Authors: is an element that has a unbounded element Author that has three attributes:
   a. Author_ID: attribute of type ID
   b. Author_First_Name: is an attribute of type string.
   c. Author_Last_Name: is an attribute of type string.
3. Validation: is an element that has a unbounded element Validator that has three attributes:
   a. Validator_ID: attribute of type ID
   b. Validator_First_Name: is an attribute of type string.
   c. Validator_Last_Name: is an attribute of type string.
4. Appendices: is an element that has an unbounded element Appendix that has two sub-elements:
   a. Title: is an element of type string.
   b. Content: is an element of type string.

5. Requirement Specification: is an element that has an unbounded element Requirements that has six sub-elements:
   a. Data Requirement: contains the input and output data to/from the system and the data should be stored in system. It could be written in different styles: data model, data dictionary, data expression and virtual window. It is an element that has two sub-elements and 12 attributes which are:

- Elements are:
    - Validation: an element that has unbounded sub-elements Validator.
        - Validator has Validator_ID of type IDREF and Validation_Status of type string attributes.
        - Data_Requirement_Content: is an element of type string. It contains the content of the requirement.
- Attributes are:
    - Data_ID: an attribute of type ID.
    - Parent_Req_ID: an attribute of type IDREF. It refers to the requirement which it composed of.
    - Category: is an attribute of type string. It determines the category of a requirement. Requirements are fall in one of these categories: data requirement, functional requirement, non-functional requirement, managerial requirement and other requirement. Since the value of the category should be one of these categories a numeration facet tool in XML was used to accept only one of these five values.
    - Level: is an attribute of type string. It shows the level of the requirement. There are four levels of requirements which are: business, domain, product and design levels. Since the value of the level should be one of these four levels, a numerations facet tool in XML is used to put a restriction on the level value so it will accept only one of these four values.
    - Style: is an optional attribute of type string. It is optional since not all categories of requirements follow a specific style. An example is “Use Case” that is one of Functional Requirement styles.
    - Notation: is an optional attribute of type string. It is optional since not all types of requirements have several notations. It shows which kind of notation is used in specific requirement style. One example of notation value is UML, which is one of the notations of “Use Case” style.
    - Source: is an attribute of type string. It shows the source of requirement. Where does the author get the requirement from?
    - Testability: is an attribute of type Boolean. It shows whether the requirement is tested before or not.
    - Traceability: is an attribute of IDREF type. It likes a foreign key which will correspond to an ID of some other requirement. It refers to a document or a link that could describe and follow the life of a requirement, in both forwards and backwards direction.

- Author_ID: attribute of type IDREF type. Like a foreign key which will correspond to an ID of the author who wrote the requirement.
- Date: Date attribute is of type Date. It shows the date and time of writing the requirement.
- Rank: attribute of type string. Its value is restricted to high, medium or low.

b. Functional Requirement: specify the functions of the system, how it records, compute, transforms, and transmits data. It could be written in different styles, some of which are: context diagram, event list and function list, feature requirement, screens, task description and many others. It is an element that has 12 attributes. These attributes are the same as in data requirement instead we replace Data_ID with FR_ID attribute and Data Requirement Content with Functional Requirement Content.

c. Non-Functional Requirement: specify how well the system performs its intended functions. Non-Functional requirement deals with system's performance, usability, maintenance and others. Performance describes how efficiently should the system work with the hardware. Usability describes how efficiently should the system work with the users. Maintenance describes how easy should it be to repair defects, add new functionality and any other maintenance issues. Non Functional Requirement attributes are the same as in data requirement instead we replace Data Requirement Content element with Non Functional Requirement Content and Data_ID attribute with Non_FR_ID attribute and add additional attribute which is Fun_ID:

- Fun_ID: an attribute of type IDREF. It makes a link with a functional that is compatible with it.

d. Managerial Requirement: contains information about deliverables like when they will be delivered, the price and when to pay it, how to check that everything is working (verification), what happens if things go wrong (legal responsibilities, penalties), etc. In addition, it specifies who owns the software and other intellectual properties. Also it includes a specification of the development process to be used. Managerial Requirement attributes are the same as in data requirement instead we replace Data_ID with Manag_ID attribute and Data_Requirement_ Content element with Managerial Requirement Content.

e. Other Requirement: contains other requirements that do not fall on the previous categories. Other Requirement attributes are the same as in data requirement instead we replace Data_ID with Other_ID attribute and Data_Requirement_Content with Other_Requirement_Content.

f. Members: is an element that has an unbounded element Member that has three attributes:

- Member_ID: attribute of type ID
- Member_First_Name: is an attribute of type string.
- Member_Last_Name: is an attribute of type string.

By combining the XML Tree and the Meta Data, a XML Schema emerged. This schema is shown in the following figures. Fig.2 shows the first part of the schema and the main divisions of it.

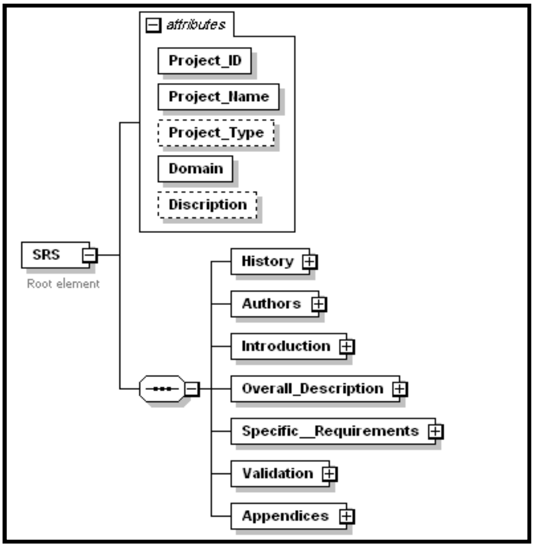


Fig. 2. SRS Schema using XML

The SRS elements with their elements and attributes are shown next. Fig. 3. Show the History element. Fig. 4. And Fig. 5 Show the Author element and the introduction element, respectively.

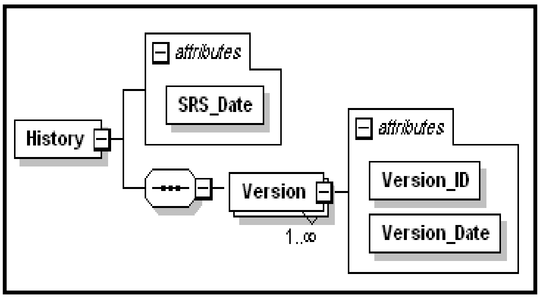


Fig. 3. History element of SRS

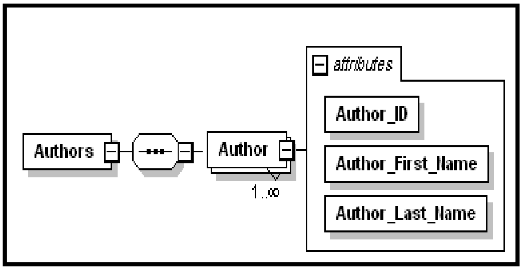


Fig. 4. Author element of SRS

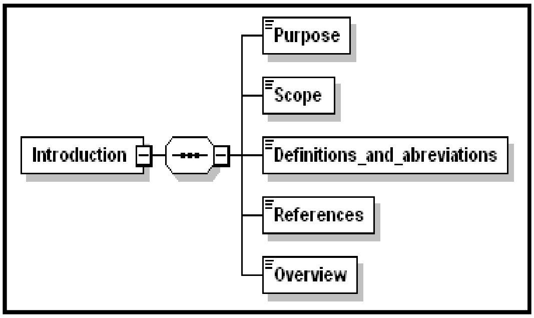


Fig. 5. Introduction element of SRS

Then, the main section that contains the actual requirement which is the specific requirement appears and shown in Fig. 6.

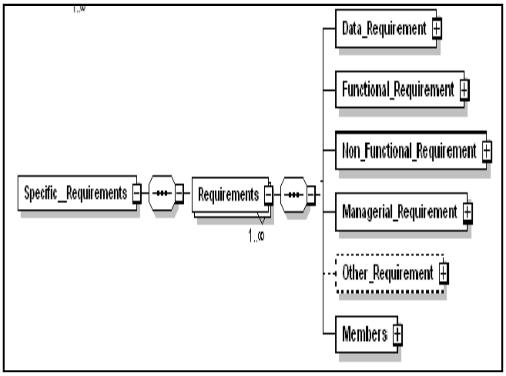


Fig. 6. Specific requirement element

Data Requirement, functional requirement, non-functional requirement, managerial requirement and other requirement are elements of the specific requirement element of SRS. They have many nested elements and attributes shown in Fig. 7, Fig. 8, Fig. 9, Fig. 10, Fig.11 and Fig. 12, respectively.

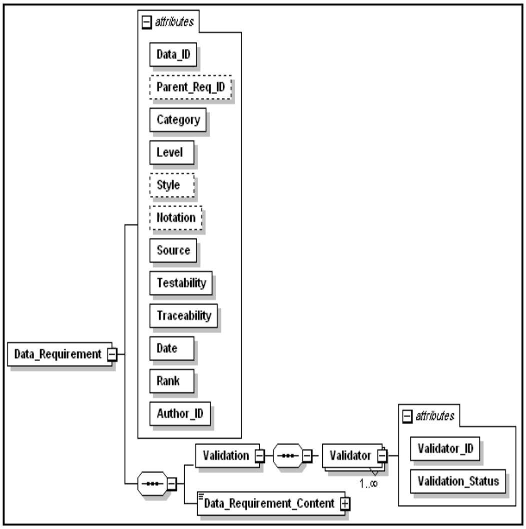


Fig. 7. Data Requirement element

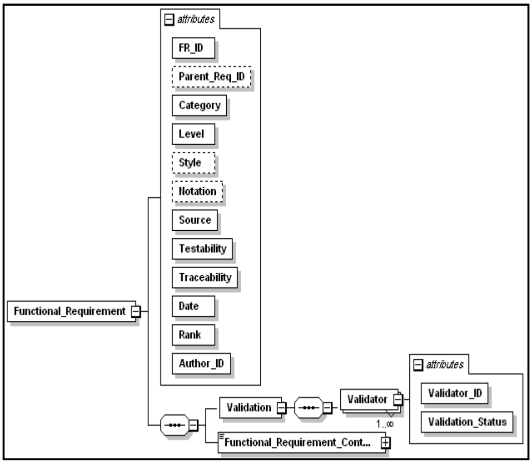


Fig. 8. Functional Requirement element

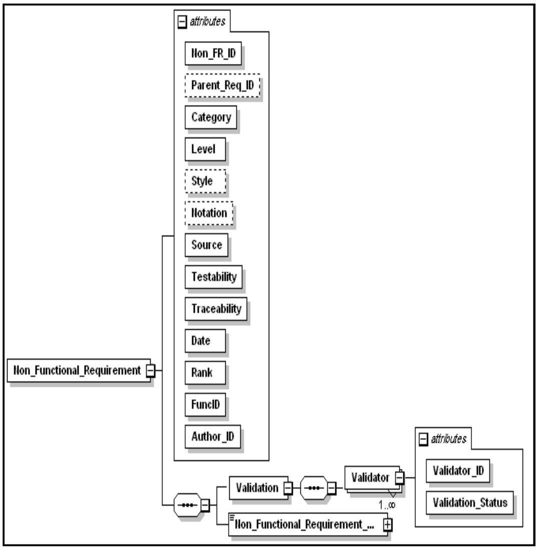


Fig. 9. Non-functional Requirement element

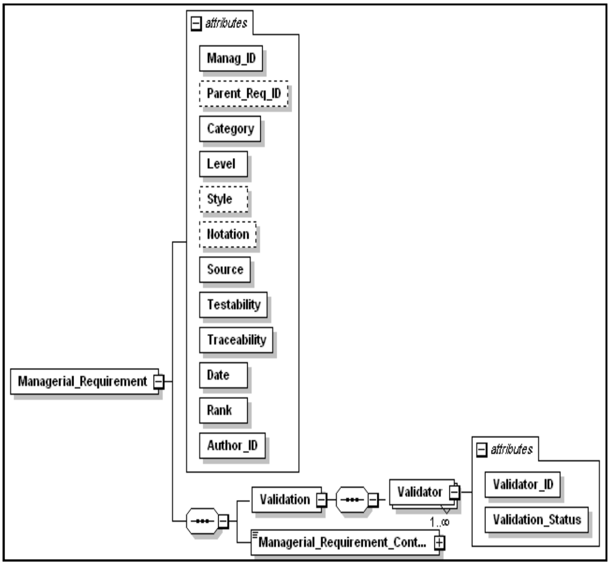


Fig. 10. Managerial Requirement element

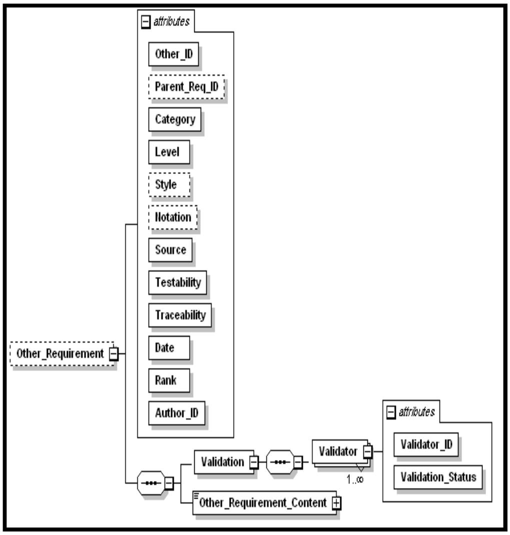


Fig. 11. Other Requirement element

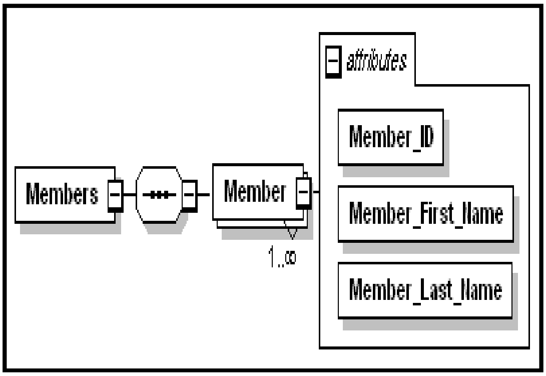


Fig. 12. Members element of Specific requirement of SRS

After that, validation and appendices elements are shown in Fig. 13 and Fig. 14.

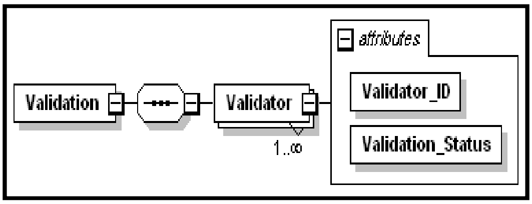


Fig. 13. Validation element of SRS

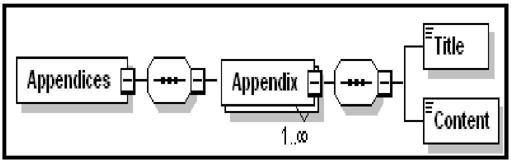


Fig. 14. Appendices element of SRS

3.3 Performance

There are ten main criteria any SRS should satisfy to consider a high-quality SRS; this is according to NASA's Software Assurance Technology Center [12]. Table 1 shows how any SRS written based on our schema met these ten criteria.

Table 1. How Schema met the ten criteria

Criteria	How it satisfied in our model
Complete	Every SRS written according our schema is automatically complete because the schema itself is complete. It defines all possible situations any SRS may have and clarifies the appropriate responses to each of these situations essentially the mandatory parts.
Consistent	The consistency of an SRS instance is ensured by the consistency of the schema itself. All system functions requirements are compatible with performance requirements. This compatibility ensured by the Non_Functional Requirement attributes Fun_ID. This attribute links each nonfunctional requirement with a functional requirement that should be compatible with.
Correct	Our schema divides the SRS to small detailed elements and each element is complete and precise. Moreover, some elements have attributes which raise the level of details. Since information written in the schema is very detailed, this will make the check operation easy to be done. Therefore, ensuring the correctness is simpler and definite. In addition, our schema has cross reference feature. For example, Author_ID and validator_ID of any requirement will reference Member_ID. This feature will ensure the correctness. For example, if a wrong Author_ID of a requirement entered, an error will happen because there is no correct reference of that Author_ID in Member_ID.
Modifiable	SRS is easy to be modified since it is divided to distinct parts. This ease the reach of the part intended to change. Since SRS written in a semi-structured format, it is easy and forward to maintain.

Table 1. (*continued*)

Ranked	The rank of requirement specification statements is specified by its importance. Our schema ensures that each requirement is ranked through the attribute Rank. Rank value could be high, medium or low.
Testable	The XML Schema makes a check upon each requirement to ensure that it is testable. This check is done in our schema for each requirement through Testability attribute. This attribute shows whether the requirement is tested before or not.
Traceable	Our schema ensures traceability for each requirement through traceability attribute.
Unambiguous	Statements of requirements are unambiguous. This is because of the allowance of using explicit knowledge regarding the requirement categories, styles, levels, etc.
Valid	Our schema offers validation element for each requirement. This element contains Validator_ID and validation status. Validation_ID referred to validator name and validation_status which has “agree” “fair” or “disagree” values.
Verifiable	Our schema ensures verification of SRS by having Parent_Req_ID as an attribute for each requirement. This attribute makes a reference to the requirement it is composed of.

3.4 Prototype

In this section a software prototype was made to just give a view on how the XML Schema works. This prototype was designed by Visual Basic 2008 Express Edition. Fig.15 shows screen shots of the SRS prototype.

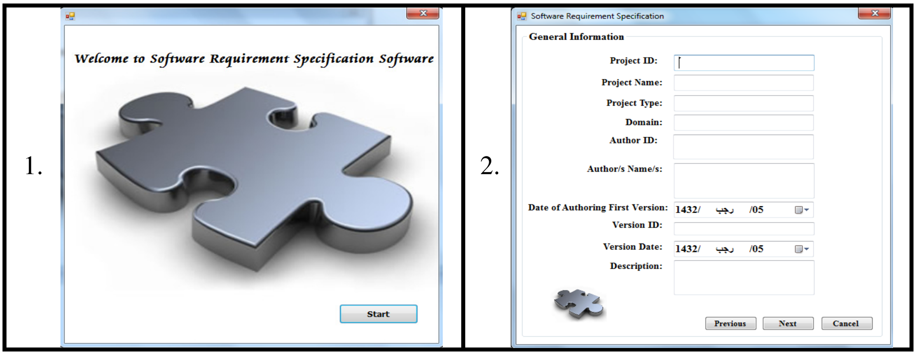


Fig. 15. SRS Prototype

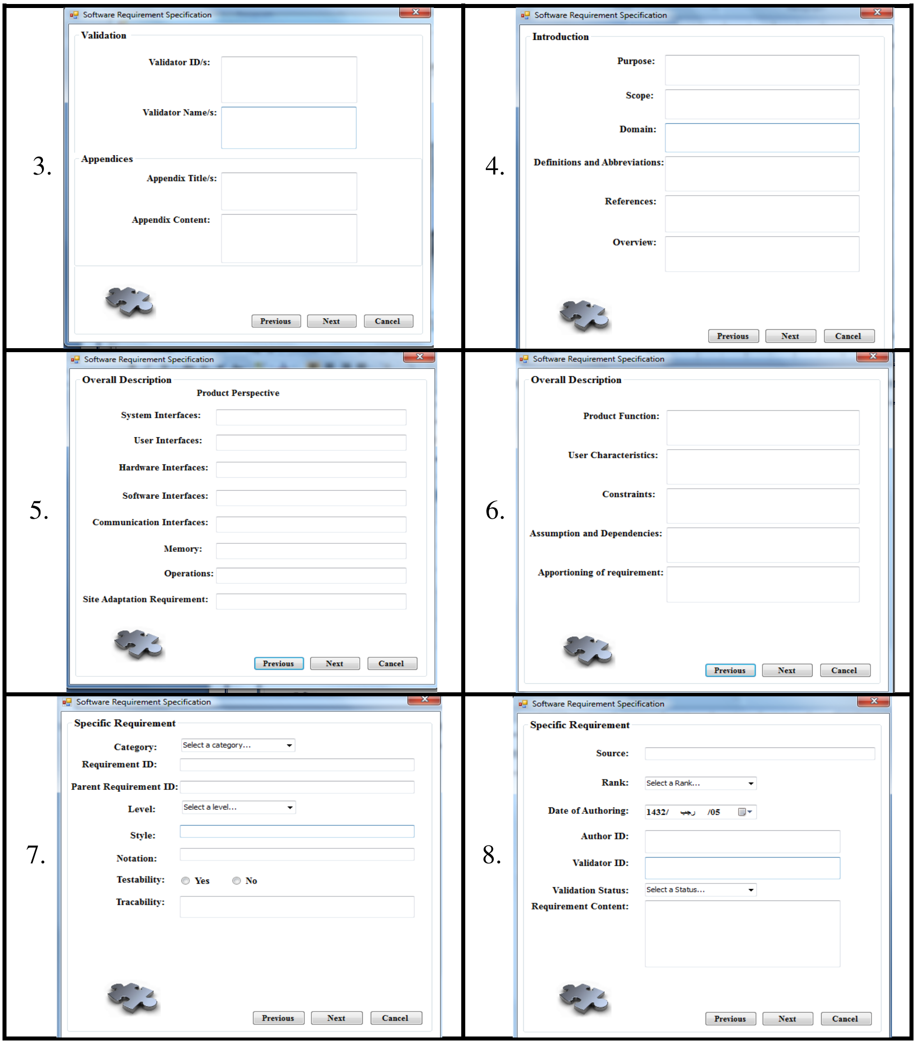


Fig. 15. (*continued*)

3.5 Case Study

A case study is selected to be the method for testing the research work. It was chosen to be the Library Management system of the University of Ballarat. A complete SRS of library system found in [13] was written using our designed schema. An example of a functional Requirement of the library SRS is shown in Fig. 16.

FR_ID:	7870	Parent_Req_ ID:	7845	Level:	Product
Category:	Functional_ Requirement	Style:	Feature Require ment	Notation:	-
Source:	LSS	Testability:	Yes	Traceability:	7845
Date:	20-May-2009	Rank:	high	Author_ID:	003
Validator_ ID:	002	Validation_ Status:	Agree		
Functional requirement Content:		The system search shall be based on various categories viz. book name or the ISBN.			

Fig. 16. An example of a functional requirement

4 Conclusion and Future Work

An overview has been given about software requirement field. The report identifies the importance of documentations and its standard. A focus has been made of the familiar SRS Standards which is IEEE 830 Standards document. An XML Schema was developed to be used as a general template for writing SRS of any software. The effectiveness of the schema was judged against the ten main criteria. In addition, a case study was applied to our schema.

In our XML Schema, the metadata language used is not definitive; it could be modified, improved or changed. So, there are some recommendations that will complete and help to improve XML Schema. Some of them are:

The proposed semi-structured of IEEE 830 SRS used in this document is at a general level. We can consider our schema as a generic schema and it needs others to complete it. Our schema needs to be integrated with other schemas like, for example, many styles of data and functional requirement need other schema to show how they are written. For example, data dictionary, data expression, context diagrams, use cases, dataflow diagram or others need a separate schema to describe how they are look like. A communication between stakeholders in SRS can be permitted by adding metadata that shows stakeholders' opinions and discussions on the SRS.

References

1. Sommerville, I.: An Introduction to software Engineering: ©Ian Sommerville (2004), `http://www.scribd.com/doc/21379604/Software-Engineering-Sommerville`
2. Surhone, L.M., Tennoe, M.T., Hens, S.F.: Software Requirements Specification. VDM Verlag Dr. Mueller AG & Co. Kg. (2010)
3. Jordan, W.: How to Write a Software Requirements Specifications (SRS) Document, `http://www.ehow.com/how_4804549_software-requirements-specifications-srs-document.html`
4. Besprozvannykh, A.V.: Extensible Markup Language (XML): Essentials for Climatologists

5. Guerrieri, E.: Software Document Reuse with XML. In: Proceedings of the Fifth International Conference Software Reuse, pp. 246–254 (1998)
6. Kyong-Ho, L., Yoon-Chul, C., Sung-Bae, C.: Logical Structure Analysis and Generation for Structured Documents: A Syntactic Approach. IEEE Transactions on Knowledge and Data Engineering 15(5), 1277–1294 (2003)
7. Ishitani, Y.: Document Transformation System from Papers to XML Data Based on Pivot XML Document Method. In: Proceedings of the Seventh International Conference on Document Analysis and Recognition, pp. 250–255 (2003)
8. Nojoumian, M.: Document Engineering of complex software specifications (2007)
9. Le Vie Jr., D.: Writing Software Requirements Specifications, http://www.techwr-l.com/techwhirl/magazine/writing/softwarerequirementspecs.html
10. IEEESocieties, IEEE Recommended Practice for Software Requirements SpeciÞcations. © 1998 by the Institute of Electrical and Electronics Engineers, Inc., New York (1998)
11. Lauesen, S.: Software requirements: styles and techniques. © Pearson Education Limited, Great Britain (2002)
12. Akinci, U.: 10 Characteristics of High-Quality SRS (Software Requirements Specifications) (2010), http://www.technicalcommunicationcenter.com/2010/07/19/10-characteristics-of-high-quality-srs-software-requirements-specifications/
13. Patil, K.B.: Software Requirements specification for Library Management System (July 2009)

ClassIN: A Class Inheritance Metric Tool

B. Ramachandra Reddy and Aparajita Ojha

PDPM Indian Institute of Information Technology, Design and Manufacturing Jabalpur, Dumna Airport Road, P.O. Khamaria, Jabalpur, 482005, India
{brreddy,aojha}@iiitdmj.ac.in

Abstract. This paper presents a comprehensive class inheritance metrics tool called ClassIN. The tool works for Java projects and presents an analysis with twenty different inheritance metrics including metrics at class level as well as at class hierarchy level. This also helps in identifying class hierarchies that may be more complex from the point of view of software maintenance. Graphical visualization of three important metrics, namely AID, specialization ratio and reuse ratio are also provided for an insight on structure of class hierarchies of a given Java project. This would help developers in identifying classes that may be more prone to faults or high maintenance costs.

Keywords: Class Inheritance Hierarchies, Inheritance Metrics, Software Maintainability.

1 Introduction

Inheritance patterns in object oriented software systems greatly affect the overall performance and maintainability of systems. Numerous metrics have been proposed by researchers and developers that measure different aspects of inheritance present in a software. Depth of inheritance (DIT), number of children (NOC) [1], average inheritance depth (AID), specialization ratio (SR) and reuse ratio (R) [2] metrics are some of the most important metrics in quantifying the effect of inheritance. Use of inheritance metrics helps in identifying possible attributes that may reduce maintenance efforts and enhance the reliability, Maintainability [27],[29],[30]and Reusability [3], Fault prediction, Defect Prediction [4],[5],[6],[28],Testability [7],[8].

This paper presents a comprehensive class inheritance metrics tool called ClassIN. The tool works for Java projects and presents an analysis with twenty different inheritance metrics including metrics at class level as well as at class hierarchy level. This also helps in identifying class hierarchies that may be more complex from the point of view of software maintenance. The main features of ClassIN are as follows –

- It provides inheritance metrics of the project at class level and also at class hierarchy level.
- It provides some insight into the depth and breadth of class hierarchies.
- It exports the values of inheritance metrics of a Java project into a spreadsheet (Excel-sheet) for analysis of results.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_11, © Springer International Publishing Switzerland 2014

- It displays all the class inheritance hierarchies of the project.
- Provides a 3D visualization of three important class hierarchy metrics for a Java project, namely AID, Specialization ratio and Reuse Ratio.
- Provides an insight into the maintainability (Modifiability and Understandability) of class hierarchies.

A comparative analysis of the proposed tool with the existing tools is also presented here. Our tool is available online with a demonstration.

2 Inheritance Metrics

Inheritance metrics are used to measure the depth, width and relative inheritance values reflecting the inheritance patterns in an object oriented system. Inheritance metrics are broadly classified into two types - class level inheritance metrics and class hierarchy metrics. The class level inheritance metrics represent the inheritance values of individual classes, whereas the class hierarchy metrics represent inheritance hierarchal structures of the related classes. Table 1 lists metrics that are commonly used for determining class level inheritance and class hierarchy level inheritance. In addition to inheritance metrics the ClassIN tool also provides values of maintainability metrics such as Average Modifiability (AM) and average understandability (AU) of class hierarchies [3].

Table 1. Inheritance Metrics

Class level metrics	Class hierarchy metrics
Depth of Inheritance (DIT) [1]	Maximum DIT (MaxDIT) [11]
Number of Children(NOC) [1]	Average Inheritance Depth (AID) [2]
Total Progeny count (TPC) [10]	Number of children for a component (NOCC) [11]
Total Ascendancy count (TAC) [10]	Total length of inheritance chain (TLI) [10]
Class-to-leaf depth (CLD) [13]	Specialization Ratio (S) [2]
Number of Ancestor classes (NAC) [14]	Reuse Ratio (U) [2]
Number of Descendent classes (NDC) [14]	Attribute Inheritance Factor (AIF) [9]
Number of Overridden Methods (NORM) [12]	Method Inheritance Factor (MIF) [9]
Number of Attributes Inherited(NAI) [12]	Specialization Index(SIX) [12]
Number of Methods Inherited(NMI) [12]	
Coupling Through Inheritance(CTI) [15]	

3 Analysis of Software Metrics Tools

Several commercial as well as open-source OO metric tools exist today. We have analyzed CKJM [16], Analyst4J [17], Eclipse plug-in 1.3.6 [18], JMT [19], VizzAnalyzer [20], Dependency Finder [21], OOMeter [22], SD metrics [23]. Table 2 shows comparative analysis of various existing tools.

Table 2. Tools and inheritance metrics used in evaluation

Metrics		TOOLS						
	CKJM	**Analyst4J**	**Eclipse plug-in 1.3.6**	**JMT**	**Vizz-Analyzer**	**Dependency Finder**	**OO-Meter**	**SD Metrics**
DIT	√	√	√	√	√	√	√	√
NOC	√	√	√	√	√	√	√	√
AID	×	√	√	×	×	×	×	×
AIF	×	×	×	√	×	×	×	×
MIF	×	×	×	√	×	×	×	×
MaxDIT	×	√	√	×	×	×	×	×
CLD	×	×	×	×	×	×	×	×
TLI	×	×	×	×	×	×	×	√
TPC	×	×	×	×	×	×	×	√
TAC	×	×	×	×	×	×	×	√
NORM	×	×	√	×	×	×	×	×
NAI	×	×	×	√	×	×	×	√
NMI	×	×	×	√	×	×	×	√
SIX	×	×	√	×	×	×	×	×
NOCC	×	×	×	×	×	×	×	×
S	×	×	×	×	×	×	×	×
U	×	×	×	×	×	×	×	×
CTI	×	×	×	×	×	×	×	×
Class hierarchies	×	×	×	×	×	×	×	×

The above table gives a clear picture of various metrics covered in some of the standard tools. One can see that DIT and NOC are covered by all the tools. Further, JMT tool also provides AIF, MIF metrics at project level whereas Analyst4J, Eclipse plug-in cover AID, MaxDIT metrics also. Whereas Analyst4J provides metrics at project level, Eclipse plug-in provides metrics at package level. SD Metrics covers maximum number of class-level metrics whereas JMT provides a mix of class level as well as class hierarchy level metrics. So there is no tool covering metrics at class hierarchy level. A comprehensive study of metrics at class hierarchy level helps determine factors that would result in better maintainability and reusability of the software. In view of this, we have developed a comprehensive tool ClassIN that covers all the metrics listed in Table 1 along with two more metrics useful to predict the maintainability through modifiability and understandability [3].

In general, designers prefer to keep the depth of inheritance low in class hierarchies, in order to improve their understandability and reusability [25]. ClassIN tool is useful for measuring depth and breadth of the class inheritance hierarchies. As it is well known, DIT, NOC metrics are two most useful measures for prediction of fault proneness and software reusability [26]. As reported in [24], higher DIT value indicates higher maintenance cost [24]. Especially for DIT values ≥ 5, software becomes highly complex from maintenance point of view. Prechelt et al. [26] have also analyzed maintainability with respect to DIT values and have concluded that modules with low inheritance depth are easier to maintain. The DIT metric gives individual class depth in the hierarchy. The NAI, NMI, DIT, NOC metrics are useful for finding the number of test cases required for determining the correctness of a software system [7], [8]. Harrison et al. [9] had concluded that MOOD metrics such as AIF, MIF provide an overall quality assessment of systems. In addition to AIF and MIF metrics, ClassIN provides two more measures, namely Specialization ratio (S) and Reuse ratio (U) to help developers assess more effectively a given software project with respect to reusability and testability. In essence, ClassIN tool provides a variety of metrics that may be used for different purposes. Metrics may be selectively used for analyzing the software performance, fault-proneness, reusability and maintainability.

4 ClassIN Tool

In this section, we shall describe the basic functionalities of ClassIN applied to a usage scenario. Initially ClassIN takes a Java project as input and after analyzing various metrics it displays the highest value of metrics in a mainframe window. The tool also generates all the class hierarchies available in the project. The tool helps in identifying various attributes of a software project. These attributes in turn help in determining if a given project needs a design review. Figure 1 displays a snapshot of ClassIN showing the results for system level inheritance metrics. After generating the metric data, the tool exports all the metrics in an Excel sheet. Next, the tool displays all the class hierarchies in the project. A snapshot of hierarchies displayed by the tool is presented in Figure 2.

In order to show the functionality of the tool and its usefulness, we considered the problem of lack of discrimination in class hierarchies. As shown in Table 1, numerous metrics have been proposed by researchers to depict and analyze the inheritance structure in class hierarchies. However, there is no standard set of metrics that helps

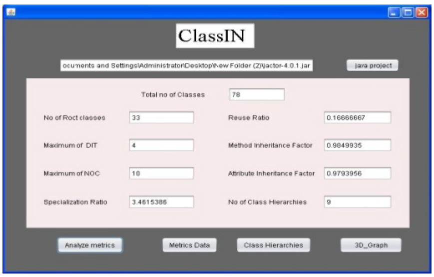


Fig. 1. Snapshot of ClassIN tool

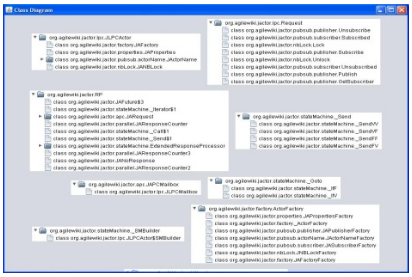


Fig. 2. Snapshot of class hierarchies

in distinguishing between inheritance patterns. In fact, very different hierarchical structures lead to the same values of some standard inheritance metrics, resulting in lack of discrimination anomaly. This prevents the developers in effectively analyzing class hierarchies for maintainability, testability and reusability of class hierarchies. As a case study, we proposed a vector valued measure DIPV= (AID, S, U) for discriminating class hierarchies. Using ClassIN Visualization module, DIPV is plotted for different class hierarchies in Figure 3. AID provides some insight into the inheritance levels in a given class hierarchy, whereas the specialization ratio S and the reuse ratio R give the idea of the breadth of the hierarchy and the depth of reuse. Thus, the triple gives a fair idea of inheritance structure of a class hierarchy. Two DIPV vectors are compared as follows. If $u = (x_1, y_1, z_1)$ and $v = (x_2, y_2, z_2)$ are two vectors then $u > v$ if and only if one of the following conditions hold (i) $x_1 > x_2$ (ii) $x_1 = x_2$ and $y_1 < y_2$ (iii) $x_1 = x_2$, $y_1 = y_2$ and $z_1 > z_2$. Otherwise the two DIPV vectors are equal. Low DIPV value of class hierarchy indicates low maintainability, high reusability and better testability of the software. The tool also helps in suggesting various measures for analyzing modifiability and understandability using its visualization module. Figure 4 shows a snapshot of the output of visuliazation module where modifiability and understanblity metric values of certain class hierarchies are plotted agaisnt the class numbers in the project. The first class hierarchy has highest maintainability vallue among the all the class hierarchies. The graph suggests the designers should reconsider the class hierarchy structure.

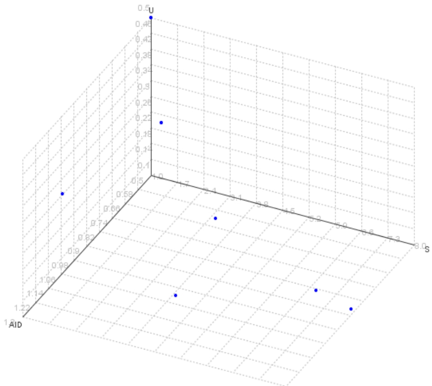


Fig. 3. DIPV values of class hierarchies

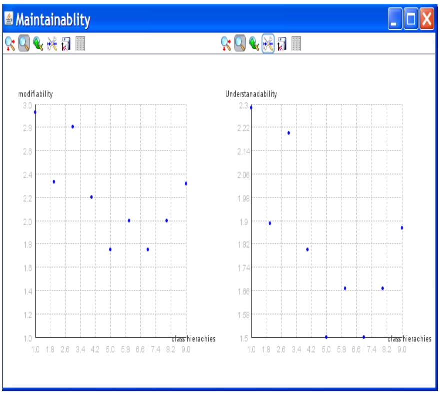


Fig. 4. Maintainability of class hierarchies

5 Tool Architecture

The architecture of ClassIN tool is presented in Figure 5. The ClassIN tool is decomposed into three modules. The first module takes a java project as an input and finds total number of classes in the project using reflection classes. In the second module, tool finds inheritance relations between classes. In the third phase, tool

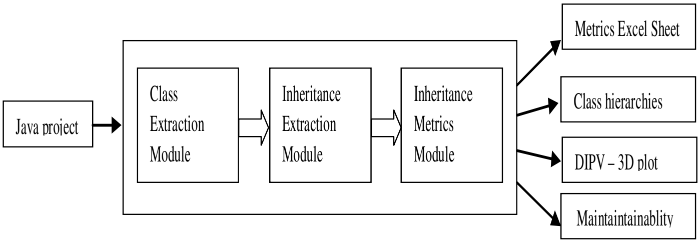


Fig. 5. ClassIN tool architecture

measures all the inheritance metrics for the project at class level as well as class hierarchy level. The tool generates four outputs namely metrics excel file, display of class hierarchies in the project, DIPV plot and graphs for analyzing maintainability of class inheritance hierarchies. Maintainability values for each class hierarchy are also displayed to ascertain class structures with higher maintenance costs. This also helps in discrimination of different class inheritance patterns in the project.

6 Tool Availability

The tool along with user guide and technical documentation, may be freely downloaded from its webpage at

https://sites.google.com/site/brcreddyse/Tools

References

1. Chidamber, S.R., Kemerer, C.F.: A Metrics Suite for Object Oriented Design. J. IEEE Trans. Soft. Eng. 20(6), 476–493 (1994)
2. Henderson-Sellers, B.: Object Oriented Metrics: Measures of Complexity, pp. 130–132. Prentice-Hall (1996)
3. Sheldon, F.T., Jerath, K., Chung, H.: Metrics for Maintainability of Class Inheritance Hierarchies. J. Soft. Main. and Evol. Res. and Pra. 14(3), 147–160 (2002)
4. Basili, V.R., Briand, L.C., Melo, L.W.: A Validation of Object-Oriented Design Metrics as Quality Indicators. J. IEEE Trans. Soft. Eng. 22(10), 751–761 (1996)
5. Subramanyan, R., Krisnan, M.S.: Empirical Analysis of CK Metrics for Object-Oriented Design Complexity: Implications for Software Defects. J. IEEE Trans. Soft. Eng. 29(4), 297–310 (2003)
6. Cartwright, M., Shepperd, M.J.: An Empirical Investigation of an Object-Oriented Software System. J. IEEE Trans. Soft. Eng. 26(8), 786–796 (2000)
7. Bruntink, M., Deursen, A.V.: An Empirical Study into Class Testability. J. Sys. and Soft. 79, 1219–1232 (2006)
8. Baudry, B., Traon, Y.L.: Measuring Design Testability of a UML Class Diagram. J. Info. and Soft. Tech. 47, 859–879 (2005)

9. Harrison, R., Counsell, S.J.: An Evaluation of the Mood set of Object-Oriented Software Metrics. J. IEEE Trans. Soft. Eng. 21(12), 929–944 (1995)
10. Abreu, F.B., Carapuca, R.: Candidate Metrics for Object-Oriented Software within a Taxonomy Framework. J. Sys. and Soft. 26, 87–96 (1994)
11. Vernazza, T., Granatella, G., Succi, G., Benedicenti, L., Mintchev, M.: Defining Metrics for Software Components. In: 5th World Multi-Conference on Systemics, Cybernetics and Informatics, Florida, vol. XI, pp. 16–23 (2000)
12. Lorenz, M., Kidd, J.: Object-Oriented Software Metrics. Prentice Hall (1994)
13. Tegarden, D.P., Sheetz, S.D., Monarchi, D.E.: A Software Complexity Model of Object-Oriented Systems. J. Dec. Sup. Sys. 13, 241–262 (1995)
14. Li, W.: Another Metric Suite for Object-Oriented Programming. J. Sys. and Soft. 44, 155–162 (1998)
15. AlGhamdi, J., Elish, M., Ahemed, M.: A Tool for measuring Inheritance Coupling in Object Oriented Systems. J. Info. Sci. 140, 217–227 (2002)
16. CKJM metric tool, `http://www.spinellis.gr/sw/ckjm/`
17. Analyst4J metric tool, `http://www.codeswat.com/cswat/index.php`
18. Eclipse plug-in 1.3.6 tool, `http://www.sourceforge.net/projects/metrics/`
19. JMT tool, `http://www-ivs.cs.uni-magdeburg.de/sw-eng/agruppe/forschung/tools/`
20. VizzAnalyzer tool, `http://www.arisa.se`
21. Dependency Finder tool, `http://www.depfind.sourceforge.net/`
22. Alghamdi, J., Rufai, R., Khan, S.: OOMeter: A Software Quality Assurance Tool. In: 9th European Conference on Software Maintenance and Reengineering, pp. 190–191. IEEE Computer Society, Manchester (2005)
23. SD Metrics tool, `http://www.sdmetrics.com/`
24. Daly, J., Brooks, A., Miller, J., Roper, M., Wood, M.: Evaluating Inheritance Depth on the Maintainability of Object-Oriented Software. J. Emp. Soft. Eng. 1, 109–132 (1996)
25. Genero, M., Piattini, M., Calero, C.: A survey of metrics for UML class diagrams. J. Obj. Tech. 4(9), 59–92 (2005)
26. Prechelt, L., Unger, B., Philippsen, M., Tichy, W.: A Controlled Experiment on Inheritance Depth as a Cost Factor for Code Maintenance. J. Sys. and Soft. 65, 115–126 (2003)
27. Harrison, R., Counsell, S.J., Nithi, R.: Experimental Assessment of the Effect of Inheritance on the Maintainability of Object-Oriented Systems. J. Sys. and Soft. 52, 173–179 (2000)
28. Catal, C.: Software Fault Prediction: A literature review and current trends. J. Expert Sys. with App. 38, 4626–4636 (2011)
29. Chen, J., Huang, S.: An Empirical Analysis of the Impact of Software Development Problem Factors on Software Maintainability. J. Sys. and Soft. 82, 981–992 (2009)
30. Genero, M., Manso, E., Visaggio, A., Canfora, G., Piattini, M.: Building Measure-based Prediction Models for UML Class Diagram Maintainability. J. Emp. Soft. Eng. 12, 517–549 (2007)

Discrimination of Class Inheritance Hierarchies – A Vector Approach

B. Ramachandra Reddy and Aparajita Ojha

PDPM Indian Institute of Information Technology, Design and Manufacturing Jabalpur, Dumna Airport Road, P.O. Khamaria, Jabalpur, 482005, India
{brreddy,aojha}@iiitdmj.ac.in

Abstract. Numerous inheritance metrics have been proposed and studied in the literature with a view to understand the effect of inheritance on software performance and maintainability. These metrics are meant to depict the inheritance structures of classes and related issues. However, in spite of a large number of inheritance metrics introduced by researchers, there is no standard set of metrics that could discriminate the class hierarchies to decipher or predict the change-proneness, defect-proneness of classes or issues that could effectively address maintainability, testability and reusability of class hierarchies. In fact, very different hierarchical structures lead to the same values of some standard inheritance metrics, resulting in lack of discrimination anomaly (LDA). In an effort to address this problem, three specific metrics have been studied from the point of view of providing an insight into inheritance patterns present in the software systems and their effect on maintainability. Empirical analysis shows that different class hierarchies can be distinguished using the trio – average depth of inheritance, specialization ratio and reuse ratio.

Keywords: Inheritance Metrics, Class Hierarchies, Software Maintainability.

1 Introduction

Metrics play an important role in measuring the software quality, cost, complexity, maintenance efforts and many other important attributes affecting the overall software performance. If these issues are addressed and analyzed properly at design level, they lead to more efficient and cost effective software development and maintenance. Numerous metrics have been proposed over a period of two decades to assess different software attributes such as size, inheritance, cohesion, coupling, testing efforts, and maintainability. Metric suites that have been widely studied and applied successfully for object oriented software include CK metrics suite [1] and MOOD metrics suite [2]. However, there is no standard set of rules that helps a software developer or a project manager in selecting specific metrics that would be more useful for a particular development. Further, these metrics often provide overlapping information. But, the use of multiple metrics becomes inevitable for large and complex software systems. This, in its turn, generates large data sets, making it

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_12, © Springer International Publishing Switzerland 2014

extremely unmanageable to analyze and interpret. Experience suggests that better inferences can be drawn by using only a small set of metrics.

Inheritance patterns present in an object oriented software system greatly affect the overall performance and maintainability of the system. Numerous metrics have been proposed by researchers and developers that measure different aspects of inheritance present in a software. Depth of inheritance (DIT), number of children (NOC), average inheritance depth (AID), specialization ratio (SR) and reuse ratio (R) metrics are some of the most commonly used metrics in quantifying the effect of inheritance. Use of inheritance metrics has claimed to reduce the maintenance efforts and enhance the reliability [3]. In an interesting study, Daly et al [4] have concluded that software systems having 3-levels of inheritance are easier to maintain than 0-level inheritance, but deeper levels of inheritance adversely affect the maintainability.

People have performed various empirical studies on design measures for object oriented software development (see for example, ([7], [8], [9], [10], [11], [12], [14], [15], [19], [22], [23], [25] and [26]). Briand et al [12] have studied relationships between coupling, cohesion, inheritance and the probability of fault detection in software. Through an empirical analysis, they conclude that accurate models can be built to predict fault-prone classes. They also show that in general, the frequency of method invocations and the depth of inheritance hierarchies are the prime factors leading to fault-proneness in the system, apart from class size. Cartwright and Shepperd [11] investigated the effect of inheritance on a specific industrial object oriented system and found that classes in an inheritance structures were approximately three times more defect-prone compared to those classes that did not participate in inheritance structure. Further, they were able to build useful prediction systems for size and number of defects based upon simple counts such as number of events or number of states per class. Similar studies were also made by Harrison et. al and Wi Li [16, 4] with a view to understand the inheritance effects on modifiability and maintainability of object oriented systems. Dallal [14] has applied the concept of class flattening to inherited classes for improving the internal quality of attributes such as size, cohesion and coupling. Elish [15] has empirically evaluated that DIT and NOC are good metrics for finding the fault tolerance, maintainability and reusability in aspect oriented systems.

Aggarwal et. al [10] have investigated twenty two different object oriented metrics proposed by various researchers and have made an empirical study of some of the selected metrics that provide sufficient information for interpretation. Metrics providing overlapping information are excluded from the set. Mishra [7] has recently introduced two inheritance complexity metrics, namely Class Complexity due to Inheritance (CCI) and Average Complexity of a program due to Inheritance (ACI). These metrics are claimed to represent the complexity due to inheritance in a more efficient way. More recently, Makkar et. al [19] have proposed an inheritance metric based on reusability of UML based software design. They have presented an empirical analysis of the proposed metric against existing reusability based inheritance metrics.

With the growing complexities of inheritance relationships and polymorphism in large object-oriented software systems, it is becoming increasingly important to concentrate on measures that capture the dynamic behavior of the system. Han [20] et. al have proposed a Behavioral Dependency Measure (BDM) for improving the accuracy over existing metrics when predicting change-prone classes. With the help of a case study, they demonstrate that BDM is a complementary indicator for change-proneness when the system contains complex inheritance relationships and associated polymorphism. Results of the case study show that the BDM can be used for model-based change-proneness prediction.

In spite of a large number of inheritance metrics introduced by researchers, there is no standard set of metrics that could discriminate the class hierarchies to decipher or predict the change-proneness, defect-proneness of classes or issues that could effectively address maintainability of the system and reusability of class hierarchies. Dallal [13] has recently made an interesting study on cohesion metrics and has demonstrated that most of the metrics reflect the same cohesion values for classes having same number of methods and attributes but distinct connectivity patterns of cohesive interactions (CPCI). This results in incorrect interpretation of the degrees of cohesion of various classes. This is termed as lack of discrimination anomaly (LDA) problem. To resolve the problem, Dallal et. al [13] have proposed a discrimination metric and a simulation-based methodology to measure the discriminative power of cohesion metrics. The proposed metric measures the probability that a given cohesion metric will generate distinct cohesion values for classes having same number of methods and attributes but different CPCIs. Motivated by the work, we have made an empirical analysis of various available class inheritance hierarchy metrics. A tool "ClassIN" is developed to study the inheritance patterns inbuilt in Java projects. Like the cohesion metrics, inheritance metrics also suffer from LDA problem. This contribution focuses on identifying distinctive features in class hierarchies of a software system having same inheritance metric values. In this paper, we propose a vector valued metric 'Discrimination of Inheritance Pattern Vector' (DIPV), to resolve the discrimination anomaly to some extent. This measure proves to be quite useful in improving the understandability of class hierarchies present in a software system. The metric captures the distinction that a particular hierarchy possesses from others. The three measures that we have chosen for defining DIPV are average depth of inheritance, specialization ratio and reuse ratio. This trio gives a good insight of the class hierarchies and related issues such as testing efforts and maintainability.

The paper is organized as follows. Section 2 presents an overview of some of the well-known inheritance metrics. The problem of lack of discrimination anomaly for inheritance metrics is also presented in this section. In Section 3, a new approach for discriminating the inheritance pattern in a software system is presented using the trio named DIPV (a vector valued metric). An Empirical analysis is also presented in this section using ClassIN tool. The tool helps in identifying inheritance patterns present in a software system with special emphasis on the software attributes such as depth and breadth of class hierarchies. This facilitates in visualizing the general inheritance pattern present in the system.

2 Inheritance Metrics and Lack of Discrimination Anomaly

Inheritance metrics measure the depth, width and relative inheritance values reflecting the inheritance patterns in an object oriented system. Inheritance metrics are broadly classified into two types - class level inheritance metrics and class hierarchy metrics. The Class level inheritance metrics represents the inheritance values of individual class, whereas the class hierarchy metrics represents hierarchal structure of the related classes. Table 1 lists metrics that are commonly used for determining class level inheritance and class hierarchy level inheritance.

Table 1. Inheritance metrics

Inheritance Metrics	**Description**
a. Class Metrics	
Depth of Inheritance (DIT)	Maximum length from the node to root.
Number of Children(NOC)	Number of immediate descendents of the class.
Number of Ancestor classes (NAC)	Total number of super classes of the class.
Number of Descendent classes (NDC)	Total number of subclasses of the class.
Total Children Count(TCC)	Number of Subclasses of the class.
Total progeny Count(TPC)	Number of classes that inherit directly or indirectly from a class.
Total Parents count (TPAC)	Number of super-classes from which the given class inherits directly.
Number of Methods Inherited (NMI)	Number of methods inherited by the class.
Number of Attributes Inherited (NAI)	Number of attributes inherited by the class.
Class-to-leaf depth (CLD)	The maximum length of the path from a class to a leaf.
b. Class Hierarchy Metrics	
Maximum DIT (MaxDIT)	Maximum of the DIT values obtained for each class of the class hierarchy.
Average Inheritance Depth (AID)	Sum of depths of classes/Total number of classes.
Number of children for a component (NOCC)	Number of children of all the classes in the component.
Total length of inheritance chain (TLI)	Total number of edges in an inheritance hierarchy graph (number of classes inherited)
Specialization Ratio (S)	Ratio of number of subclasses to the number of super classes.
Reuse Ratio (U)	Ratio of number of super classes to the total number of classes.
Method Inheritance Factor (MIF)	Ratio of number of inherited methods in a class to the number of visible methods in a class.
Attribute Inheritance Factor (AIF)	Ratio of number of inherited attributes in a class to the number of visible attributes in a class.

DIT and NOC are most commonly used class level metrics. While these metrics do give some idea on the inheritance complexity of a system, they do not discriminate between different hierarchical patterns. AIF, MIF metrics do provide some insight into the internal structure of a hierarchy. Specialization ratio and reuse ratio are commonly used for discriminating the class hierarchies. However, in many instances, these metric values turn out to be the same for relatively very different inheritance patterns. Although class hierarchies can be understood to some extent using specialization ratio, reuse ratio and average inheritance depth, none of the measures alone is capable of providing a good understanding of the inherent hierarchies. In what follows, we shall present some specific cases that exhibit the LDA in class hierarchy inheritance metrics.

Consider the metric MaxDIT defined by

$$MaxDIT = \max_{C_i \in K} \{ DIT(C_i) \}$$

where C_i is a class in a class hierarchy K. Note that MaxDIT [21] for all the three hierarchies in Figure 1 turns out to be 2, but the three hierarchies have different structures. MaxDIT does not indicate any inheritance pattern and the relative overriding complexities. Moving onto Average Depth of Inheritance (AID), it is defined by

$$AID = \frac{\sum_i DIT(C_i)}{TC}$$

where TC denotes total number of classes in the hierarchy. AID is a ratio of sum of DIT values to the total number of classes. It may be noted that AID value for Figure 1(a) is 1.5 and is 1.33 for hierarchies shown in Figure 1(b) and 1(c). Hence this also does not give any clear picture of the hierarchical structure present in a module. Turning to the metric NOCC [21], all the class hierarchies shown in Figures 1 have the same value 5.

Thus, the possibility of discriminating class hierarchies using NOCC is very low. Similarly, the TLI metric (total length of inheritance chain) value for all the hierarchies of Figure 1 is 5. The Specialization Ratio S (Table 1) gives some idea about the width of class hierarchy, higher the specialization ratio, wider would be the class hierarchy. The hierarchy having higher value of S is considered to be better than that having value close to 1. The specialization ratio for hierarchies of Figure 1(a) and 1(b) is the same and is equal to 2.5, while for Figure 1(c) it turns out to be 1.67.

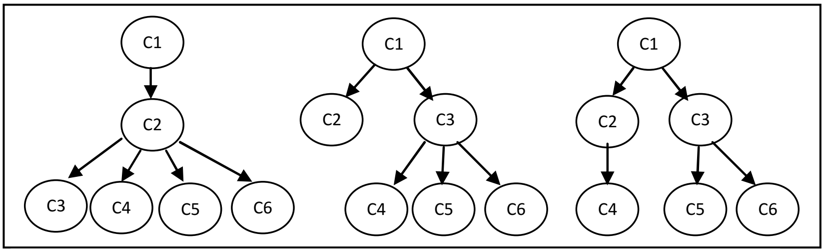


Fig. 1. Class hierarchies

Another related ratio that helps in discriminating different hierarchies is the Reuse Ratio U (Table 1). The reuse ratio is always less than 1. The higher reuse ratio reflects that the system is having deep hierarchy and high reuse value. The reuse ratio for hierarchies of Figure 1(a) and 1(b) is 0.33, 0.33, while the same for 1(c) turns out to be 0.5.

An empirical study was performed to analyze the performance of different class hierarchy metrics using six different projects from sourceforge.net [24]. These projects are Bloat (12) JActor (9), Oxygene (24), Pandora Project Management (19), SGLJ (23) and OpenNLP (41). Numbers in the brackets indicate the number of class hierarchies present in each project. Thus a total number of 128 class hierarchies were taken to analyze the discriminating power of inheritance metrics. Six metrics, namely MaxDIT, NOCC, TLI, AID, S, and U were chosen to examine the discriminating power.

Table 2. Statistics of class hierarchy metrics for 128 classes of six different projects

Metrics	**Min**	**Max**	**Mean**	**Median**	**Standard Deviation**	**Percentage of distinct metric values (%)**
Max DIT	1	5	1.48	1	0.87	3.12
NOCC	1	80	6.9	3	12.63	19.53
TLI	1	80	6.92	3	12.63	19.53
AID	0.5	2.61	1.12	1	0.4	29.68
S	1	69	4.05	2	8.03	24.21
U	0.01	1	0.48	0.36	0.31	20.31

3 Class Hierarchy Metric: A Vector Approach

Let us define the Discrimination of Inheritance Pattern Vector (DIPV) as a triple using the combination of AID, specialization ratio and reuse ratio as follows.

$$\text{DIPV} = (\text{AID, S, U})$$

AID provides a good understanding of the inheritance levels present in a given class hierarchy. But, this alone is insufficient in giving details of inheritance pattern. Combining it with specialization ratio gives the idea of the breadth of the hierarchy, whereas the reuse ratio also gives the information about the depth. Thus DIPV gives a fair amount of indication on the maintenance efforts needed. Comparison of two DIPV is performed as follows. If $u = (x_1, y_1, z_1)$ and $v = (x_2, y_2, z_2)$ are two vectors then $u > v$ if and only if one of the following conditions hold (i) $x_1 > x_2$ (ii) $x_1 = x_2$ and $y_1 < y_2$ (iii) $x_1 = x_2$, $y_1 = y_2$ and $z_1 > z_2$. Otherwise $u <= v$.

To ascertain better maintainability, we propose the following algorithm.

Algorithm

Input: DIPV1 = (x_1, y_1, z_1), DIPV2= (x_2, y_2, z_2)
Output: Lower DIPV (lower maintainability)
Begin:

1. If $x_1 > x_2$ then
2. return DIPV2.
3. else if $x_1 = x_2$ and $y_1 < y_2$ then
4. return DIPV2..
5. else if $x_1 = x_2$ and $y_1 = y_2$ and $z_1 > z_2$
6. return DIPV2.
7. else
8. return DIPV1.
9. end if.
10. end.

Table 3. LDA cases in some class hierarchies

Class Hierarchy	MaxDIT	NOCC	AID	S	U	AM	AU
1	2	3	1	1.5	0.5	2.5	2
2	2	3	1.25	1.5	0.5	2.875	2.25
3	2	4	1.2	1.33	0.6	2.8	2.2
4	2	4	1.2	2	0.4	2.8	2.2
5	2	5	1.5	2.5	0.333	2.5	3.25
6	2	5	1.333	2.5	0.333	2.33	3
7	2	10	1.54	3.33	0.27	3.31	2.54
8	2	10	1.63	5	0.18	3.45	2.63
9	3	6	2.14	2	0.42	4.21	3.14
10	3	6	1.85	1.5	0.57	3.78	2.85
11	3	8	1.88	1.33	0.66	3.83	2.88
12	3	8	1.55	2	0.44	3.33	2.55
13	3	11	1.5	2.75	0.33	3.25	2.5
14	3	11	1.75	2.2	0.41	3.625	2.75
15	4	16	2.117	2.28	0.41	4.17	3.11
16	4	16	1.82	3.2	0.29	3.73	2.82

In order to develop some understanding on the class hierarchy metrics and their revelation about the inheritance structures, a tool "ClassIN" has been developed in the present study. The tool aims at finding out various inheritance metrics for Java projects. We have considered only abstract classes and concrete classes of the projects in the tool. Using the tool, AID is compared first, lower value gives lower maintainability and hence, we do not go further to compare the second and third components. If both AID values are the same then the second and third attributes of DIPVs are compared to ascertain the breadth and depth of inheritance patterns. Higher specialization and lower reuse ratio indicate a better design from the maintenance point of view.

Table 3 shows some of the LDA cases of class hierarchies taken the six projects mentioned in Section 2. In the first and second case, values of MaxDIT and NOCC are 2 and 3 respectively for both the class hierarchies. Applying the algorithm one gets the first DIPV. The resultant hierarchy gives lower maintenance cost, low testability and better reusability. This can also be asserted by looking at the values of AM and AU, which are precisely the maintainability metrics. In some cases all the values turn out to be the same. It indicates that both the hierarches have same maintainability cost.

3.1 Validation of DIPV

After applying DIPV, the algorithm returns either DIPV1 or DIPV2. For the validation of DIPV results, statistical discriminant analysis test was performed. The resultant discriminate function is

$$D_i = -7.146 + 1.788 * AID + 0.583 * S + 5.536 * U.$$

The discriminate function coefficients are positively correlated with discriminant function. So, AID, S and U are suitable for discrimination of class inheritance hierarchies. More than 61% grouped cases are correctly classified in all the class hierarchies present in the six mentioned projects from sourceforge.net [24].

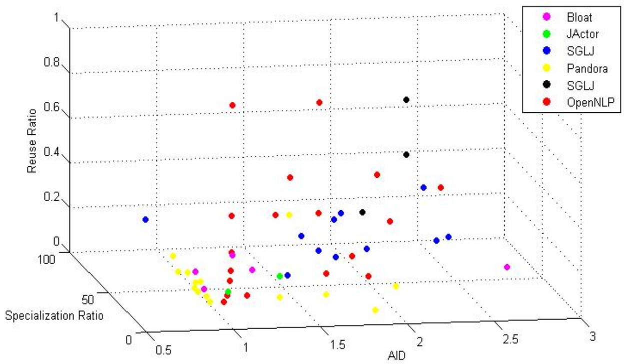


Fig. 2. A visualization of DIPVs of class hierarchies

Figure 2 shows a plot of DIPV for a total of 128 class hierarchies taken from six projects from sourceforge.net as mentioned in Section 2. For the project Pandora, one may observe that specialization ratio of one of the class hierarchies is around 70, while its reuse ratio is well within the limit (0.5) and AID is close to 2. Such hierarchies may lead to higher maintenance and testing efforts and should be reviewed at the design level only. Moving onto the class hierarchies of the Project Oxygen, all the class hierarchies have specialization ratio within the interval range of [1, 4], highest reuse ratio is 0.6 and the highest AID is 2.25. Thus the project is well designed; maintenance and testing efforts would be within the manageable limits. Thus, in general, the triple gives a good understanding of the hierarchical pattern and helps in visualizing the blueprint of the design for analysis and refinement, if needed. However, it cannot be claimed that in all cases, DIPV will always give a good insight on maintainability, reusability or testability. Nonetheless, it may be treated as a primary measure to identify class hierarchies that may create issues in the project maintenance.

4 Conclusion

Different inheritance metrics are analyzed for discriminating the class hierarchies to understand their effect on maintenance efforts. The metrics MaxDIT, AID, reuse ratio, specialization ratio can be used for discriminating the hierarchies. However, a single metric does not suffice to give any insight on the inheritance pattern. To enhance the understanding on the inheritance pattern and related software attributes such as average depth and breadth of hierarchical structure and their overriding complexities a vector DIPV has been proposed in this paper. It helps in providing a better picture of the blueprint of the inheritance pattern present in a software system. Accordingly, the design could be reviewed from the point of view of maintenance, testing efforts and reusability. A tool ClassIN is developed to visualize the inheritance patterns present in a software system. This helps in quick analysis of the software design and the inheritance patterns used in the system. The tool can be downloaded from https://sites.google.com/site/brcreddyse/Tools

References

1. Chidamber, S.R., Kemerer, C.F.: A Metrics Suite for Object Oriented Design. J. IEEE Trans. Soft. Eng. 20(6), 476–493 (1994)
2. Harrison, R., Counsell, S.J.: An Evaluation of the Mood set of Object-Oriented Software Metrics. J. IEEE Trans. Soft. Eng. 21(12), 929–944 (1995)
3. Sheldon, F.T., Jerath, K., Chung, H.: Metrics for Maintainability of Class Inheritance Hierarchies. J. Soft. Main. and Evol. Res. and Pra. 14(3), 147–160 (2002)
4. Daly, J., Brooks, A., Miller, J., Roper, M., Wood, M.: Evaluating Inheritance Depth on the Maintainability of Object-Oriented Software. J. Emp. Soft. Eng. 1, 109–132 (1996)
5. McCabe, T.J.: A Complexity Measure. J. IEEE Trans. Soft. Eng. 2(4), 308–320 (1976)
6. Abreu, F.B., Carapuca, R.: Candidate Metrics for Object-Oriented Software within a Taxonomy Framework. J. Sys. and Soft. 26, 87–96 (1994)

7. Mishra, D.: New Inheritance Complexity Metrics for Object-Oriented Software Systems: An Evaluation with Weyuker's Properties. J. Comp. and Info. 30(2), 267–293 (2011)
8. Henderson-Sellers, B.: Object Oriented Metrics: Measures of Complexity, pp. 130–132. Prentice-Hall (1996)
9. Li, W.: Another Metric Suite for Object-Oriented Programming. J. Sys. and Soft. 44, 155–162 (1998)
10. Aggarwal, K.K., Singh, K.A., Malhotra, R.: Empirical Study of Object-Oriented Metrics. J. Obj. Tech. 5(8), 149–173 (2006)
11. Cartwright, M., Shepperd, M.J.: An Empirical Investigation of an Object-Oriented Software System. J. IEEE Trans. Soft. Eng. 26(8), 786–796 (2000)
12. Basili, V.R., Briand, L.C., Melo, L.W.: A Validation of Object-Oriented Design Metrics as Quality Indicators. J. IEEE Trans. Soft. Eng. 22(10), 751–761 (1996)
13. Dallal, J.A.: Measuring the Discriminative Power of Object-Oriented Class Cohesion Metrics. J. IEEE Trans. Soft. Eng. 37(6), 788–804 (2011)
14. Dallal, J.A.: The impact of Inheritance on the internal Quality Attributes of Java Classes. Kuw. J. Sci. and Eng. 39(2A), 131–154 (2012)
15. Elish, M.O., AL-Khiaty, M.A., Alshayeb, M.: An Exploratory case study of Aspect-Oriented Metrics for Fault Proneness, Content and fixing Effort Prediction. Inter. J. Qua. and Rel. Mana. 30(1), 80–96 (2013)
16. Harrison, R., Counsell, S.J., Nithi, R.: Experimental Assessment of the Effect of Inheritance on the Maintainability of Object-Oriented Systems. J. Sys. and Soft. 52, 173–179 (2000)
17. Li, W., Henry, S.: Object-Oriented Metrics that Predict Maintainability. J. Sys. and Soft. 23(2), 111–122 (1993)
18. Zhang, L., Xie, D.: Comments On the applicability of Weyuker's Property Nine to Object-Oriented Structural Inheritance Complexity Metrics. J. IEEE Trans. Soft. Eng. 28(5), 526–527 (2002)
19. Makker, G., Chhabra, J.K., Challa, R.K.: Object Oriented Inheritance Metric-Reusability Perspective. In: International conference on Computing, Electronics and Electrical Technologies, pp. 852–859 (2012)
20. Han, A., Jeon, S., Bae, D., Hong, J.: Measuring Behavioral Dependency for Improving Change-Proneness Prediction in UML-based Design Models. J. of Sys. and Soft. 83(2), 222–234 (2010)
21. Vernazza, T., Granatella, G., Succi, G., Benedicenti, L., Mintchev, M.: Defining Metrics for Software Components. In: 5th World Multi-Conference on Systemics, Cybernetics and Informatics, Florida, vol. XI, pp. 16–23 (2000)
22. Basili, V.R., Briand, L.C., Melo, L.W.: How Reuse Influences Productivity in Object-Oriented System. Commun. ACM 39(10), 104–116 (1996)
23. Briand, L.C., Wst, J., Daly, J.W., Porter, D.V.: Exploring the Relationships between Design Measures and Software Quality in Object-Oriented Systems. J. Sys. and Soft. 51(3), 245–273 (2000)
24. Java Projects, http://www.sourceforge.net
25. Radjenovic, D., Hericko, M., Torkar, R., Zivkovic, A.: Software fault prediction metrics: A systematic literature review. J. Inf. and Soft. Tech. 55, 1397–1418 (2013)
26. Zhou, Y., Yang, Y., Xu, B., Leung, H., Zhou, X.: Source code size estimation approaches for object-oriented systems from UML class diagrams: A comparative study. J. Inf. and Soft. Tech. 56, 220–237 (2014)

Transformation of Coloured Petri Nets to UML 2 Diagrams

Ayman Yassin and Hoda Hassan

Computer Science, Arab Academy for Science and Technology, Cairo, Egypt
Ayman.Yassin@Gmail.com, huda.hassan@aast.edu

Abstract. Business Process Modeling Notation is used by business modelers to model business processes logic and artifacts. However, it is inadequate in expressing the execution semantics of business processes and takes a process-oriented approach for modeling systems. UML, on the other hand, is known for its expressiveness to present the object-oriented approach for modeling software-based system. There is a rising need to transform business process models to flawless UML models. This paper proposes a modeling transformation technique for transforming a business process-modeling notation model to different UML diagrams, using Coloured Petri Nets (CPN) as a formal intermediate step to ensure flawless transformation. This transformation would allow modeler to take advantages of the presentation power in BPMN as well as the implementation power in UML. Furthermore, this step will bridge the gap between the different modeling notations previously mentioned.

Keywords: Business Process Modeling Notation, Coloured Petri Nets, UML, Transformation, Use Cases Diagram, Activity Diagram.

1 Introduction

BPMN is one of the most popular Business Processes modeling language nowadays. BPMN is more suitable for business process modeling due to its high presentation power [1, 2]. With A Business Process Modeling Diagram, modelers can simulate the business artifacts to know the bottlenecks in the system being visualized, as well as identify the metrics of the system under development [3, 4].

BPMN is a Process Centric Approach that lacks the execution semantics required for process modeling implementation [2, 4]. In that respect, UML, which is an object oriented modeling language, is more suitable and preferred by software engineers [2, 4]. Yet, most of the stakeholders come from a non-technical background, and are ALIEN towards UML notations and the different OOP constraints that the language expresses. Furthermore, the UML model generated based on the BPMN process model is usually validated at a late stage by the business modelers leading to late detection of flaws, which results in more effort and cost to correct [5].

Therefore, there is a pressing need to bridge the gap between the modeling tools used by business modelers (BPMN) and that used by the software engineers (UML),

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_13, © Springer International Publishing Switzerland 2014

while ensuring flawless transformation from one tool to the other. Defining formal semantics for business process modeling notation (BPMN) using Coloured Petri Nets (CPN) helps the modeling tools vendors to automatically validate the business model. Further, it helps business analysts simulate the business process behavior to enable detection of flaws [5].

In this paper, we propose a model to transform a formal CPN model transformed and mapped from BPMN model to different common UML diagrams as shown in Fig. 1. This would allow modeler to be able to take advantages of both the presentation powers in BPMN and the implementation power in UML. Accordingly, business modelers will only have to focus on their expertise in process modeling, while software engineers will only have to focus on implementation tasks. Furthermore, it will help to bridge the gap between the formal model represented by CPN and the semi-formal model represented by UML. The transformation will be based on one-to-one mapping. Our focus in this paper will be on the mapping of CPN to use cases and activity diagrams, which are considered two of the common dynamic UML models.

1.1 Coloured Petri Nets

According to [5] Coloured Petri Net (CPN) is a graphical language for constructing models of concurrent systems and analyzing their properties. CPN is a discrete-event modeling language combining Petri Nets and the functional programming language CPN ML that is based on Standard ML. Standard ML provides the primitives for the definition of data types, describing data manipulation, and for creating compact and parameterized models. CPN is the most well known kind of high-level Petri Nets. CPN incorporates both data structuring and hierarchical decomposition - without compromising the qualities of the original Petri Nets.

A CPN model consists of data, places, transitions and arcs. Data are defined as data types, data objects and variables that hold data values. A place is a location for holding data. A transition is an activity that transforms data. An arc connects a place to a transition to specify data flow paths. A CPN model uses data types, data objects and variables. A CPN data type is called a color set such as integer, real, string, and Boolean that is available in a computer language [6, 9].

1.2 UML

Unified Modeling Language (UML) is a standardized general purpose modeling language in the field of object-oriented software engineering. UML is maintained and developed by the object management group (OMG). UML is widely adopted and considered as de facto standard for software modeling and design. Thus, UML provides us with an integrated way to specify the problems, facts, requirements, or software structure, using a set of expressive UML diagrams. Since recent large-scale software development focuses more on modeling than on programming, modeling is one of the most important key activities for successful software projects [4]. UML diagrams can be divided into two groups, namely static or structural diagrams and dynamic or behavioral diagrams.

Use cases diagram is one of the most common UML behavioral diagrams. Use cases diagrams are a means for specifying required usages of a system. Typically, they are used to capture the requirements of a system. The key concepts associated with use cases diagram are actors, use cases, and the subject. The users and any other systems that may interact with the subject are represented as actors. The required behavior of the subject is specified by one or more use cases, which are defined according to the needs of actors. An instance of use case refers to an occurrence of the emergent behavior that conforms to the corresponding use case type. Such instances are often described by interaction specifications [7].

UML provides mechanisms for reusing and adding on to use cases and actors. Actor capabilities can be expanded or can replace entire use cases using generalization. Use cases common elements can be factored by using included use cases, or can be added on to base use cases using use case extension [8].

UML 2 activity diagrams are important structured visual modeling notations useful for describing different types of behavior found in computer and information systems.

UML 2 activities are suitable for modeling the diverse requirements of many traditional scenarios. Activities provide for visual modeling that can be easily understood. UML 2 activities are based on Petri Net like semantics [11].

An activity diagram is a directed graph that consists of actions and flows. The basic structure of an activity model is similar to that of a Petri net, however there are supplementary model components provided in UML, which CPN does not provide. Those components include initial/final nodes, fork/join nodes, and decision/merge nodes. In addition to those components, there could exist expansion regions in UML, which manipulate collections of input data [4].

The choice of using UML diagrams as the target model in this paper is motivated by the fact that UML is one of the most widely used industry-standard modeling tools for software development. Furthermore, the proposed model attempts to bridge the gap between the modeling tool used by the business modelers (BPMN) and that used by the software engineers (UML), with the early detection of the business process flaws using the semantic model of CPN as shown in Fig. 1. The mapping of BPMN model into its corresponding CPN model was proposed in [5].

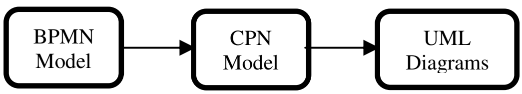


Fig. 1. The transformation process from BPMN to UML

The rest of this paper is organized as follows. Section 2 presents the mapping from CPN to UML diagrams. Examples of such mapping are shown in section 3. Finally, Section 4 and Section 5 discuss the related work and the conclusion respectively.

2 The Mapping of CPN to UML Diagrams

The proposed mapping process aims to extract the use cases diagram and the activity diagram. In particular, activity diagrams can be used to describe the dynamics of the system while use cases diagram can be used to describe the functionality of the system components. This will produce formal and structured use cases that can be used directly in the use case driven approach [10].

2.1 Mapping CPN to Use Cases Diagrams

In UML, use case diagram shows actors, use cases and their dependency relationships. In this paper, all relationships between the actor and the use case and between different use cases have been considered. A CPN model can be transformed to the use case diagram using the mapping process shown in Fig. 2 that should be applied in order and taking in consideration the following:

- As the transition of the CPN expresses a task, action or computation thus, it will be mapped to a use case and consequentially, CPN place will be mapped to an actor as shown in Fig. 3(a) and 3(b).
- The additional information of CPN (e.g. the transition guard and the arc expressions) will be mapped to the use case pre-conditions.
- The color function of an initial place, which is the place with no-inbound arcs, will be also mapped to the use case pre-condition.
- A final place is the place with no outbound arcs, and its color function represents the post-condition or the results of alternative steps within the use case.
- The predicate of the transitions will be also mapped to post-conditions.
- When a marked token is in a place and connected to a transition this means that the place is interacting with the transition. This place will be mapped to the caller actor and the trigger of the use case. The transition firing with this place means that the use case is ready for its task and will return to the caller actor after execution as shown in Fig. 3(c).
- The mapping from CPN to the use cases dependences (e.g. include and extend) based on the token functions, arc expressions, transition predicate and the multiple fusions of CPN. An include relationship is a relationship in which one use case (the base use case) includes the functionality of another use case (the inclusion use case). Include relationship can be extracted from CPN as shown in Fig. 3(d) such that P1, T2 represents use case (UC1) and P2, T2 represents use case (UC2). T1 output arc with its arc expression marking P2 with the required token and value, which are required for firing the transition T2. Respectively, an extend relationship is used to specify that one use case (extension) extends the behavior of another use case (base) as shown in 3(e) such that when UC1 is being executed, the execution of UC2 is optional. The condition of this selection is assigned to transition T1.

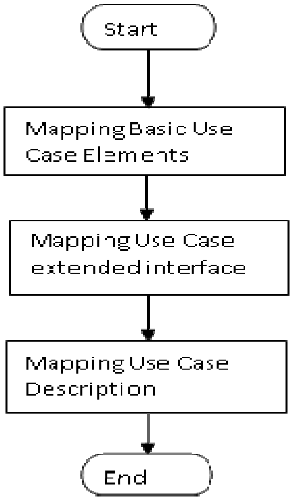


Fig. 2. The mapping Process of Use Case

Fig. 3(a). Mapping CPN Places

Fig. 3(b). Mapping CPN Transition

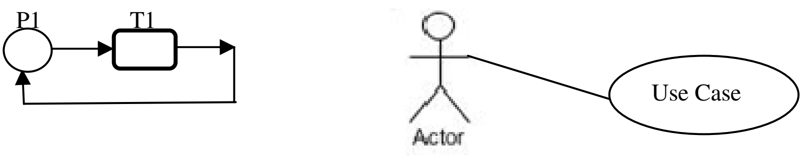


Fig. 3(c). Mapping CPN to Simple Use Case

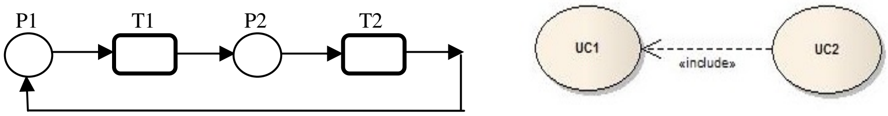


Fig. 3(d). Mapping CPN to Use Cases Include Relationship

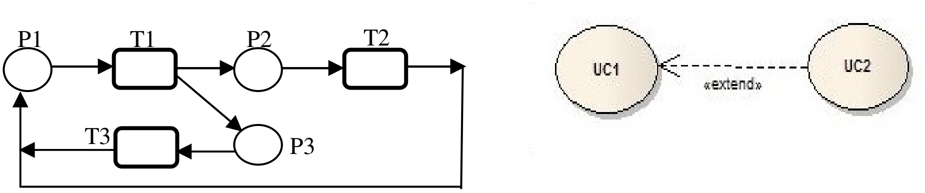


Fig. 3(e). Mapping CPN to Use Cases Extend Relationship

2.2 Mapping CPN to Activity Diagrams

In order to transform CPN to activity diagram (AD) the proposed rules and the idea presented by [12] is applied with a little adjustment while, using the reverse engineering of the rules. In all presented diagrams character P denotes places, character m denotes arc expression and character F denotes transition fusion. We cannot elaborate more description due to the limited paper size.

1) General CPN to Activity

In general, each CPN transition is mapped to an action. A place with a single inbound and a single outbound arcs is mapped to a control flow. The activity diagram presented in Fig. 4(a) represents an interaction without high-level operators. There are two Lifelines and one message between them. Each action represents activity state.

2) CPN to Activity (forming fork)

CPN transition branched to parallel transitions are mapped to correspondence AD parallel interaction as shown in Fig. 4(b). Fork node is a control node that splits a flow into multiple concurrent flows. A fork node has one incoming edge and multiple outgoing edges.

3) CPN to Activity (forming join)

CPN parallel transitions combined into one transition later are mapped to two parallel ADs combined later into one AD that denotes the end of parallel processing as shown in Fig. 4(c). Join node is a control node that synchronizes multiple flows with multiple incoming edges and one outgoing edge.

4) CPN to Activity(forming decision)

CPN with transition condition function is mapped to AD decision. AD decision node accepts tokens on an incoming edge and presents them to multiple outgoing edges. Which of the edges is actually traversed depends on the evaluation of the guards on the outgoing edges as shown in Fig. 4(d). The AD represented shows that if condition Action1, then do action 2, else do action 3. This signifies If-Else statement.

5) CPN to Activity (forming merge)

CPN transition with merging places is mapped as transition from two Parallel Activities to one Activity using merge as shown in Fig. 4(e).AD Merge node is a control node that brings together multiple alternate flows. It is not used to synchronize concurrent flows but to accept one among several alternate flows. A merge node has multiple incoming edges and a single outgoing edge.

6) CPN to AD Looping Transition

AD Looping Transition node is a structured activity node that represents a loop with setup, test, and body sections. Looping occurs in second condition of decision i.e. in Action3 indicates that While condition true do Action 3 as shown in Fig. 4(f).

7) CPN to AD Precedence Transition

AD Precedence means Action 1 should precede action 3 as shown in Fig. 4(g).

8) CPN to AD Timing Transition

As shown in Fig. 4(h) the result value contains the time at which the occurrence transpired. Such an action is informally called a wait time action.

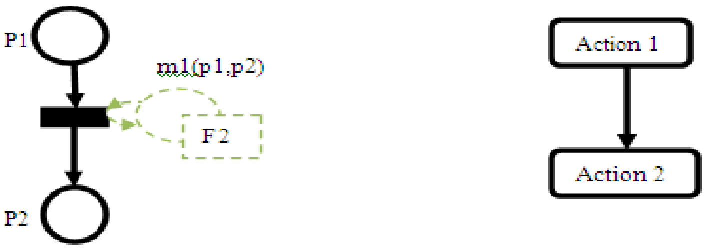


Fig. 4(a). General CPN to Activity

Fig. 4(b). CPN to Activity (forming fork)

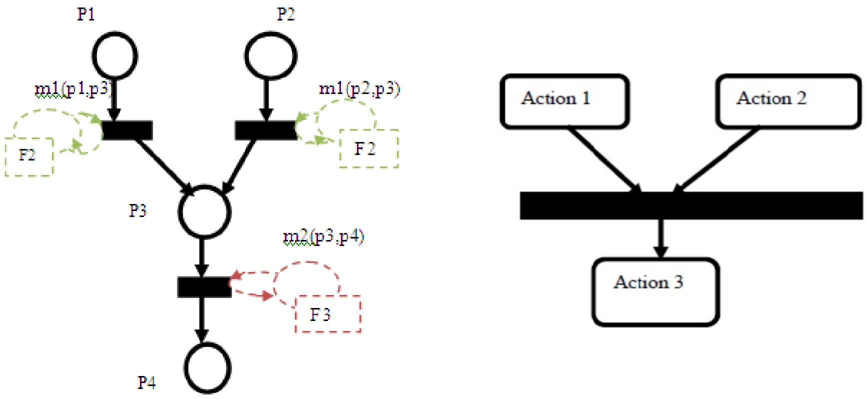


Fig. 4(c). CPN to Activity (forming join)

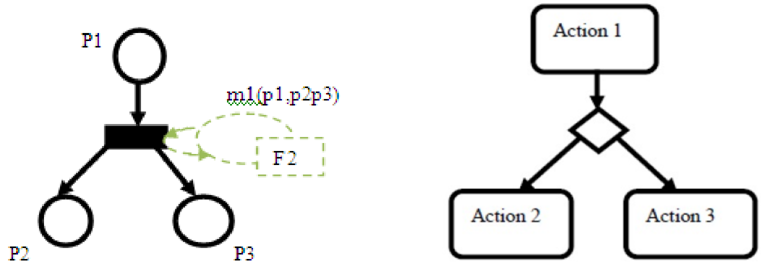


Fig. 4(d). CPN to Activity (forming decision)

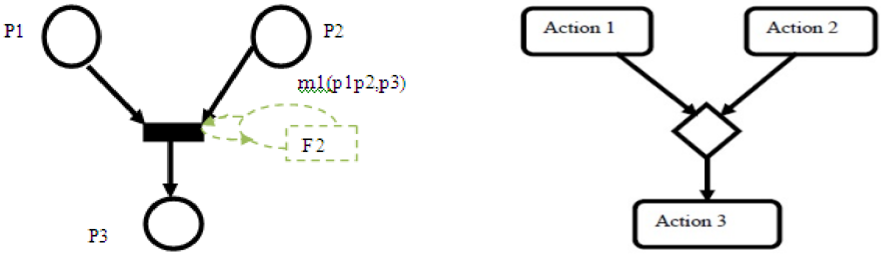


Fig. 4(e). CPN to Activity (forming merge)

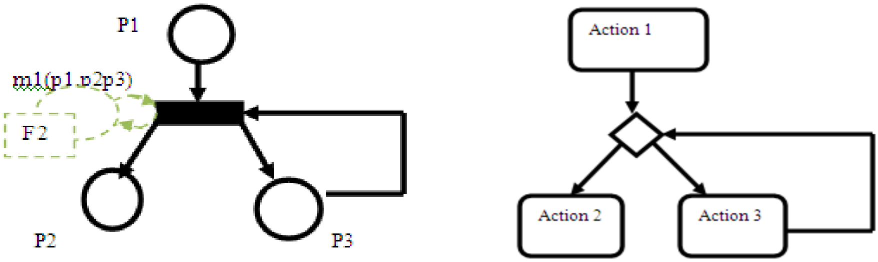


Fig. 4(f). CPN to AD Looping Transition

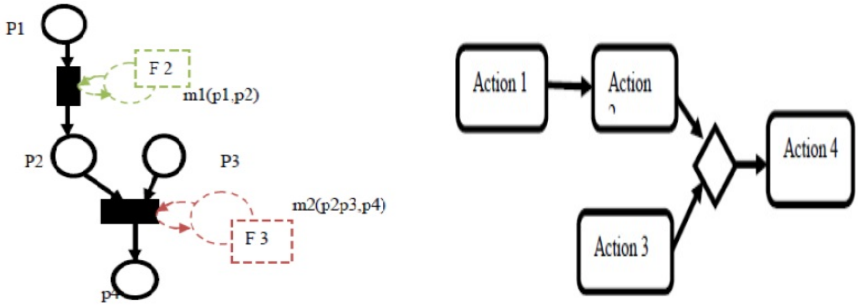


Fig. 4(g). CPN to AD Precedence Transition

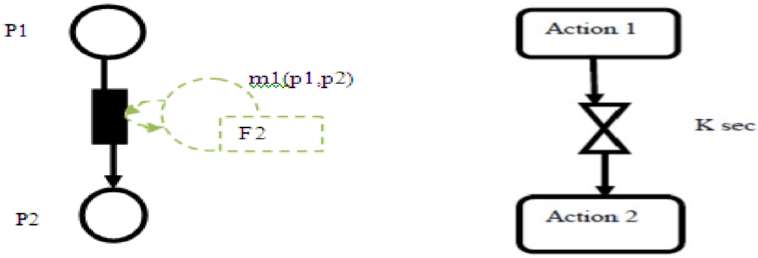


Fig. 4(h). CPN to AD Timing Transition

3 Example

The coloured petri net in Fig. 5 models a part of an automatic teller machine (ATM) with a bank at the backend. When the ATM is in the ready state, a client can ask for a certain amount of money. The ATM communicates this amount to the bank and waits for approval. When the approval arrives, the money is given to the client. Meanwhile, the requested amount is deducted from the client's account. If the account is deficient, approval will not be granted. In the model, a rejection message would be sent from the bank to the ATM, leading to an error message from the ATM to the client. Using the mapping from CPN to use case and activity diagrams described in the previous section, ATM System process in Fig. 5 is mapped to the use case diagram in Fig. 6 In addition, the activity diagram Fig. 7 is produced.

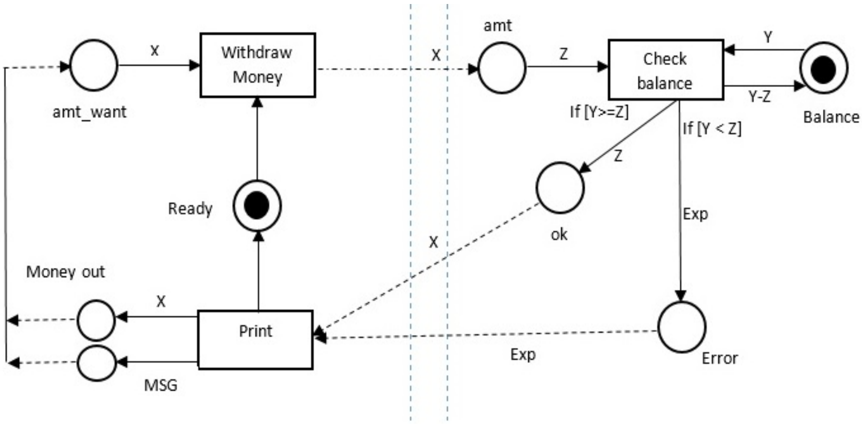


Fig. 5. CPN ATM Example

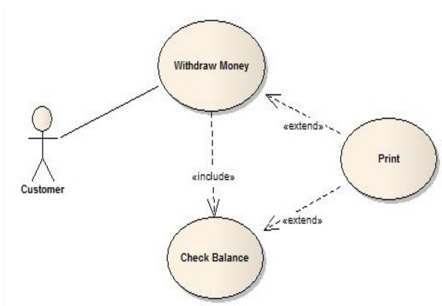


Fig. 6. Use Case for ATM Example

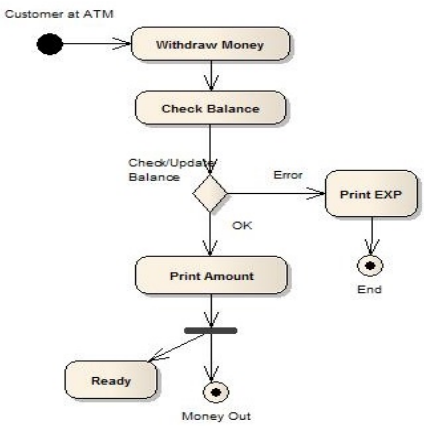


Fig. 7. Activity Diagram for ATM Example

4 Related Work

Based on our knowledge most of the researches and approaches are transforming and presenting different UML diagrams notations to Petri Nets or Coloured Petri Net (CPN) for simulation, verification and validation purposes depending on its formality and analysis capabilities in contrast to the work proposed in this paper. There are two mainstream approaches to transform UML constructs into CPN, which are informal approaches and formal approaches. Not all Approaches use CPN.

In [11], the authors present the transformation of UML 2.0 activity diagram into Petri Nets and Coloured Petri Nets. One of the formal mappings proposed was using the triple graph grammar (TGG). Although TGG can represent forward and backward transformation but the set of rules used in the backward transformation is not the same as those used in the forward one. Accordingly, in complex models the transformation presented will require the application of different set of rules recurrently.

In [4], the authors propose a complementary modeling process to UML modeling, which is used to keep the consistency between heterogeneous UML models based on CPN. The proposed informal transformation model is abstract as the focus was on preserving the consistency between UML models.

In [13], the authors propose a simple rule based bi-directional transformation of UML 2 activity diagram to Petri Net. In that transformation, the rules can be implemented in a generic informal manner or formally using triple graph grammar (TGG) notations. However, the idea presented was not validated through implementation. In addition, the proposed mapping is to Petri Net not to CPN. The backward transformation needs elaboration on the rules required.

In [14], the authors propose an algorithm to transform a software architecture described by use case, sequence and component diagrams into an executable model based on different extensions of Petri Nets. This work tries to fill the gap between software architect and non-functional requirement analyst. However, the idea presented in the mapping of use case diagrams is to Petri nets not CPN that yields to extra places and transitions in the Petri Nets. In addition, the idea did not present clearly how the Use case additional information (e.g. pre-conditions, post-conditions and trigger) are mapped to its corresponding Petri Nets.

5 Conclusion

This paper presents a formalization and a mapping process of Coloured Petri Nets (CPN) model to its corresponding UML models with respect to use case and activity diagrams. The proposed process is a step towards bridging the gap between different modeling notations, as that exists between BPMN 2.0 and UML 2. BPMN is used by business process modelers, and UML 2.0 is used by software engineers for further software development. The transformation proposed uses CPN for validation and simulation of the previously mentioned models. This helps vendors of modeling tools to validate the business process model automatically and generate the corresponding software UML model. Future work will aim to better tune the proposed mapping rules, transforming CPN to other common UML diagrams and build an automated tool for this transformation. Such automation should help to reduce the communication time between business process modelers and software engineers as well as, to synchronize business process with implementation seamlessly.

References

1. White, S.A.: Introduction to BPMN. IBM Corporation
2. Bào, N.Q.: A proposal for a method to translate BPMN model into UML activity diagram. Vietnamese-German University, BIS 2010 (2010), 2010,contact@nqbao.com
3. Birkmeier, D.: An Empirical Comparison of the Usability of BPMN and UML Activity Diagrams for Business Users. University of Augsburg, Germany, dominik.birkmeier@wiwi.uni-augsburg.de
4. Shinkawa, Y.: Inter-Model Consistency in UML Based on CPN Formalism. Faculty of Science and Technology, Ryukoku University, 1-5 Seta Oe-cho Yokotani, Otsu 520-2194, Japanshinkawa@rins.ryukoku.ac.jp
5. Ramadan, M.: BPMN Formalization using Coloured Petri Nets. Computer Science. Arab Academy for Science and Technology, Cairo, mohamed.e.ramadan@aast.edu
6. Shin, E., Levis, A.H., Wagenhals, L.W.: Transformation of UML-based System Model to Design/CPN Model for Validating System Behavior, Michael. Department of Computer Science, Texas Tech University, Lubbock, TX 79409-3104, USA, Michael.Shin@coe.ttu.edu
7. OMG Unified Modeling LanguageTM (OMG UML), Superstructure Version 2.4.1 (August 2011)
8. Pilone, D., Pitman, N.: UML 2.0 in a Nutshell A desktop Quick Reference. O'REILY (2005)
9. Jensen, K.: Coloured Petri Nets – Basic Concepts, Analysis Methods and Practical Use, vol. 1-3. Springer (1992, 1994, 1997)
10. Jacobson, I.: Object-Oriented Software Engineering, A Use Case Driven Approach. Addison-Wesley (1992)
11. Staines, A.S.: Intuitive Transformation of UML 2 Activities into Fundamental Modeling Concept Petri nets and Colored Petri Nets. University of Malta, Malta
12. Agarwal, B.: Some Rules to Transform Activity Diagrams into Colored Petri Nets. International Journal of Recent Technology and Engineering (IJRTE) I(5) (November 2012) ISSN: 2277-3878
13. Spiteri Staines, A.: Rule Based Bi-Directional Transformation of UML2 Activities into Petri Nets. International Journal of Computers 5(2) (2011)
14. Emadi, S., Shams, F.: A new executable model for software architecture based on Petri Net. Indian Journal of Science and Technology 2(9) (September 2009) ISSN: 0974- 6846; Computer Engineering Department, Science Research Branch, Islamic Azad University, Tehran, Iran, Computer Engineering Department, Shahid Beheshti University, GC, Tehran, Iran emadi@srbiau.ac.ir1; f_shams@sbu.ac.ir2

mEMR: A Secure Mobile Platform Integrated with Electronic Medical Records

Waqar Haque, Dan Horvat, and Leon Verhelst

University of Northern British Columbia, Prince George BC, Canada
{waqar.haque,horvat,verhelst}@unbc.ca

Abstract. A platform for secure communication between mobile devices and Electronic Medical Records (EMR) is proposed. This is an essential component of a shared care environment when multiple physicians/specialists are involved in providing care to a patient and consultation is necessary. Further, when physicians travel to provide care in sparsely populated regions, secure communication between their mobile devices and remote EMRs is necessary to securely obtain health information and push events recorded during the visit. The proposed solution provides an EMR-agnostic communication platform which adopts the standardized Clinical Document Architecture. The proposed solution has been demonstrated to work in both online and offline modes thereby making it useful not only in urban populations but also in remote areas where network connectivity may be less reliable.

Keywords: Mobile Platforms, Clinical Document Exchange, Secure Communications.

1 Introduction

The healthcare industry has benefitted greatly from recent advances in technology. Successful implementation of systems for management and internal exchange of data, and automation of the creation and consumption of relevant health data has reduced the amount of effort required to maintain Health Information Systems (HIS). This also results in quicker turnaround of medical results and related information thus enhancing the quality of patient care. Access to reliable health information systems reduce errors, speed up complex processes and reduce the workload of toiling caregivers by making information accessible at the point-of-care [1]. However, since mobile caregivers frequently rely on notes or their own memory, it is common for a caregiver to forget some items of interest or not be able to read their scrawled notes [1]. Many healthcare agencies have implemented electronic information systems which are highly customized for their own environments. Though these systems are internally successful, there is a growing need for external exchange of clinical information. The expansion of health information systems outside their environments is generally limited due to lack of standardization and security concerns. In order to more effectively facilitate exchange of data between health care agencies, a Clinical

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_14,

Document eXchange (CDX) project was initiated. This project seeks to standardize message formats and centralize the document exchange web service so that health agencies with disparate systems can communicate with each other. It has been recognized that the use of clinical documents which conform to Health Level 7 (HL7) standards, specifically the Clinical Document Architecture (CDA) specifications [2], is one of the preferred ways to implement message exchange. CDA is a standard for structured documents that describes the syntactic and semantic rules for clinical documents in order to easily facilitate the exchange of such documents across health information systems [2] [3].

Secure communication and access to patient's Electronic Medical Record (EMR) is an essential component in a shared care environment when multiple physicians and specialists are involved in patient's care. This requires that non-compatible health information systems (HIS) must be able to communicate with each other using conformant messages. The functionality must also be augmented for scenarios when mobile devices are used and the HIS is not readily accessible. In this paper, we propose a solution which demonstrates how mobile devices can securely use the CDX system to send/receive messages across multiple mobile platforms and with instances of CDX compatible EMR systems. The proposed solution is not itself an EMR system, but serves as an extension to any EMR which uses CDX to implement external clinical document exchanges.

1.1 Background

Interior Health (IH) and Northern Health (NH) are two of the five health authorities in British Columbia, which provide healthcare to the general populace. Both are spearheading the adoption of CDX system as they move towards enhancing secure exchange of clinical information. IH is developing the CDX web services while NH is developing applications which will use the CDX services. NH is faced with a unique set of challenges as it serves a sparsely populated region consisting of a mere 300,000 people within the largest Health Authority in British Columbia which spans about 600,000 km^2 [4]. Thus, physicians and other healthcare providers often need to travel to remote locations to provide clinical services. This removes them from their respective offices and effectively cuts them off from their EMR system particularly when services are being rendered in locations without internet access. In order to access patient's health information at point-of-care, a physician either takes patient charts, prints the necessary information or downloads it to a mobile device. These methods present security risks, with files or papers going missing and basic mobile devices being non-secure. Additionally, the physician is required to record the events of the visit which are noted and must be transcribed later. This is more costly and can introduce errors, specifically human errors. During an offsite visit, the physician may also need to consult with other physicians or specialists, provide them with necessary information including attachments which may not be available or, if available, appropriate to transmit using the unsecure, conventional email systems. This calls for a secure, structured, electronic means of creating care records and securely communicating those to the patient's EMR or sharing with other providers involved in patient's care. Due to connectivity issues, this functionality is desirable to work in both online and offline scenarios. The offline mode poses more security challenges as messages must be stored locally until the mobile device syncs with the selected EMR upon coming online.

2 Related Work

Currently, no mobile platforms exist which provide secure and reliable EMR-agnostic communication with electronic health records. An XML-based Clinical Document Architecture prototype to move discharge summaries from hospitals to family practice locations was proposed in [5]. The proposed implementation is web-based, lacks mobile capabilities and is limited to Discharge summaries and Referral forms. The developed prototype is a stand-alone system that does not integrate with existing health records or EMR systems. Security issues or concerns are also not addressed. An infrastructure used to bring health records to mobile phones in a small clinic has been proposed in [6]. A web-based application displays patient information on mobile devices. However, the implementation is location specific, does not adhere to HL7 standards and is not available for interfacing in a generic manner with any EMR system; additionally, the web-based implementation implies that a user must have an internet connection to use the application. They also do not address any security issues or concerns. A cross-institutional implementation of a web-service to view and transfer clinical data between institutions is presented in [7]. Though this implementation adheres to HL7/CDA specifications and addresses security issues, the work is limited to non-mobile computing.

3 The Proposed mEMR Solution

We identified two methods which could potentially be used for retrieving and uploading EMRs using the CDX system. The first method is to use a web-based application that is accessible through the browser of any device, computers, tablets or phones. This allows greater versatility and choice for the physician about which device to bring to what situations. However, the devices themselves may be insecure thus putting any downloaded records at a risk for compromise. Another disadvantage of the web-based method is its reliance on internet connections which severely limits the use. The vast expanse of the region under consideration only extenuates the need for an offline-capable solution. The second method is to develop an application for a mobile device, primarily on tablets because the message templates and views are fairly large and contain many fields. In mobile environments, a key challenge is to create a secure environment for the patient information and records. Mobile devices must also be able to function in an offline mode. This gives the physicians a persistent, secure access to records in their EMR system and the capability to securely consult with other physicians when needed. It is important to note that some basic information from EMRs must be downloaded beforehand if use in offline mode is anticipated. We chose to use the second method due to the additional security and offline capability which allows the physician to create messages while offline; since the messages are CDX compatible, the devices transparently syncs these with the selected EMR when the application comes online.

3.1 Mobile EMR (mEMR)

mEMR has been developed in parallel for two platforms: iOS and Android. We present our work with reference to iOS platform and note that the concepts presented are almost identical across the two platforms. Users of the mobile application must be registered as a provider in CDX and are required to login for authentication. Each CDX account has multiple providers. Once logged in, the user selects the appropriate provider. The application has two main views; the message list split view and the message creation view. The message list split view has two sub components: the message list table view and the message viewing pane. The user is presented with the split view upon login.

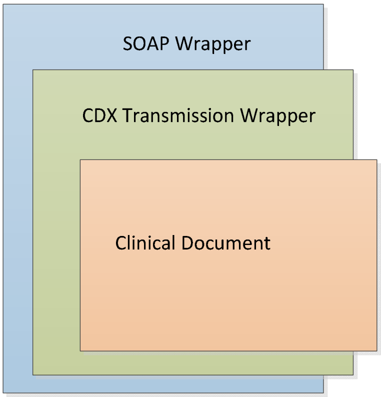


Fig. 1. Message Structure

3.1.1. Messages

mEMR sends and receives messages that strictly adhere to CDX specifications. A typical message consists of the following hierarchy: a SOAP wrapper wraps the CDX transmission wrapper that in turn contains the Clinical Document (Fig. 1). The SOAP wrapper defines the transmission protocol which the message adheres to and specifies how to process the message [8]. The CDX transmission wrapper holds information relevant to the message being sent to and from CDX, such as the message identifier, the creation time of the current CDX message and the sending and receiving CDX accounts. We use eight different CDX document types in our solution. These message types allow us to communicate with CDX to get message lists, retrieve individual messages, submit messages and retrieve provider information. The SOAP wrapper, the CDX wrapper and CDA are all XML-based. The CDX wrapper and clinical documents use XML according to the HL7 v3 specifications. XML is both human readable and computer processable [3]. This allows for the implementation of a system that generates and consumes messages efficiently and logically.

CDA is an XML-based standard defined by HL7 that has an express goal of facilitating the transmissions of clinical documents across Health Information Systems. It is a three-layer architecture with each incremental layer adding increased specificity to the document [5]. CDA Level One includes the structured CDA header and a human readable unstructured body. CDA Level Two extends the unstructured narrative blocks into sections that are identified with a section code. CDA Level Three adds entries to CDA Level Two and adds the full vocabulary of codes according to HL7's Reference Information Model (RIM) structure [2] [9]. The mEMR application allows viewing of CDA Levels One, Two and Three, and it allows for the creation, replying, forwarding and saving of messages conforming to CDA Level One. Data quality and richness increases as the CDA Level increases [9]. This means that a physician can more accurately and clearly describe a patient's state of health through a higher-level document.

In our solution, we represent the Clinical Document Architecture as a hierarchy of classes. Each class represents one element of the document; attributes of the class represent the attributes and values for that element. The code to encapsulate the CDX structure consists of 48 classes, and the CDA structure required 39 classes. The CDA structure is currently being expanded as we move to support CDA Level 2 and CDA Level 3 messages. Each structure is recursively defined which enables recursive printing.

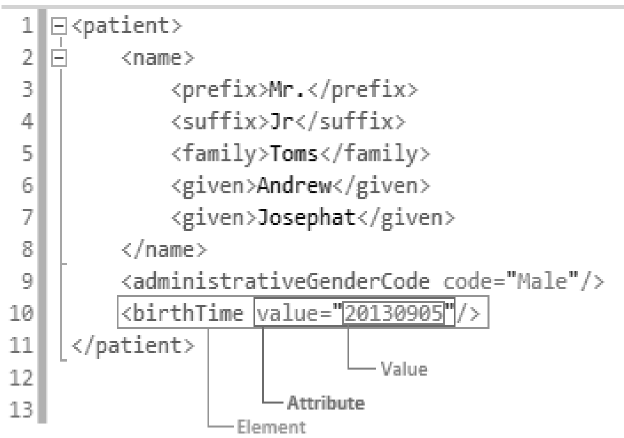


Fig. 2. XML Patient Structure

Fig. 2 shows the XML structure of the patient element and its subcomponents as defined by the EHR specification. We can read from this structure that a patient has a name, an administrative gender code and a birth time. The name element consists of additional elements: prefix, suffix, family and multiple given names. These sub elements all hold a value. The administrative gender code and the birth time elements both have additional attributes, code and value, respectively. In our implementation, this structure transforms into the class structure shown in Fig. 3. In order to retrieve the XML string from the class, we override the public String toString() method in Java and the -(NSString *) description method in Objective C. We print the elements with their attributes and values using these methods; if an element consists of other constructs, their respective printing methods are called. In this manner we can print a message recursively by printing the root component of the message. We chose to represent the XML Structure as a class structure to ensure structural integrity and extendibility.

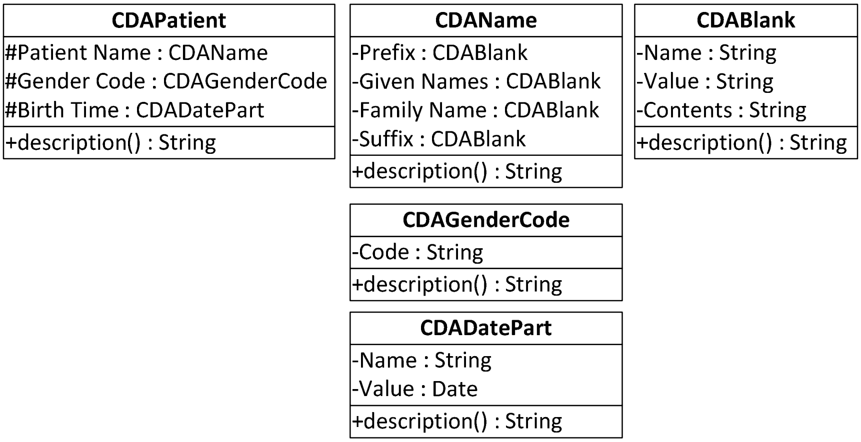


Fig. 3. CDA Patient Classes

3.1.2. Structure

The application's view structure is illustrated in Fig. 4. The user is first presented with the login view, and on successful login the user is taken to the main view, which contains two sub views: the message list view and the message display view. The message list view shows a table of message metadata that is used to choose the mailbox or message which the user wants to view. The message display view shows the contents of a selected message in an aesthetic format. There are two viewing options for the main message view. The default option is a custom view for displaying message contents and associated attachments. There is also an additional view that the user can switch to so that an XSLT file [10] formats the XML string of the CDX message. This XSLT file is provided by Interior Health and is used as the standard format for displaying CDX specific messages. We implemented the custom view to interact nicely with our class structure and to display attachments, as the XSLT view did not lend itself to storing/using message information or displaying attachments. From this main view the user can simply navigate to the other related views, such as the create message view, and the settings view. The Quick Look Preview View and the Camera View are used for viewing attachments and taking pictures, respectively.

The message list table view lists the metadata of the message list retrieved from CDX. This metadata is stored in the on-device database so that offline users can view their message lists. When a user selects an unread message from the list, the application requests the message from CDX web service. Once the message is retrieved from the server it is encrypted and saved in the onboard database. If the message already exists in the database, it is read locally, decrypted and displayed to the user. This is much faster than putting a request to the web service for the message each time the user wants to view the message and has the additional advantage of lowering the amount of data transferred between the CDX system and the application.

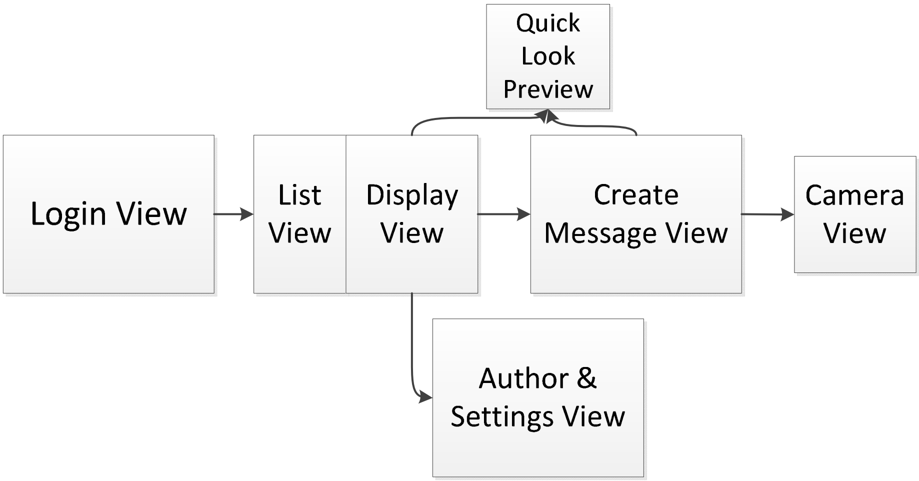


Fig. 4. mEMR View Structure

The message creation view allows the user to create unstructured reports and discharge summaries which are conformant to CDA Level One. The user can choose between the two templates. A user is able to reply to a message or save drafts for later retrieval, editing and transmission. The messages can also be composed using speech-to-text feature enabled on the keyboard. This feature is only available in online mode. We note that there is a privacy and security concern with this feature as the speech data has to be sent off to a remote server [11] [12]. Therefore, we advise caution when enabling this feature.

The application is able to send messages to and receive messages from any EMR system when the messages are transported through CDX. To avoid message retrieval collisions, the EMR and mEMR instances should be provided with distinct credentials or messages should be flagged as for mobile or for the EMR system. Messages sent by the mEMR conform to CDX message standards and the clinical document payload conforms to HL7 v3 CDA R2 standards. The authentication of the documents is done on server side, with incomplete or invalid documents being rejected. The mobile application hides this from the user, instead allowing them to easily enter patient information while the application itself ensures conformance to the CDX and HL7 standards.

3.2 CDX Challenges

One of the main difficulties encountered with the development of our mobile EMR solution was the lack of control over the communications system. Since the web services were implemented with statically located EMR systems in mind, it is understandably difficult to create a messaging-like application on a mobile platform using the provided web services. There are certain features that are generally provided to messaging applications but are not provided by the CDX system, such as authentication services and push notifications. In order to effectively differentiate between messages for a user, the following metadata needs to be present: author, creation time, title, type, and message identifiers. Unfortunately, the development environment for CDX does not supply these fields. CDX currently supplies the message identifier and the message type. This is not sufficient for the user to identify the sender or the message's subject. Thus, we must resort to a more complicated message metadata retrieval method. Similarly, CDX does not have an authentication method, so it relies solely on the clients to ensure that the users are legitimate. We implemented a pseudo-authentication method by asking the user for the identifier, password and CDX account used for querying the provider list. If the provider query returns a number of contacts greater than one, the message and its credentials are accepted as valid by CDX and the user is authenticated.

3.3 Security

We secure mEMR in a variety of ways. First, a triplet is used for credentials to access the application. The individual records within the database are encrypted followed by encryption of the entire database. For encryption, we use a user-specific master key that is generated from the user's login identifier and password. Since the credentials are provided for each session, the master key is not permanently stored, but generated for the session. As part of generating the master key, we check if the master cryptographic salt exists in the keychain of the device; if it does not exist, the salt is generated and stored in the keychain. The salt and a string of concatenated username and password is then run through the Password-Based Key Derivation Function 2 (PBKDF2) [13]. We use PBKDF2 with 10000 iterations; the salt length is 256 bits. The master key for the current session is stored in memory.

To encrypt a value in the database we use CCCrypt [14] encryption library as provided by Apple. We use the CCCrypt() function to encrypt a single database value as this function acts as a one shot encryption or decryption operation [14]. The CCCrypt() function is provided with the proper parameters so that it uses the kCCAlgorithmAES128 algorithm to encrypt and decrypt the data. CCCrpyt() also requires an Initial Vector, that is, an additional random number that helps with ensuring duplicate sequences do not exist after encryption. We provide a 128 bit Initial Vector. For convenience, the Initial Vector is stored in a separate field instead of being prepended to the encrypted message. Once the message's raw data has been encrypted it is stored in the database with its unencrypted metadata.

The mEMR application offers offline capabilities. This requires secure storage of messages and their associated attachment(s). Currently, we store the messages in an encrypted on-device *SQLite3* [15] database. On the iOS platform, we use the CCCrypt library to encrypt the message. Attachments can be encrypted on device using this same method and are viewable using a custom UIView. Attachments are limited to the following filetypes: .tiff, .jpeg, .png, .rtf and .pdf as per the CDX specifications. In order to use Apple's UIDocumentController or QuickLookPreviewer, the attachments need to be stored unencrypted on the file system. However, this was unsuitable for our security requirements, so we implemented a custom UIView that allowed us to encrypt the attachments before writing to the disk and decrypt upon retrieval for display. This ensures that the attachments are never stored unencrypted on the disk, with the notable exception of photos taken from the on-device camera and attachments from outside of the application (such as on the camera-roll and public libraries). CDX will soon move to securing its SOAP messages through X.509 certificates or Kereberos certificates. This will fully secure the transmission of documents to and from the CDX system. Documents that are sent by the system will be fully encrypted during the transmission process and will only be able to be decrypted by a known client with the required certificate.

4 Conclusion

We have demonstrated that mobile devices can securely exchange messages across mobile platforms and with heterogeneous healthcare information systems in both online and offline modes. The clinical documents are required to be CDA compliant. This allows healthcare providers to view, create and share health records in order to enhance the quality of patient care. The proposed infrastructure is specifically useful in a shared care and point-of-care environments. All processes and patient data are secured by sophisticated encryption. We examined the issues surrounding working with a web service that is beyond our direct control. The solution presented is robust and has been demonstrated to link with existing EMRs without requiring any modification at their end. One of the limitations is the lack of ability to download patient information on demand as CDX currently restricts this functionality. The proposed solution, however, has the capability to initiate a query to a web service for patient information in order to easily populate the forms. As an alternative, one can also ping the EMR to obtain an empty message which can auto-populate the required fields in the response. As future work, we intend to expand the functionality to incorporate higher levels of CDA such as including templates for lab reports.

Acknowledgments. This work was funded by a research grant from Northern Partners in Care (NPiC). NPiC is an initiative funded through the BC Ministry of Health and Doctors of BC which aims to improve the quality of care across Northern BC by improving the effectiveness of communication between family and specialist physicians. The underlying EMR was provided by Applied Informatics for Health Society (AIHS).

References

1. Abraham, C., Richard, W., Boudreau, M.-C.: Ubiquitous Access: on the front lines of patient care and safety. Communications of the ACM 51(6), 95–99 (2008)
2. Health Level Seven International: HL7 CDA Release 2 (2005), http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7 (accessed November 12, 2013)
3. Muller, M., Frankewitsch, T., Ganslandt, T., Burkle, T., Prokosch, H.-U.: The Clinical Document Architecture (CDA) Enables Electronic Medical Records to Wireless Mobile Computing. Stud Health Technol Inform 107(2), 1448–1452 (2004)
4. Northern Health: About Us, http://www.northernhealth.ca/AboutUs.aspx (accessed 2013)
5. Paterson, G., Sheperd, M., Wang, X., Watters, C., Zinter, D.: Using XML-Based Clinical Document Architecture for Exchange of Structured Discharge Summaries. In: Proceedings of the 35th Hawaii International on System Sciences, Maui (2002)
6. Ghose, A., Bhaumik, C., Agrawal, A.K.: Mobile Healthcare Infrastructure for Home and Small Clinic. In: Proceedings of the 2nd ACM International Workshop on Pervasive Wireless Healthcare, New York, pp. 15–20 (2012)
7. Müller, M., Ückert, F., Bürkle, T., Prokosch, H.-U.: Cross-institutional data exhange using the clincal document architecture (CDA). International Journal of Medical Informatics 74(2-4), 245–256 (2005)
8. W3C: SOAP Version 1.2 (2007), http://www.w3.org/TR/soap12-part1/ (accessed October 07, 2013)
9. Brull, R.: CDA Levels of Interoperability (2011), http://www.hl7standards.com/blog/2011/06/02/cda-levels-of-interoperability/ (accessed June 2, 2013)
10. W3C: XSL Transformations, http://www.w3.org/TR/xslt (accessed August 3, 2013)
11. Banks, T.: Cloud Computing And The USA Patriot Act: Canadian Implications, http://www.mondaq.com/canada/x/191964/IT+internet/Cloud+Computing+And+The+USA+Patriot+Act+Canadian+Implications (accessed August 20, 2012)
12. British Columbia: Freedom of Information and Protection of Privacy Act, http://www.bclaws.ca/EPLibraries/bclaws_new/document/LOC/freeside/-%20F%20-/Freedom%20of%20Information%20and%20Protection%20of%20Privacy%20Act%20RSBC%201996%20c.%20165/00_Act/96165_03.xml#part3_division1 (accessed September 11, 2013)
13. Internet Engineering Task Force (IETF): Network Working Group, http://tools.ietf.org/html/rfc2898 (accessed August 12, 2013)
14. Apple: CCCryptor(3cc), https://developer.apple.com/library/ios/documentation/System/Conceptual/ManPages_iPhoneOS/man3/CCCryptor.3cc.html#//apple_ref/doc/man/3cc/CCCryptor (accessed November 12, 2013)
15. SQLite: SQLite, http://www.sqlite.org/ (accessed September 12, 2013)

Mobile Botnet Attacks: A Thematic Taxonomy

Ahmad Karim*, Syed Adeel Ali Shah, and Rosli Salleh

Department of Computer Science and Information Technology,
University of Malaya, Malaysia
ahmadkarim@siswa.um.edu.my

Abstract. Mobile botnets have recently evolved owing to the rapid growth of smartphone technologies. The implications of botnets have inspired attention from the academia and industry alike, which includes vendors, investors, hackers and researcher community. Above all, the capability of botnets is exploited in a wide range of criminal activities, such as, Distributed Denial of Service (DDoS) attacks, stealing business information, remote access, online/click fraud, phishing, malware distribution, spam emails, and building mobile devices for illegitimate exchange of information/materials. In this paper, we investigate mobile botnet attacks by exploring attack vectors and a subsequent presentation of a well-defined thematic taxonomy. Through identification of significant parameters from the taxonomy, we conduct a comparison to explore effects of existing mobile botnets on commercial as well as open source mobile operating system platforms. The parameters for comparison include mobile botnet architecture, platform, target audience, vulnerabilities/loopholes, operational impact and detection approaches. Related to our findings, we present open research challenges in this domain.

Keywords: Mobile botnet, smartphone, attacks, malware.

1 Introduction

Mobile attacks are the most critical and emerging threats due to increasing market penetration of smartphones and handheld devices. These smartphones use full-featured operating systems incorporated with powerful hardware that supports manifold interfaces and sensors. At present, personal computers (PCs) have declined as a standout choice of computing. Recent statistics show that global shipments of mobile devices immensely exceeded as compared to personal computers (PCs) since 2011 [4]. Furthermore, in the near future, the wide-scale deployment of 4G technologies e.g LTE and WiMAX will become the major source of broadband Internet access for general public. During 2012-2013, 4G enabled devices represent only 0.9 percent of all global mobile connections and they acquired for 14 percent of mobile data traffic [6]. This technological shift has inspired cyber criminals to exploit the vulnerabilities of smartphone devices through off-the-shelf malware creation tools

* Corresponding author.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_15,

[8]. Similarly, the burst of mobile applications have enabled the dissemination of malicious code to potentially wide range of audience. The majority of current mobile threats replicate the behavior of attacks on desktop machines through the Internet. Therefore, many of the existing solutions can be considered applicable to the malicious mobile attacks as well. In spite of that, mobile devices have their own constraints such as limited processing, less data storage capabilities and heterogeneity of operating systems (OS) (Android, Apple, Windows etc.), that restricts the security solutions to be programmed efficiently.

Botnet is a network of compromised machines, aiming to collectively perform some activity on the recommendation of botnet creator (bootmaster). The intentions of botmaster are to disrupt legitimate services over the Internet or deceive the private information. In the similar context, with the rapidly growing mobile computing world, mobile botnets are evolving as a serious threat towards targeting mobile phone devices such as smartphones. The motive of this attack is somewhat similar to that of traditional botnet attacks --- to gain access to the resources, interpret contents of mobile user device and transfer control to the botnet initiator. This only comes true when the hacker takes advantage of the exploited area/loopholes of mobile devices to gain unauthorized access to the compromised mobile devices. Eventually, the hacker's goal is to perform malicious and unauthorized activities including illegal phone calls, accessing control panel, sending emails, initialization of worm code and unauthorized file access or photos [12].

Andbot [12] is a mobile bot which employs URL flux and it is considered as a stealthy, low-cost, and resilient bot, which uses botmaster for illegal activities in mobile environment. This botnet uses microblogs to send malicious commands. It is found that Andbot can be easily implemented on smartphones for longer durations without being noticed or detected. Andbot integrates several other schemes to make it efficient and stealthy. Cloud Based Push-Styled Mobile Botnets [16] is a new type of botnet in mobile environment that uses push-based notification services to disseminate the commands.

Recently, a number of mobile botnets have evolved that can degrade the performance of mobile device. For instance, ZeuS [17] is a botnet that focuses the Blackberry, Symbian and Windows platform users and DreamDroid [19] botnet affects the Android based devices. Similarly, IKee.B [21] is a botnet that is used to scan IP address for iPhones, whereas BMaster and TigerBot specifically target Android application frameworks. Similarly, Epidemic mobile malware is a new terrifying threat for mobile users [22] which disseminates rapidly in smartphones. The malware affects the older version of iOS, however still, epidemic mobile malware is a predominant threat for mobile users. Mobile botnets is a relatively new research domain which constitutes a number of problems. Consequently, the detection, analysis and mitigation have become hot issues nowadays for the industry and research.

To the best of our knowledge, in the existing literature, a comprehensive survey on mobile botnet attacks does not exist. This is the first comprehensive review on mobile botnet attacks exploiting mobile botnet architecture, platform, target audience, vulnerabilities/loopholes, operational impact and detection approaches. Therefore, the contribution of this review is three fold: (a) from the extensive literature survey we conclude that, in order to comprehend mobile botnet's threatening effect, a comprehensive understanding of the features of malicious mobile attacks is essential, i.e. Type of attack, platform, category, target audience, loopholes, dissemination techniques, operational impact and defensive approaches. Therefore we timeline the mobile botnet/malwares according to the above mentioned properties, (b) we propose thematic taxonomy of state-of-the-art mobile botnet attacks to highlight different aspect of attacks as well as recovery techniques to avoid this growing threating phenomenon(c) further, we highlight open challenges and issues pertaining to the dissemination of these malicious mobile botnet threats.

The rest of the paper is organized as follows; a thematic taxonomy is presented in section 2 to classify the mobile botnet attack vector; in section 3, we compare the existing mobile botnet attacks based on the significant parameters derived from the taxonomy. Finally, section 4 highlights issues and challenges that require further research in avoiding mobile botnet attacks.

2 Thematic Taxonomy Based on Mobile Botnet Attack Vector

In this section, we present a thematic taxonomy of mobile botnets based on the mobile botnet attack vector as shown in Figure 1.

2.1 Thematic Taxonomy

In Figure 1, based on the exhaustive survey of botnet attacks, we present a thematic taxonomy on the basis of architecture, platform, attack types, loopholes, target audience, operational impact, and defensive approaches. In the following section, we briefly highlight the existing contributions in mobile botnet attack vector, as well as open areas for research.

Architecture: PC based botnet are considered as the most compromised platforms for botnet attacks as compared to the recently evolved mobile botnets due to several reasons: (a) limited battery power (2) resource constraints (3) limited Internet access etc. Mobile botnet architecture has similarities with the traditional PC based botnet architecture. For instance, similarities in the underlying C&C communication protocols exist that includes IRC, HTTP [23], and P2P [24]. These protocols provide coordination between botmaster and bots in a desktop computing environment. In addition to these, SMS [25] [26] [27], Bluetooth [28] or MMS mechanisms have also emerged in mobile botnets

Loopholes: Institutions and consumers keep themselves up-to-date about mobile threats because mobile devices are vulnerable to new threats. In particular, Android OS is a victim of malware attacks as reported by[29]. Its increased market share and open source architecture is the enabling factor in the exploitation of various attacks. Although new versions of android operating systems are more resistant against security vulnerabilities, nevertheless, 44 percent android users do not update and hence compromise their mobile security. SIM cards are the tiny computers inside most mobile devices that allow them to communicate with the service provider. According to one security research [30], flaws in SIM card technology and implementation make hundreds of millions of mobile devices susceptible to being hacked. The root of the problem is the fact that encryption in most SIM cards relies on DES (Data Encryption Standard)- an algorithm created by the US government four decades ago. DES was secure in its day, but the day has long since passed. Now, DES is considered insecure, and is relatively trivial for a skilled hacker or crack.

Attack Types: A recent study by Kespersky [15] reported that the most common attacks targeting Android platform are SMS Trojans, adware, viruses, spywares and root exploits. Moreover, mobile botnets are becoming serious threat focusing different mobile platforms that can perform various tasks at the instructions of botmaster. Botnet typically uses DNS to retrieve IP addresses of the servers, therefore targeting DNS service is initial point of attack. It results in activation of incredibly robust and stealthy mobile botnet C&C [31]. Moreover, the key feature of mobile communication relates to the exchange of traffic load and its constant observation for billing and accounting. Consequently, mobile botnets have the potential to affect the call charging detail records (CDR) of the infected mobile systems [4]. Another approach used to reduce such activities is known as the rootkit [32]. It is a type of malicious code specially designed to hide the unwanted activities and virus propagations from the system. In this case, the C&C of a botnet instructs the bots to carry out a malicious activity including, sending spam messages or acquiring authorized control over to smartphone devices and hijack business activities.

Target Audience: Mobile botnets are focusing target audience from diverse environments ranging from public audience to government, enterprises and organizations. Profitable organizations like banks (which are shifting majority of their services to the mobile environment e.g. payment of bills, generating account statements and funds transfer). Therefore, the primary focus of mobile botmaster is to gain access to those mobile devices which are dedicated for business activities and tries to launch various activities for instance, DDoS, remote control, and hijacking of private/confidential information. According to Information Week[33], Bank of America, U.S bank, Wells Fargo and JPMorgan Chase were among those U.S banks that were slowed down by DDoS attacks. As a consequence of this attack, thousands of their customers filed complaints for site down and could not access their normal banking activities e.g. account checking, saving, bill payments, mortgage accounts and other similar services through mobile applications.

Operational Impact: Overall operational impact of mobile botnet can be seen in two different perspectives: 1) relevant to the host device itself and, 2) relevant to the service provisioning model. The direct impact related to the host mobile device includes privacy violation, data theft, root access, location identification and battery consumption. Similarly, the concept related to service provisioning model includes, disruption of services, channel occupation, outage of resources, and content compromise.

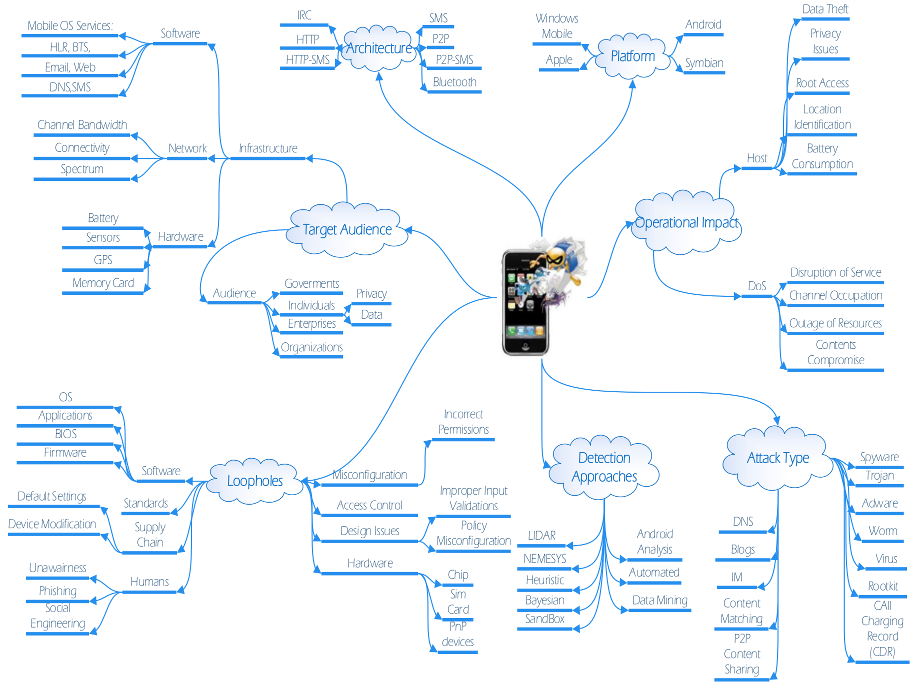


Fig. 1. Thematic Taxonomy based on Mobile Botnet Attack Vector

Detection Approaches: A mobile botnet detection approach based on "pull" style C&C was presented in [34]. Through investigating flow features (total packets, total bytes) of C&C traffic passing through VPN, the authors investigated the abnormalities of these traffic flows. This approach can also detect mobile botnets residing within signatures, abnormal models and whitelists.

A layered IDS and Remediation framework was proposed by [35] which automatically detects, analyzes, protects and removes security threats in smartphones. This study aims to overcome smartphone vulnerabilities and threat detection in three stages: a) behavioral and threat modeling technique, b) implementation and deployment of stochastic and machine learning techniques to automatically detect intrusion detection that can span among different network layers, applications and social media, and c) builds an automatic threat detection model to lessen or even overcome the security risks.

A behavioral model for the detection of smartphone malwares based on ontology techniques was proposed in [36]. Another approach [4] focused to design a network based anomaly detection method based on analytical modeling, learning and simulating, together with the billing and control-plan data to detect mobile attacks and anomalies. Furthermore, authors in [37] created a virtual lab environment for the purpose of analysis and detection of Android malwares through emulating the environment. In addition to that, a signature based mobile botnet detection algorithm considering Bayesian spam filter mechanism as the key component was proposed in [r10].The authors concluded that this system is capable of identifying 87% of spam message from the dataset.

A prototype named Airmid [29] was designed to automatically identify and respond to malicious mobile applications through analyzing the behavior of the network traffic. It identifies malicious network traffic through the help of cooperation between smartphones and in-network sensors. To detect Android malwares and malicious behavior of applications, a hybrid (static & dynamic) model was proposed in [38] which detects malicious activity through the following stages: a) static analysis --- to parse Manifest file of applications and decompile using reverse engineering tools, and b) dynamic analysis--- execute an application and log all performed actions.

3 Comparison of Existing Mobile Botnet Attacks Based on Taxonomy

In this section, we compare the existing mobile botnet attacks based on the significant parameters derived from the taxonomy. Table 1 shows the comparison.

3.1 DreamDroid

A mobile botnet based malware DroidDream [16] appeared in spring 2011. The purpose to launch this attack is to gain root access to Android mobile devices in order to acquire unique identification information (model number, product ID, EMI number, provider, language etc) of the mobile phone. Moreover, after infection, the compromised device could also download and install additional executable programs and features without being noticed by the user, while providing a backdoor root access for attacker. Later, Google managed to remove the effected applications from its official marketplace and had implemented a “kill switch” mechanism to remotely clear Android handheld devices that had been malfunctioned by DroidDream malware.

3.2 SymbOS.Yxes[2]

This worm targets Symbian mobile devices with OS 9.1S60 3rd Edition, but can also run on wider range of Symbian Operating systems. The potential capabilities of this worm are (a) sending messages to those phone numbers that were harvested from infected devices' SMS inbox. (b) steal information from the victim device e.g. serial number of phone, subscription information and redirect this information to servers controlled by cybercriminals (c) search installed applications from application manager and attempt to kill those tasks or applications. This worm uses valid but revoked certificate, therefore it is required for a device to avoid this attack through enforcing online verification of certificate.

3.3 IKee.B

A standalone malicious program that infects iPhone in different ways, such as (a) the device is 'jailbroken'-hacked and installed a software that is not signed by Apple (b) installation of unsigned secured shell (SSH) with remote access enabled capability (c) default root password('alpine') has not been changed from the default factory setting. Similarly, this worm can infect other vulnerable iPhones by scanning over 3G or wi-fi networks. Moreover, its dispersion takes place in three stages when active on the iPhone (a) changes default password (b) establish connection with remote server 92.61.38.16 via HTTP and download and install additional components (c) send banking information incorporated in SMS messages to remote server. The only defensive reaction is to reset iPhone and restore all setting to its factory default.

3.4 BBproxy

A Trojan malware was detected in blackberry smart phones in Mid-2006, which targeted the enterprise data and network. Initially it creates a trust relationship between a blackberry device and company’s internal server. Once the connection is established, it hijacks and establishes a connection with the company’s internal server. In addition to that, data tunnel established between both entities is based on a secure tunnel. Therefore it is difficult to detect any suspicious activity for intrusion detection system which is installed on the perimeter of the network. The recommendations to avoid this malicious act are: (a) keep blackberry server in demilitarized zone (DMZ) (b) the communication between blackberry server and device should be restricted.

Table 1. Mobile Botnet Threat Analysis: A TimeLine

Botnet/ Malware	Type	Platform	Category	Target Audience	Loophole	Dissemination Technique	Operational Impact	Defensive Reaction
DroidDream [1]	Root Exploitation	Android	Trojan	Android users	Alter code for Root access	Games	Root access, steal data	Android App Kill switch
SymbOS. Yxes[2]	Service disruption	Symbian OS 9.1	Worm	Symbian Users	Invalid certificate registration	Sending SMS, Redirect to cybercriminal website	Abnormally high phone bills, battery power loss	FortiGate Systems, FortiClient Systems
IKee.B [3]	Root Access	Apple	Worm	Systems and Networks, iPhone users	Unapproved SSH, setting default SSH password	scan and infect other iPhones by Wi-Fi or 3G networks	Stole financial sensitive information	Restore firmware via Apple iTunes
Geinimi [5]	Personal Information Theft	Android	Trojan	Android Users	Exploit backdoor	Games	Send private information to C&C via HTTP	Symantec Power Eraser Tool (SPE)
RootSMART [7]	Root Exploitation	Android	Malware	Android <2.3 or 3.0	GingerBreak Root Exploit	Through two Helper-Scripts	Establish connection with C&C	Use reputable app store
SMiShing[9]	Spam, Fraud	Any	Phishing	Any	Phishing to humans	Monetized by signing up	Steal personal information	Educate People
Snooping [10]	Privacy/ Snooping	Android <2.3.4 and 3.0	ftware ult	Users using synchronization services	Misusing Google's ClientLogin Protocol	Attacker snoops AuthToken in clear text	Impersonate user to change his personal info	Minimize timeout of AuthToken
SpySmart Phone [11]	Spy Software	Any	Sensors	Any	Phishing to humans	Installation on victim machine	Steal personal information,	Educate People
SSL Renegotiation DoS [13]	DoS, Asymmetric Processing	Any	Generic Attack	SSL/TSL servers	TSL Operations	Massive TLS renegotiation requests	Deplete Server resources	Disable SSL/TSL renegotiation
BBproxy [14]	Infrastructure	Blackberry/ RIM	Trojan	Enterprise Internal data and network	Exploit the trust relationship	Games, Email	Steal companies' Information	Separate DMZ, limited access
Foncy [15]	SMS Trojan	Android	Trojan	Any	Sending random messages to victims	Working with IRC bot and a root exploit	Malicious activates initiated by C&C	Already Dead
Cawitt [15]	SMS Trojan	Android	Trojan	Twitter Users	Posting Message on Twitter	Unknowingly sending SMS to premium users	Information Threat	Antivirus Scanner
SpamSold[18]	SMS Spam	Android	Spam	Any	Deceptive Android Permissions	Fee games	Establish connection with C&C	Various Antivirus Software
Obad[20]	Admin Exploitation	Android <v4.3	Trojan	Android cell holders	Google Play fake stores	Spam text messages	Attain admin rights to hack a firm	Patch in v4.3

4 Issues and Challenges

As a result of exhaustive survey on the existing botnets, we identify open issues for progressive security of mobile devices against botnets. With the proliferation of mobile technology and cloud computing services, the following issues are of concern for academia and industry alike:

- Initially, manifestation of a cross-functional group is essential that involves researchers and the stakeholders (e.g. enterprises, governments, networks, and ISPs) for identification and effective confiscation of botnets. A clear and transparent policy on mobile equipment and use must be documented and socialized across the enterprise. Moreover, the public audience should be aware of the means by which mechanisms are designed and developed to overcome mobile botnet threats.
- There is no way for security and risk leaders to ignore the increasing demand and proliferation of mobile into the enterprise at this point. The demand isn't just being driven from by the mass adoption and use consumer devices, but businesses are also leveraging the power of mobile computing to strengthen their value to their clients and customers, making them more agile, relevant, and able to respond to the needs of their customers.
- Scanning and blocking of malicious code in cloud can be implemented to preempt the code or information sharing centers in cooperation with antivirus vendors identify and plan to block the threats. When the malicious code is preempted it may not be possible for providers to predict how devices with more operating platforms receiving the code will behave with traffic. But in case of detection and block management of threats it can be applied in blocking solutions.
- As compared to desktop operating system, smartphone device operating system has less capability in terms of processing, memory and storage, which ultimately restricts the security policy to be implemented at its best.
- Network operators have remarkable control on the software employed for smartphones, which are using their network. The case happens especially when mobile phones are sold as part of wireless subscription. The operators should provide built-in anti-virus scanning facility and should enforce updating and patching in response to any malicious activity.
- User awareness with respect to security threats is a key contribution towards persistent solution of the problem. Therefore, a relevant and determined education and awareness campaign should be introduced that targets mobile users on the risks, policies, and procedures.

5 Conclusion

In this survey, we have conducted an exhaustive survey of existing botnet attacks on mobile devices. Through an investigation of botnet attack vectors, we have presented a well-defined taxonomy and used it to explore the acute features of existing botnet attacks. This review aims to serve as a roadmap for researchers to study and enforce secure communication patterns that are focused at various aspects of attack vectors.

Related to our observations about mobile botnet attacks, we conclude that Android has the minimum resistance against mobile botnets for two main reasons: 1) being open source that makes it a free to contribute digital contribution platform and 2) augmented market penetration that makes it suitable for the spread of botnet.

Addressing mobile botnet attacks have become a challenge for information security professionals and researchers. Therefore, it is necessary that, stakeholders must implement some cooperative and legislative actions to eliminate this hazard. Similarly, it is also important to negotiate on possible international legislative issues and establish global policies to systematically avoid this harmful threat.

References

1. DroidDream, DroidDream (2012), http://www.webopedia.com/TERM/D/droiddream.html (accessed on: November 30, 2013)
2. Center, F.: SymbOS/Yxes.A!worm!worm (2009), https://www.fortiguard.com/ve?vn=SymbOS/Yxes.A (accessed on: November 30, 2013)
3. Worm: iPhoneOS/Ikee.B, http://www.f-secure.com/v-descs/worm_iphoneos_ikee_b.shtml (accessed on: November 30, 2013)
4. Abdelrahman, O.H., et al.: Mobile Network Anomaly Detection and Mitigation: The NEMESYS Approach. arXiv preprint arXiv:1305.4210 (2013)
5. Android. Geinimi, http://www.symantec.com/security_response/writeup.jsp?docid=2011-010111-5403-99 (accessed on: November 30, 2013)
6. Arbor Networks: Worldwide Infrastructure Security Report (2012), https://www.arbornetworks.com/news-and-events/press-releases/recent-press-releases/4737-the-arbor-networks-8th-annual-worldwide-infrastructure-security-report-finds-ddos-has-become-part-of-advanced-threat-landscape
7. Xuxian, J.: Security Alert: New RootSmart Android Malware Utilizes the GingerBreak Root Exploit (2012)
8. Ollmann, G.: The evolution of commercial malware development kits and colour-by-numbers custom malware. Computer Fraud & Security 2008(9), 4–7 (2008)
9. Musthaler, L.: How to avoid becoming a victim of SMiShing (SMS phishing), pp. 10–11 (2013) (accessed on: November 10, 2013)
10. Mills, E.: Report: Android phones vulnerable to snooping attack
11. Kiley, S.: Spy Smartphone Software Tracks Every Move
12. Xiang, C., et al.: Andbot: towards advanced mobile botnets. In: Proceedings of the 4th USENIX Conference on Large-Scale Exploits and Emergent Threats. USENIX Association (2011)
13. Orchilles, J.A.: SSL Renegotiation DOS
14. Zetter, K.: BlackBerry a Juicy Hacker Target

15. SecureList: Mobile Malware Analysis: Part-6, http://www.securelist.com/en/analysis?calendar=2013-02
16. Zhao, S., et al.: Cloud-based push-styled mobile botnets: a case study of exploiting the cloud to device messaging service. In: Proceedings of the 28th Annual Computer Security Applications Conference. ACM (2012)
17. Zeus Botnet Eurograbber Steals $47 Million, http://www.informationweek.com/security/attacks/zeus-botnet-eurograbber-steals-47-millio/240143837 (accessed on: November 19, 2013)
18. Microsoft, Malware Protection Center, Trojan: AndroidOS/SpamSold.A (2013)
19. Android DreamDroid two: rise of laced apps, http://www.itnews.com.au/News/259147,android-dreamdroid-two-rise-of-lacedapps.aspx (accessed on: November 19, 2013)
20. Donovan, F.: Botnet of mobile devices used for first time to distribute Trojan (2013)
21. F.: securel Virus and threat descriptions
22. Szongott, C., Henne, B., Smith, M.: Evaluating the threat of epidemic mobile malware. In: 2012 IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). IEEE (2012)
23. HTTP-Botnets: The Dark Side of an Standard Protocol!, http://securityaffairs.co/wordpress/13747/cyber-crime/http-botnets-the-dark-side-of-an-standard-protocol.html
24. Grizzard, J.B., et al.: Peer-to-peer botnets: Overview and case study. In: Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets (2007)
25. Geng, G., et al.: An improved sms based heterogeneous mobile botnet model. In: 2011 IEEE International Conference on Information and Automation (ICIA). IEEE (2011)
26. Hamandi, K., et al.: Android SMS botnet: a new perspective. In: Proceedings of the 10th ACM International Symposium on Mobility Management and Wireless Access. ACM (2012)
27. Geng, G., et al.: The Design of SMS Based Heterogeneous Mobile Botnet. Journal of Computers 7(1), 235–243 (2012)
28. Singh, K., Sangal, S., Jain, N., Traynor, P., Lee, W.: Evaluating bluetooth as a medium for botnet command and control. In: Kreibich, C., Jahnke, M., et al. (eds.) DIMVA 2010. LNCS, vol. 6201, pp. 61–80. Springer, Heidelberg (2010)
29. Nadji, Y., Giffin, J., Traynor, P.: Automated remote repair for mobile malware. In: Proceedings of the 27th Annual Computer Security Applications Conference. ACM (2011)
30. Forbes, https://viaforensics.com/resources/reports/best-practices-ios-android-secure-mobile-development/mobile-security-primer/
31. Open DNS, security whitepaper, http://info.opendns.com/rs/opendns/images/OpenDNS_SecurityWhitepaper-DNSRoleInBotnets.pdf
32. Rootkit, http://en.wikipedia.org/wiki/Rootkit
33. Bank Site Attacks Trigger Ongoing Outages, Customer Anger, http://www.informationweek.com/attacks/bank-site-attacks-trigger-ongoing-outages-customer-anger/d/d-id/1106615?
34. Choi, B., Choi, S.-K., Cho, K.: Detection of Mobile Botnet Using VPN. In: 2013 Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS). IEEE (2013)

35. Roshandel, R., Arabshahi, P., Poovendran, R.: LIDAR: a layered intrusion detection and remediationframework for smartphones. In: Proceedings of the 4th International ACM Sigsoft Symposium on Architecting Critical Systems. ACM (2013)
36. Chiang, H.-S., Tsaur, W.-J.: Identifying Smartphone Malware Using Data Mining Technology. In: 2011 Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN). IEEE (2011)
37. Andrews, B., Oh, T., Stackpole, W.: Android Malware Analysis Platform. In: 8th Annual Symposium on Information Assurance, ASIA 2013 (2013)
38. Spreitzenbarth, M., et al.: Mobile-sandbox: having a deeper look into android applications. In: Proceedings of the 28th Annual ACM Symposium on Applied Computing. ACM (2013)

Evaluation of an Integrated Mobile Payment, Ticketing and Couponing Solution Based on NFC

Marta Campos Ferreira[1], João Falcão e Cunha[1], Rui José[2], Helena Rodrigues[2], Miguel Pimenta Monteiro[1], and Carlos Ribeiro[3]

[1] Faculdade de Engenharia, Universidade do Porto, Portugal
[2] Centro Algoritmi, Universidade do Minho, Portugal
[3] Instituto Superior Técnico, Universidade de Lisboa, Portugal
{mferreira,jfcunha,apm}@fe.up.pt, {rui,helena}@dsi.uminho.pt, carlos.ribeiro@tecnico.ulisboa.pt

Abstract. The emergence of Near Field Communication technology added more functionality to mobile phones. People can now access to new mobile services such as payment, ticketing, couponing, and control access to cars, homes and hotels. Although these services seem to complement each other, most studies have focused on the development and evaluation of each mobile service individually. In this paper, we propose an integrated mobile service solution – payment, ticketing and couponing – based on NFC. This solution was evaluated in a two-phase user study, which was run in real environment. Participants have largely appreciated the integration of tickets and discount coupons on the payment process. However, despite the integration of services being one of the strongest points of the solution, it also revealed to be one of the main challenges that any payment systems may have to address. Findings suggest that adding additional services to a solution may become very complex from a cognitive perspective.

Keywords: NFC, user study, user experience, field trial, mobile payments, mobile ticketing, mobile couponing.

1 Introduction

The general adoption of mobile devices combined with their increasing functionalities is changing the way people use mobile phones. Currently, it is already possible to make payments with mobile phones in several countries. Mobile payment can be defined as the use of a mobile device (mobile phone, PDA, wireless tablet) "to initiate, authorize and confirm an exchange of financial value in return for goods and services." [1].

Service Providers started to use the most basic mobile phone features, like phone calls and text messages, to make payment services available. Nevertheless, the emergence of contactless technologies such as Near Field Communication (NFC) added more functionality to smartphones. NFC enables other services and applications, beyond payment, such as mobile ticketing, couponing, and control access to cars, homes and hotels. Several studies about NFC-enabled mobile services

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_16, © Springer International Publishing Switzerland 2014

have been conducted in recent years. However, most studies focused their research on the development and evaluation of each mobile service individually.

In this paper we present an integrated mobile service solution (payment, ticketing and couponing) based on NFC that was developed under a research project called MobiPag. The MobiPag Project resulted from a Portuguese consortium of universities, technology companies, mobile phone operators and banks, which main aim was to develop a mobile service solution that represents a further step towards the dematerialization of money, tickets and coupons. The solution was developed for Android devices allowing customers not only to make payments but also to receive and redeem meal and bus tickets and discount coupons.

For this project two mobile applications were created: one for the costumers and the other for the merchants. Another important outcome of the project was the development of the MobiPag backend platform, which acts as a broker interlinking every service and device. It also contains several payment adaptors that connect with each financial account service.

The solution was tested in real environment at the University of Minho (UM) Campus. The user study was divided in two phases: 9 customers in the first phase and, 16 customers and 5 merchants in the second phase. The experiments allowed us to test the reliability and usability of the solution as well as the user experience in the interactions between costumers and merchants.

One of the main contributions of this study was the evaluation of a complex solution that combines payment, ticketing and couponing. We provide an analysis of the impact of this integrated solution on users' mental model and on their interaction with the system. Design guidelines regarding more user-friendly and ease-to-use mobile payment interfaces are also pointed out, as well as overall considerations about NFC performance.

The outline of the current paper is as follows: the related work section introduces a literature review about NFC-enabled mobile services and its prior study. Then the mobile payment solution is described, as well as the evaluation methodology. Finally, the evaluation results are presented, followed by discussions and conclusions.

2 Related Work

NFC is a short range wireless communication technology that when combined with smartphones enable several innovative mobile services such as payment, ticketing, couponing and access control [2]. NFC was considered a good choice for mobile payments in terms of speed, security and usability when compared with traditional mobile payment service concepts, such as Interactive Voice Response, Short Message Service, Wireless Application Protocol and One Time Password Generator [3] . In fact, NFC allows two way contactless communication, offers faster connection between devices, less chance of interference, and has a shorter range, making it more secure for use in crowded places.

In recent years several mobile payment, ticketing and control access systems based on NFC were developed and tested by researchers. Rehman and Coughlan [4] developed a mobile payment system to purchase goods at groceries. Geven et al. [5] tested the interaction of users with NFC enabled services, such as door access, payment and information consulting, which showed that user experience was negatively affected by functional failures, missing feedback and inconsistent interaction models.

The application domains for NFC are also quite diverse. Hoy and Brigham [6] explored alternative uses for NFC in libraries like linking physical items to digital versions, self-checkout and access control. Iglesias and her colleagues [7] tested an health monitoring system based on NFC, which allows users to identify themselves and send health information that is monitored by physicians and caregivers. Broll et al. [8] evaluated the usability and user experience of multi-user interactions on dynamic NFC-displays, providing preliminary results to the design and development of touch-based mobile interactions with large displays for multiple users .

Other studies have investigated mobile systems based on combinations of NFC and other technologies. The authors Ghìron et al. [9] prototyped and tested a ticketing solution for public transport based on NFC and SMS technologies. While W. Chen et al. [10] proposed a mobile payment system for merchant micropayments based on GSM and NFC technologies, which allows the integration with existing Point-of-Sale (POS) equipment and simplifies the integration with existing mobile infrastructures.

Further issues addressing NFC-based mobile services have been researched, such as peer-to-peer systems [11], loyalty models [12], NFC mobile ticketing business models [13], security [14] [15] [16], and factors affecting consumer acceptance of NFC mobile phone services [17].

3 Mobile Payment System

The tested mobile payment system was developed for Android devices with NFC technology and was targeted at the community of the UM Campus where it was evaluated. It allows customers to make payments at the University canteens, cafeterias and student support office. Users can also receive and redeem bus and meal tickets and discount coupons. For this deployment two mobile applications were created (see Figure 1): the User App for customers and the Merchant App for merchants.

A major bank acted as the financial institution, supplying quality accounts for the merchants and customers, identified by a virtual card (VCard). The transactions between devices, and their financial accounts, are brokered and implemented by the MobiPag backend platform [18]. This backend platform has the particularity to support multiple applications since it was designed in a modular and extensible way.

Both customers and merchants used Samsung Galaxy S III devices to perform the payment transactions. This particular smartphone model was chosen because it already contains the SIMAlliance Open Mobile API, allowing the access to Universal Integrated Circuit Card (UICC) applets. Acting as secure element the UICC is capable of storing sensitive information, and performing the appropriate cryptographic operations, assuring the other needed security requirements.

The generic interaction between the customer and the merchant may be described in three steps. They include:

1. Preparation step: both peers prepare the transaction (select the VCard to use);
2. First tap step: the merchant device communicates what the customer is going to buy (quantity and price) and the customer device communicates the coupons and tickets to be redeemed (if any);
3. Second tap step: the customer agrees to the payment and the money transfer

occurs; also both customer and merchant receive a completion notification (i.e. operation successful or not) and a receipt of the transaction is sent to the customer.

In this protocol the payment process is done in two taps and the validation processes are completed in a single tap. To be possible to exchange multiple NFC messages in a single tap, changes were made in the operating system NFC service.

The User App and Merchant App main interfaces are represented in Figure 1. The User App (Figure 1(a)) allows customers to pay for goods and services selecting the "Direct Payment" menu. In this menu, they can choose a VCard and the coupons to redeem. It also allows customers to validate meal and bus tickets, by selecting the "Meal Tickets" and "Bus Tickets" menus respectively. Moreover, customers may check the discount coupons they have stored in their virtual wallet ("Check Coupons" menu) and the transactions they have performed ("Reports" menu).

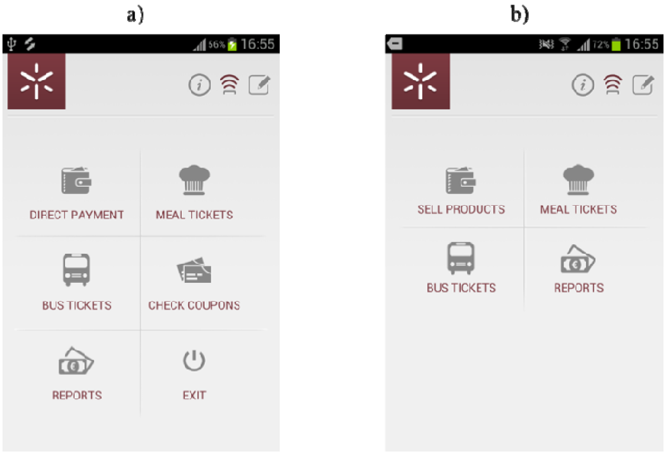


Fig. 1. MobiPag main application interfaces: (a) customer interface; (b) merchant interface

The Merchant App interface is represented in Figure 1 (b). The menu "Sell Products" is used in the University cafeterias. The merchant selects the goods he wants to sell (type and quantity) and the application calculates the total price. This option will accept money and discount coupons from customers' devices. Further discount coupons can be sent to customers during the second tap. With the "Meal Tickets" menu, merchants can sell and validate meal tickets. They only specify which tickets they are going to sell/validate and the quantity. A similar service is performed with the "Bus Tickets" menu. The menu "Reports" shows the performed transactions.

The proposed solution was tested in a two-phase user study and the next section details the methodology used.

4 Evaluation Methodology

The MobiPag solution was evaluated in two phases. The first phase (Laboratory Pilot) involved 9 participants and consisted in performing a set of tasks with the mobile application in a laboratory environment. The main goals of this phase were to test the usability of the application interface and to identify any related problems.

Considering the results of the first test, the application was improved and a new testing phase was initiated. The second phase (Experimental Pilot) involved 16 customers and 5 merchants in real-world scenarios. In this phase we also tested the usability of the interface and collected opinions about the application and about the concept of making payments with mobile phones.

4.1 Selection of Participants

Before launching the pilot to evaluate the MobiPag solution, a survey was applied to the academic community of the University of Minho in order to: 1) present the MobiPag project; 2) assess respondents' perception about mobile payments; and 3) recruit volunteers to participate in the pilot. The survey was applied online (Google docs platform) and was available for a month. During this period 308 responses were collected, from which 112 have expressed interest in participating in the pilot.

A diverse sample of participants in terms of gender, age and occupation (students, professors and non-teaching employees) was recruited, consisting of 9 people for the first phase and 16 for the second. Despite our effort to select a representative sample, it was difficult to maintain the same characteristics of the original group of respondents. For example, it was not possible to find participants available in the age groups of 17-21 years old and 48-52 years old. Male individuals were more willing to participate than female.

For the experimental pilot, we have also recruited 5 merchants for the different payment scenarios: Student Support Office, Cafeteria, Canteen and Bus. Merchants' selection was done considering the payment situations we wanted to test and their availability to participate.

4.2 Experimental Procedure

The User App and the Merchant App were installed in the Samsung smartphones and provided to the participants. Both user study phases comprised three moments:

1. Questionnaire for sample characterization with seventeen structured questions;
2. User test with ten tasks. Tasks correspond to actions performed on the prototype;
3. Interview with nine unstructured and one structured questions.

The main objective of the questionnaire was to characterize the sample in terms of socio-demographic characteristics, smartphone usage, University cafeterias, canteens and buses frequency of use, and mobile payments opinion.

The second moment consisted in performing tasks with the mobile phone. The tasks were defined a priori and represented typical daily transactions, including not only payments but also tickets and discount coupons. So, ten tasks were performed: 1) Open MobiPag application; 2) Buy bus tickets; 3) Check bus tickets stored in the mobile phone; 4) Validate bus tickets; 5) Buy a portion of fruit; 6) Buy meal tickets; 7) Check meal tickets stored in the mobile phone; 8) Check coupons stored in the mobile phone; 9) Validate meal tickets; and 10) Buy coffee (and chocolate).

Throughout the experimental pilot, all prices, products and services were real, but users did not pay with their money, using an account credited with an appropriate initial amount. The threshold for requesting a PIN from users was set to 1 euro, creating situations with and without PIN request.

We observed users during the execution of the tasks to get additional usage data, such as task completion time, verbal comments, and number and types of errors. After performing each task users rated, according to a 5-point Likert scale, how easy it was to perform the task. At the end of the experimental procedure an individual interview was carried out, in order to collet opinions about the application, main difficulties and most interesting features, and comparison with other payment methods.

5 Evaluation Results

The results of the MobiPag evaluation were grouped in three main sections. Section 5.1 describes the main socio-demographic characteristics of phase 1 participants and details the outcomes. Section 5.2 characterizes phase 2 participants and details the results. Finally, section 5.3 give some conclusions of both phases' interviews.

5.1 Phase 1 – Laboratory Pilot

The laboratory pilot took place in a meeting room at UM and was performed by 9 participants and 4 researchers (1 as a merchant). The pilot comprised three moments: questionnaire, user test and interviews. Each test was done individually and lasted about 40 minutes. The participants played the customer role.

Most of the participants (7 in 9) were aged between 27 and 41 years and 6 were male. The majority of participants (5) had a smartphone and almost all respondents found interesting or very interesting to use the mobile phones to pay (88,9%).

Table 1 shows the average time users took to complete each task and the average rating they assigned to tasks. Laboratory pilot users classified almost all tasks as "easy" or "very easy" to perform, except for tasks 2 and 10. These tasks were also those that took longer to be completed. Participants successfully completed 86 tasks in 90 but there were 4 tasks (4,4%) not completed. Two people failed to complete task 10 and another person failed to complete task 2 and 3.

During the experiments some problems concerning the application were identified:

- Menu names were misleading: users selected the wrong menu in 20 tasks (22% of the total tasks);
- Buttons that didn't work: users tried to use buttons that didn't work;
- Lack of feedback: for instance, users didn't understand why and when they received coupons. Users were also confused about the payment process itself:
- Users didn't realize the need of two taps to complete a payment transaction. The 9 users were asked to perform a total of 36 payment tasks and some attempted to purchase with a single tap in 8 of the tasks (22% of the total tasks).
- Users associated mobile payments with purchases over the internet. They thought to be more logical to select first what they wanted and then make the payment.

After this testing it was important to improve the application based on users' suggestions and on our observations. Hence, the names of the menus were changed, in order to become more intuitive, and further feedback messages were included.

Table 1. Laboratory and Experimental pilot: average time and average rate per task

	Phase 1 – Laboratory Group		Phase 2 – Experimental Group	
Tasks	Avg. Time (mm:ss)	Avg. Rate	Avg. Time (mm:ss)	Avg. Rate
T1. Open MobiPag application	00:03	4,6	00:01	5,0
T2. Buy bus tickets	02:11	3,4	00:32	4,6
T3. Check bus tickets stored in the mobile phone	00:21	4,2	00:02	4,6
T4. Validate bus tickets	00:29	5,0	00:15	4,9
T5. Buy a portion of fruit	01:39	4,8	00:47	4,4
T6. Buy meal tickets	00:41	5,0	00:39	4,8
T7. Check meal tickets stored in the mobile phone	00:06	4,9	00:01	4,8
T8.Check coupons stored in the mobile phone	00:04	4,9	00:01	5,0
T9. Validate meal tickets	00:52	4,8	00:30	4,5
T10. Buy coffee (and chocolate)	01:52	2,4	00:51	4,5

5.2 Phase 2 – Experimental Pilot

The experimental pilot consisted on three evaluation sessions in different days and with distinct users. In total, 16 users (customers) and 5 merchants participated in this phase and all payment situations were executed in real contexts. Most of the users (10 in 16) were aged between 22 and 31 years, other 3 were aged from 32 and 41 years and the last 3 were aged between 42 and 46 years. The majority of them (62,5%) were male and 14 of the total 16 had a smartphone. When asked about the usefulness of paying with the mobile phones, 87,5% of users find it interesting or very interesting.

Considering the merchants that participated in the pilot, 2 of them were aged between 26 and 35 years old and 3 were aged between 36 and 50 years old. All of the merchants were male and 4 of them do not have an academic degree. When asked about the level of smartphone usage, 1 merchant answered that they have satisfactory knowledge, 2 have reasonable knowledge and the other 2 do not have any smartphone usage knowledge. When asked about if they were interested in receiving payments through the mobile phone, 80% answered they were interested or very interested.

Merchants involved in the experimental pilot participated on a training session before the experiment. Users also had their training session, in order to explain the payment and validation processes and for them to get used to the application.

According to Table 1 the average time that users took to perform each task decreased dramatically, taking less than a minute to complete them. Users completed successfully all 160 tasks (16 users x 10 tasks per user) and found it “easy” or “very easy” to perform, which means an improvement comparing with the previous phase.

Improvements regarding the previous experiment problems were also observed:

- Misleading menu names: the number of times users selected the wrong menu dropped from 22% to 2,5% (4 tasks in 160).
- Need of two taps: the number of times users attempted to make a purchase with a single tap dropped from 22% to 1,5% (1 in 64 purchasing tasks).

The major problems identified in this testing phase were related to the performance of the NFC connection. Users found it slow and complicated to perform a touch. Sometimes transactions had to be reinitiated, due to NFC connection problems. These impoverished user experience, and led users to assign a lower rating to the tasks.

5.3 Analysis of the Interviews

We applied 9 interviews to the laboratory pilot participants, 16 to the experimental pilot participants and 5 to the merchants. The interviews were recorded and latter transcribed. The information gathered from the interviews was grouped in four main categories: 1) perceived value; 2) security; 3) concerns and issues; and 4) cost.

Perceived Value. Participants have perceived the value of our payment solution in very diverse ways. A first group has mainly focused on the payment itself, highlighting the importance of performance, its value as a cash replacement and the advantages of not having to carry the wallet, cash and credit cards.

It's a way of saving time to people, because it is very fast. It is environmentally friendly, which is very important, it does not use paper. [Lab_Cust9]

On the contrary, another set of participants have referred that they would prefer to pay with cash and, despite being helpful, they would not see the mobile payments application as necessary.

It is helpful, but not necessary (...). And besides, you need to have specific devices that not everyone has or is willing to have. [Lab_Cust1]

While some users seemed to associate value to the possibility of replacing cash, problems with battery duration and with the non-ubiquitous nature of NFC payments at points of sale, will create a situation where this type of value proposition is not enough. Even if deployed in a large scale, mobile payments will have to share the payment space with other payment methods and, at least for a long time, with cash. Therefore, it is important to develop mobile payment systems with unique value propositions that help to differentiate from traditional payment methods. In our case, this additional value proposition was mainly centred on the integration of tickets and discount coupons directly in the payment process. These features were largely appreciated by participants, particularly the possibility of buying tickets, carrying them on the mobile phone and validating them for gaining access to services.

Security. The fact that participants had to introduce a PIN to complete a transaction make them felt secure to pay with a mobile phone. They even compared the system with debit cards in terms of security. Feedback mechanisms like “Operation successfully concluded” messages, beep and mobile phone vibration during NFC connection and after payment completion also increased customers' perception of security.

It is secure, because besides hearing the sound, we receive a message. And then we can see that the transactions were made. [Lab_Cust9]

Concerns and Issues. The integration of tickets and discount coupons in the payment process introduced some complexity in the experience. While the overall procedures were considered simple to perform, coupon redemption and to a lesser extent the use of tickets, were clearly the activities in which participants have experienced more difficulties. The lack of an obvious reference and established practices means that people do not have an understanding about how the system is supposed to work.

Network limitation, battery duration, NFC connection problems, non-ubiquitous nature of NFC payments at points of sale and high cost of NFC equipment also raised some concerns regarding mobile payments adoption.

Cost. Opinions about the willingness to purchase the application were not consensual among the participants. Some were able to pay for it, if the price was very low and if the application was useful for several daily payment situations. Though, others considered that the cost should be supported by merchants.

On the other hand all merchants were willing to pay for the application, perhaps because they already pay for other payment methods and for billing software.

6 Discussion and Conclusions

The first general observation is the very positive attitude that participants, both customers and merchants, have shown in regard to the use of the technology. While we cannot extrapolate this into acceptance of the technology in real world daily usage, we can assert that people are open to the new possibilities offered by mobile payments and are willing to try how they can use them in realistic payment situations.

The design of mobile application interfaces has also been proven as an important issue to take into account in the setup of NFC-based payment deployments. Carefully choosing terminology, buttons and symbols meaning or providing feedback to users are examples of user interface design issues that may greatly influence application usability and system performance. Furthermore, customers' perception of security may be increased by adding PIN requests and system feedback such as system messages about the state of transactions or beeps and phone vibration.

The choice of the technology in which the system will rely on is very important. NFC technology is relatively new and demonstrated not to be stable yet. Potential NFC connection problems as well as wireless network connectivity issues may lead to users' frustration and negatively influence their perception of the system.

Findings also suggest that such a system may be suitable to be used in payments, but not for validation scenarios, since it proved to be much more time consuming than the current validation process. Obviously some improvements could be implemented regarding the mobile validation, yet hardly as efficient as the current system.

Finally, another issue was identified during the experiments, which may be seen as a fundamental challenge that any payment system may have to address. The added flexibility leading to a myriad added services integrated with payments is simultaneously one of the strongest points and one of the biggest challenges in interaction design. What could be a simple payment procedure may become very complex from a cognitive perspective when multiple types of coupons, discounts and loyalty schemes become part of the payment process. Further research must be done in order to analyse the impact of complex mobile service solutions on users' mental model. Also, a larger deployment of an improved version of the solution should be considered in order to overcome the explorative nature of the present study.

References

[1] Au, Y., Kauffman, R.: The economics of mobile payments: Understanding stakeholder issues for an emerging financial technology application. Electron. Commer. Res. Appl. 7(2), 141–164 (2008)
[2] Coskun, V., Ozdenizci, B., Ok, K.: A Survey on Near Field Communication (NFC) Technology. Wirel. Pers. Commun. 71(3), 2259–2294 (2012)
[3] Massoth, M., Bingel, T.: Performance of Different Mobile Payment Service Concepts Compared with a NFC-Based Solution. In: 2009 Fourth Int. Conf. Internet Web Appl. Serv., pp. 205–210 (2009)
[4] Rehman, S., Coughlan, J.: An Efficient Mobile Payment System Based On NFC Technology. World Acad. Sci. Eng. Technol. 78, 1695–1698 (2013)
[5] Ferro, B., Tscheligi, M.: Experiencing Real-World Interaction – Results from a NFC User Experience Field Trial. In: 9th Intl. Conf. on Human Computer Interaction with Mobile Devices and Services (MobileHCI 2007), pp. 234–237 (2007)
[6] Hoy, M.B., Brigham, T.: Near field communication: getting in touch with mobile users. Med. Ref. Serv. Q. 32(3), 351–357 (2013)
[7] Iglesias, R., Parra, J., Cruces, C., de Segura, N.G.: Experiencing NFC-based touch for home healthcare. In: Proc. 2nd Int. Conf. PErvsive Technol. Relat. to Assist. Environ., PETRA 2009, pp. 1–4 (2009)
[8] Broll, G., Vodicka, E., Boring, S.: Exploring multi-user interactions with dynamic NFC-displays. Pervasive Mob. Comput. 9(2), 242–257 (2013)
[9] Ghiron, S.L., Sposato, S., Medaglia, C.M., Moroni, A.: NFC Ticketing: A Prototype and Usability Test of an NFC-Based Virtual Ticketing Application. In: 2009 First Int. Work. Near F. Commun., pp. 45–50 (February 2009)
[10] Chen, W., Hancke, G.P., Mayes, K.E., Lien, Y., Chiu, J.-H.: NFC Mobile Transactions and Authentication Based on GSM Network. In: 2010 Second Int. Work. Near F. Commun., pp. 83–89 (2010)
[11] Balan, R.K., Ramasubbu, N., Prakobphol, K., Christin, N., Hong, J.: mFerio: The Design and Evaluation of a Peer-to-Peer Mobile Payment System, pp. 291–304 (2009)
[12] Ozdenizci, B., Ok, K., Coskun, V.: NFC Loyal for Enhancing Loyalty Services Through Near Field Communication. Wirel. Pers. Commun. 68(4), 1923–1942 (2012)
[13] Juntunen, A., Luukkainen, S., Tuunainen, V.K.: Deploying NFC Technology for Mobile Ticketing Services – Identification of Critical Business Model Issues. In: 2010 Ninth Int. Conf. Mob. Bus. 2010 Ninth Glob. Mobil. Roundtable, pp. 82–90 (2010)
[14] Pasquet, M., Reynaud, J., Rosenberger, C.: Secure payment with NFC mobile phone in the SmartTouch project. In: International Symposium on Collaborative Technologies and Systems, pp. 121–126 (2008)
[15] Hu, J.-Y., Sueng, C.-C., Liao, W.-H., Ho, C.C.: Android-based mobile payment service protected by 3-factor authentication and virtual private ad hoc networking. In: 2012 Comput. Commun. Appl. Conf., vol. 1, pp. 111–116 (January 2012)
[16] Kadambi, K.S., Li, J., Karp, A.H.: Near-field communication-based secure mobile payment service. In: Proc. 11th Int. Conf. Electron. Commer., ICEC 2009, p. 142 (2009)
[17] Chen, K.-Y., Chang, M.-L.: User acceptance of 'near field communication' mobile phone service: an investigation based on the 'unified theory of acceptance and use of technology' model. Serv. Ind. J. 33(6), 609–623 (2013)
[18] Consortium, M.: MobiPag - Relatório Técnico Final (2013)

Augmented Reality Mobile Tourism Application

Flávio Pereira[1,2], Daniel Castro Silva[1,3], Pedro Henriques Abreu[1], and António Pinho[2]

[1] UC-DEI – University of Coimbra - Department of Informatics Engineering / CISUC – Centre for Informatics and Systems, University of Coimbra
Pólo II, Pinhal de Marrocos, 3030-290 Coimbra, Portugal
fpereira@student.dei.uc.pt, pha@dei.uc.pt

[2] Ubiwhere, Lda.
Rua Cristóvão Pinho Queimado, 47, 3800-129 Aveiro, Portugal
apinho@ubiwhere.com

[3] FEUP-DEI – Faculty of Engineering, University of Porto – Department of Informatics Engineering / LIACC – Artificial Intelligence and Computer Science Laboratory
Rua Dr. Roberto Frias s/n 4200-465 Porto, Portugal
dcs@fe.up.pt

Abstract. The growing use of smartphones has revolutionized the way people live, fostering the use of mobile application in the most diverse situations. These applications take advantage of the mobile device's capabilities to provide the user with useful and contextualized information, being equipped with increasingly intuitive interfaces, and offering richer contents in an attractive manner. Augmented reality emerges as one of the technologies that can be used in these applications, allowing for an improved user experience. This paper describes a tourism-oriented mobile application, in this case to be used in a botanical garden, which uses current mobile device's capabilities to provide the visitor with several ways to obtain the desired information. The results obtained from this application are shown, including images of the implemented features, and highlighting the results related to the use of augmented reality.

Keywords: Mobile Application; Augmented Reality; Usability.

1 Introduction

The proliferation and growing use of smartphones have revolutionized the way we live today. There has been a great technological development in this area, which is becoming more accessible to the ordinary citizen. Consequently, this has led to an increased use of mobile applications in the most diverse situations and areas. These applications take advantage of the capabilities of mobile devices (accelerometers, GPS, 3G connectivity, among others) to provide useful and context information to the user in a ubiquitous manner. Using the pervasive and ubiquitous computing, it becomes possible to obtain information from the environment and use it to configure or adjust the application and, consequently, meet the needs of users.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_17,

These applications are increasingly present in everyday life of smartphones users, and have been used in the most diverse situations. Accordingly, the use of these applications require interaction between the application itself on the mobile device and the user. In order to make this interaction more enjoyable, applications must possess robust and usable interfaces, providing information to the user intuitively. Human-Computer Interaction (HCI) is the area of computing which investigates the design, evaluates and implements interfaces for humans to interact with computer systems in an efficient and intuitive way [1]. A Graphical User Interface (GUI) is one of the approaches to HCI, which allows direct manipulation of objects or menus on a screen [2]. Another important approach to HCI is to use augmented reality, to provide users with information in a more appealing manner.

Augmented reality is a technology that combines elements of the real world with virtual elements in three dimensions (3D), allowing interaction between objects (real and virtual) in real time. Although research on augmented reality has advanced in recent decades, this technology has yet to reach the mass market. The minimum requirement to achieve augmented reality is a device (e.g., a smart phone), a camera and a processing unit. There is a wide range of areas for augmented reality applications. In addition to the medical, paleontology or gaming areas [7, 8, 9], augmented reality applications can be used in museums, galleries or open spaces (eg gardens) where objects such as paintings, sculptures, artifacts, ruins, plants, among others, can be "enhanced" with information (texts, images, sounds or videos) to be shown via a device when they approach the object. One of the areas where augmented reality is of great value is tourism. By using this technology, the user can get information about a particular place with virtual information and virtual images (three-dimensional) for example in natural parks or gardens. Furthermore, it is also possible to obtain georeferenced information with augmented reality, taking advantage of the potential of GPS devices of mobile devices [3].

With the help of augmented reality systems, the users can get relevant information needed for the enrichment of their culture. It’s possible to say that augmented reality provides an improvement in the perception and interaction in the combination of the virtual world with the real world. In addition, the user’ experience is improved using this type of technology [6].

In this article we propose the use of augmented reality in a mobile application for the Botanical Garden of the University of Coimbra, using the Android operating system. The application in question allows for the visualization of multimedia content on various plants present in the garden. The use of augmented reality becomes an alternative to information visualization in two dimensions, improving the user's visit.

This document is structured as follows: Section 2 presents some works related to the one proposed here. Section 3 presents the system architecture, as well as key features for the mobile application. In Section 4, the test scenario is presented and in Section 5 some results are presented. Finally, Section 6 summarizes the developed work and presents some lines of future work.

2 Related Work

This section provides an overview of related projects and applications.

America's National Parks[1] is an application that provides information about each of the 58 national parks in the United States. This application tells the user how to get to a particular park, provides a map of each park and topological maps for hiking. Once loaded, the walks can be used without network connectivity. Figures 1.a and 1.b show images of the application. The leftmost image show the use of 2D maps and the provision of user location within the park, while the second image shows the use of the multimedia gallery.

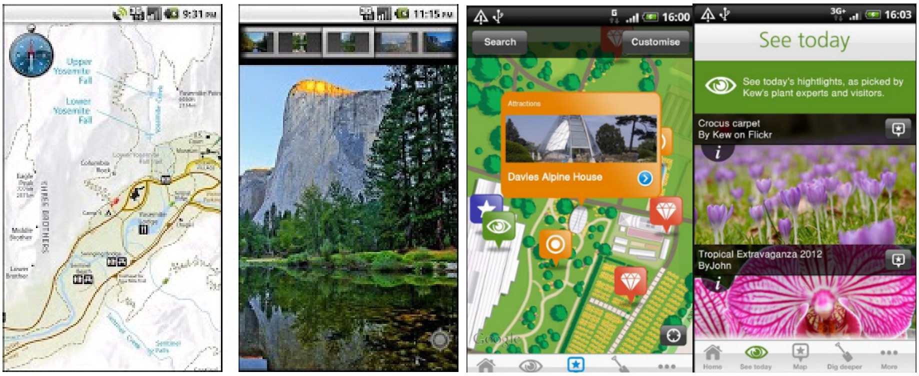


Fig. 1. America's National Parks (a and b) and Kew Garden Application (c and d)

Kew Gardens[2] is a smartphone application that allows for the discovery of plants within Kew Gardens, in London, as well as finding greenhouses and discovering hidden treasures. It provides the user with an interactive map of the garden, also allowing him to customize the map to display only what he wants. Another interesting feature involves the use of a QR code scanner, allowing the user to scan codes beside the plants in order to obtain detailed information. The user can leave his feedback of the various attractions using a feedback system. The application integrates with social networks such as Facebook and Twitter, allowing a visitor to share his experience with his friends. It is possible to see in Fig. 1.c and 1.d some images of this application. In Fig. 1.c we can see the map of the application containing the various points of interest in the Garden, represented by icons. Figure 1.d shows photos about plants, taken by visitors or added by the managing staff of the Garden.

GeoTravel Guide[3] is a mobile tourism guide with augmented reality, which can be used anywhere. With this application, it is possible to use augmented reality to display

---

[1] `https://play.google.com/store/apps/details?id=com.lbs.android`

[2] `http://www.kew.org/visit-kew-gardens/visit-information/garden-guides/mobile-app/index.htm`

[3] `http://www.augmentedworks.com/app/travel-guide-with-ar-augmented-geotravel`

the desired points of interest as well as the name and the distance to a chosen point of interest. Figures 2.a and 2.b show images of the application. The left image shows the use of augmented reality in the application, and the location of points of interest in the camera can be seen, as well as the distance to them. Figure 2.b shows the possibility of downloading Wikipedia articles about points of interest, as well as saving them for offline use.

Wikitude Augmented Reality Travel Guide[4] is an augmented reality mobile guide based on context and location. It uses the GPS receiver and compass to determine the position and orientation and, using a global database containing thousands of points of interest, displays graphics and text with information about the image captured by the camera. The user is also informed of the distance to the point of interest – see Fig. 2.c and 2.d.

Fig. 2. GeoTravel Guide application (a and b) and Wikitude application (c and d)

After analyzing the applications presented above, it appears that each contains components that can be used in the Botanic Gardens Applications. Kew Garden and America's National Parks offer the ability to view 2D maps of the venues in which it operates as well as multimedia elements. GeoTravelGuide offers the possibility of viewing points of interest as well as the distance to each through augmented reality. The Botanical Gardens application shall combine elements used in these applications. With it, it is possible to view data across interfaces in 2D as well as through an interface using augmented reality. Comparing with the Kew Garden application and America's National Parks, the most similar to the proposed application, the introduction of augmented reality becomes a differentiating component.

3 System Architecture

This section addresses the complete architecture of the system. Performing a high-level analysis first, the system architecture follows a client-server paradigm. The client is the mobile application (using the Android operating system), and the server

[4] `http://www.wikitude.com/app/`

provides a REST API, exchanging data with the mobile application. The two components exchange data via REST requests, using JSON format for the data.

The mobile application works offline, and network connection is only required to update the content in the application, allowing it to be synchronized with the data in the server database. For this, whenever the user connects to the internet, the application detects the network connection and informs the user of the need to upgrade the contents of the application. The synchronization becomes therefore a significant requirement for the application, and data exchange should be fast. Another feature that requires a network connection is data sharing over social networks. The server component has a REST API, which allows the application to download the available content as necessary. Figure 3 shows the described architecture. The mobile application has several modules, each grouping a set of functionalities of the application. More details about each module can be found in section 3.2.

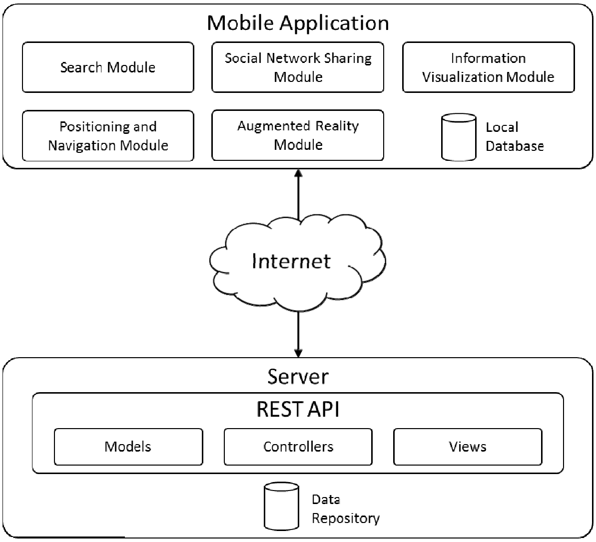


Fig. 3. System architecture

3.1 Synchronization

One of the mechanisms identified and implemented for data synchronization consists in storing a server-side tag that represents the version of the data the server stores as well as keeping the same tag in the client in order to check for data consistency between server and application. When the application detects a network connection, it sends a message to the server requesting the tag, invoking a method from the API. The server sends the tag to the mobile application and it checks whether or not synchronization is required. After synchronization, the tag on the client is updated to the latest version.

This mechanism works well for transferring small-size text content. However, for exchanging images between server and application, the solution was not as suitable, as sending all images from the server to the application would take a long time in case of large contents. Therefore, for sending images, the server sends the application a list of the images in the server. When the mobile application receives the list, it compares the names of the contents with the names of the contents present in the list of images of the application and transfer only the images that are not listed on the client. With this mechanism, only the new images that have been added on the server are sent to the application, minimizing the overhead in terms of time and traffic, when compared to the case where all pictures are sent to the client application.

3.2 Main Application Features

The features presented in this section are grouped by modules in order to identify and differentiate the application components. Some images of the described features are presented in Section 5.

Positioning and Navigation: Using the application, the user can access the map of the Garden, and view in it his position within the Garden. For this, the mobile application takes advantage of the potential of the mobile GPS equipment. In addition, the user can view in the map of the garden the plants that surround it. The map is also used to allow the user to take a certain route inside the Garden or search for the location of a specific plant. Using the routes component, a user can choose a route and follow it, visualizing only plants of the route in the 2D map or the augmented reality component. There are pre-defined routes for the application, but the user can also create his own routes with the plants he wishes to visit.

Information Visualization: This module refers to the visualization of all the information the application provides to the user about existing plants. Using the application, the user can access a list of plants. For each plant, the visitor can view details and multimedia content such as text or images. In order to make the application more attractive, thus improving the user experience, it is possible to view information about the plants using augmented reality. Turning on the camera of the mobile device, the user can visualize the location of the various plants in the garden, as well as some information about them. With this, the user can be guided to a particular plant, using augmented reality. The aspect of georeferencing in augmented reality is explored in this module. Through the GPS coordinates of plants and GPS equipment of the mobile device, the location of a plan is shown to the user as well as the distance to it. The mobile device keeps on its database the existing Garden plants and consequently the GPS coordinates of the plants. When the user moves, the presented distance to a plant in changed in real time. This procedure can be seen in Section 5, in the obtained results.

Social Networks Sharing: using the application, the user can share data about his visit to the Botanical Garden in three social networks: Facebook, Twitter and Google+. This is considered an important feature of the developed application, due to the ever-growing use of social networks, and constitutes both a feature appreciated by users and a means of divulging the application through the community.

Search Module: This module allows the user to search for specific plants, after which he can either visualize information regarding said plant (be it text or other media content), or set the location of the plant as a destination point of a route to be used by the navigation module.

4 Experimental Setup

As mentioned above, the application is designed to be used in the Botanical Garden of the University of Coimbra, which has hundreds of plants and little information on most of the plants in the Garden. Therefore, to conduct the tests for the application about 30 plants in a specific area of the garden were considered.

Two types of tests were defined to validate the developed application. The first type concerns the use of the application to check the proper functioning of the implemented features so as to verify the usability of the application. The second type relates to the use of augmented reality component inside the Garden. For this, the user must walk through the garden, choosing to use the augmented reality option. Once this option is turned on, the user must move towards a plant that will appear on the screen and verify that the element followed in the camera coincides with the right plant. It should be noted that the results for this test depend largely on the accuracy of the mobile device GPS.

5 Results

This section presents the results obtained during the tests. Some images of the application are presented for some of the most relevant components.

5.1 Results from the Information Visualization Module

Figures 4, 5 and 6 show the results obtained from the information visualization module of the application. It is possible to see in Fig. 5 (left) that for each plant there is a description as well as images. If the user wishes to view the location of the plant within the Garden, he could press the "Find" button in order to view a map with the location of the plant in the garden, as well as the position of the user at every moment. With this, the user can easily move to reach the place where the plant is located at. The figure on the right represents the list of all plants that exist in the garden.

Analyzing the information visualization with the augmented reality component, it is possible to see in Fig. 4 and 6, through the camera of the mobile device, that the user can see where a particular plant is within the Botanical Garden and therefore can

Fig. 4. Location of plants belonging to a route, using Augmented Reality

be guided to that plant. It can also be observed that the distance between the user and the plant is provided to the user. In addition, when pressing a plant in the view, detailed plant information is provided. The information present in the camera of the device depends in part of the accuracy of the GPS mobile device, since information is georeferenced at each instant, depending on the position and orientation of the user.

5.2 Usability Tests

Usability tests are needed to identify shortcomings and improvements in an application of this kind. The solution to test the degree of satisfaction of users as well as the usability of the application tests was using a System Usability Scale (SUS) test, a questionnaire consisting of 10 questions with a five-value scale, five of them worded positively and the other five negatively [4]. The questionnaire result can determine the quality level of the application in terms of usability. After collecting and analyzing the responses of twelve users, an average score of 80 was obtained. For a usability test of this type, the average rating is set to 68. Any result below this level is considered below average and any results above this value is considered above average [5]. Therefore, the obtained result (80) is considered to be good. Based on this test and the obtained results, it can be considered that the application provides good results in terms of usability and user satisfaction. The 12 participants were between 19 and 24 years of age. Users who had already visited the Botanical Garden without the use of any application during their visit were chosen. Some of the users had already made use of mobile applications related to tourism in other situations, thus making this user group diverse in terms of experience with mobile tourism-related applications, and representative of prospective users of the application.

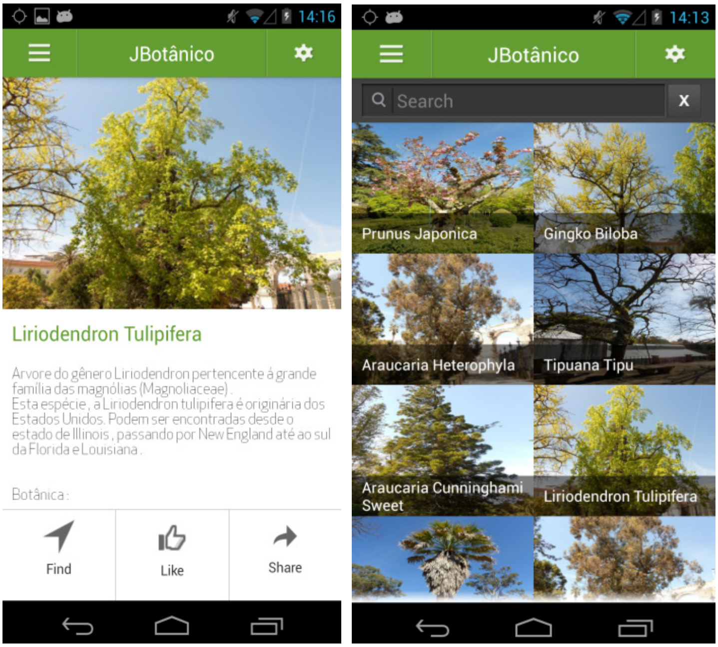


Fig. 5. Screenshot for the information visualization

5.3 Augmented Reality Accuracy Tests

During the tests, particular attention was given to the results obtained for the augmented reality component. It was important that the application was able to display the position of a plant as accurately as possible. At best, the icon representing a plant should completely coincide with the corresponding plant. Of course, the results obtained depend on the accuracy of the collected data regarding the location of each plant, as well as the precision of the GPS on the mobile device being used. It was found that, on average, there was a precision error of 3 to 4 meters from the data presented by the application.

Fig. 6. Representation of all plants, with Augmented Reality

6 Conclusions and Future Work

In this paper we discuss the use of augmented reality in a mobile application for tourism, offering a set of features that enhance the user experience during a visit to a Botanical Garden. The application has features similar to those presented in related applications but extending and bringing them together, thus providing a unique application that presents all features that are valued in such an application, such as support for augmented reality, or integration with social networks.

Using the navigation capabilities of the application, the user can follow a given route, be it in a 2D map of the garden or using the augmented reality mode, visiting a set of plants. For each plant, information is provided by means of text and other media (images and/or videos), which also serves as a means to educate the visitor in a more appealing manner than more traditional non-interactive forms.

The application was designed having in mind that the interface is a major concern, as to improve user experience, and increase user satisfaction when using the application. Tests performed with a group of twelve volunteers show that these goals were attained, having achieved a score of 80 in the SUS scale, which can be considered a good result.

After developing this application, it can be seen that it is possible and interesting to improve the augmented reality component present in the application. Some possibilities still being considered involve the use of image recognition techniques or pattern recognition, to allow for the recognition of plants from the camera of the mobile device, showing the plant in 3D in the device. With this, the usability of the application would improve, making the use of the application more attractive to the user, enhancing his experience during his visit to the Botanical Garden.

Acknowledgments. This work is partially supported by iCIS project (CENTRO-07-ST24-FEDER-002003) which is co-financed by QREN, in the scope of the Mais Centro Program and FEDER.

References

1. Teixeira, F.G., Santos, S.L.: Designing an Interface for Three-Dimensional Interaction with a Focus on Usability and Graphical Performance, pgdesign, Porto-Alegre, pp. 39–50 (2010) (in Portuguese)
2. Garbin, S.M.: Study on the Evolution of Human-Computer Interface. Master's Thesis, São Carlos School of Engineering, University of São Paulo, São Carlos, Brazil (2010) (in Portuguese)
3. Henrysson, A.: Bringing Augmented Reality to Mobile Phones. Dissertations. Department of Science and Tecnology Linkopings Universitet, Norrkoping, Sweden (2007)
4. Brooke, J.: SUS – A Quick and Dirty Usability Scale. In: Jordan, P.W., Thomas, B., McClelland, I.L., Weerdmeester, B. (eds.) Usability Evaluation in Industry, pp. 189–194. Taylor & Francis (1996)
5. Lewis, J. R. and Sauro, J.: The Factor Structure of the System Usability Scale. In Proceedings of the 1st International Conference on Human Centered Design (Held as Part of HCI International 2009), San Diego, CA, USA, pages 94–103,Springer-Verlag (2009)
6. Oliveira, I.A.: User Interface: Human-Computer Interaction Over Times. Olhar Científico Journal 1(2), 178–184 (2010) (in Portuguese)
7. Lee, J.-D., Huang, C.-H., Lan, T.-Y., Lee, S.-T., Wu, C.-T.: A Medical Augmented-Reality System for Image-Guided Surgery using Marker-Added ICP. International Journal of Innovative Computing, Information and Control 7(11), 6523–6539 (2011)
8. Silva, D.C., Vinhas, V.: An interactive Augmented Reality Battleship Game Implementation. In: Taisch, M., Cassina, J. (eds.) Proceedings of Learning with Games 2007, Sophia Antipolis, France, September 24-26, pp. 213–219 (2007)
9. Abreu, P., Mendes, P.: Mastermind: an Augment Reality Approach: Portinga Legacy Game to new Interaction Paradigms. In: Tsekeridou, S., DavidCheok, A., Giannakis, K., Karigiannis, J. (eds.) DIMEA. ACM International Conference Proceeding Series, vol. 349, pp. 205–210. ACM (2008)

MusE Central: A Data Aggregation System for Music Events

Delfim Simões[1], Pedro Henriques Abreu[1], and Daniel Castro Silva[1,2]

[1] UC-DEI – University of Coimbra - Department of Informatics Engineering / CISUC – Centre for Informatics and Systems, University of Coimbra
Pólo II, Pinhal de Marrocos, 3030-290 Coimbra, Portugal
delfim@student.dei.uc.pt, pha@dei.uc.pt
[2] FEUP-DEI – Faculty of Engineering, University of Porto – Department of Informatics Engineering / LIACC – Computer Science and Artificial Intelligence Laboratory
Rua Dr. Roberto Frias s/n 4200-465 Porto, Portugal
dcs@fe.up.pt

Abstract. Along with the evolution of the Internet and its ever-increasing use, a significant trend to develop solutions to provide information is noticeable. This scenario increases information decentralization in certain contexts. Nowadays, information regarding music is available online in several online locations. In this context, this paper reports on the development of a web platform that centralizes existing information regarding events, retrieving other contextually related data available. Validated with both information retrieval quality and interface usability metrics, this project attained a more effective and complete concert search service according to the data provided to the user. It has revealed itself capable of retrieving information on more events, when compared to other platforms without a data centralization approach. According to the usability survey results attained, such as an 89 SUS scale score, it was proved that the developed service is provided with a simple and intuitive interface as well.

Keywords: Information aggregation, musical events search, usability, web development, web-scraping.

1 Introduction

Nowadays, it is possible to find music data in many web platforms. Focusing on concerts, there are many applications providing information, which increases data decentralization, since their knowledge base may vary. Apart from that, the degree of information presented to the users may also vary according to the platform, making it sometimes difficult to conveniently inform a user about some concerts. To mitigate this problem and improve the efficiency and quality of online concert searching, we propose the development of a web solution that centralizes existing data on the Web from several sources and provides it through a single platform. Some web platforms already integrate concert data retrieved from external sources, but they usually rely on a single source of information. This approach may enhance the service provided by

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_18,

the platform, but does not improve the online concerts search, since it doesn't achieve an effective data centralization. As for the information degree variance among the several information providers, some of these platforms already aim to provide a more complete event searching service, by providing contextual data related to events. Nevertheless, many of these platforms present little context-related data.

Along with the evolution of the Web, it has become very common to find platforms that retrieve information from external sources, to provide an enhanced service. Several scientific studies and web platforms have been developed, covering several data retrieval approaches. The news information context is one example where such advance has been verified. A very popular information retrieval approach over this context of data is RSS (Really Simple Syndication) feeds, usually provided by news platforms, whose content is constantly updated, and that allow for data to be extracted and processed [1]. There are also popular solutions, such as Google News, that gather news from well-known providers, centralizing the information for the user. In the music context, there are also some services based on data retrieved from other platforms. Several techniques can be used to exchange data between entities, one of them being the integration based on RESTful web services, where a system designed according to REST (Representational States Transfer) architecture [2] provides services that may be invoked to perform actions. It is possible to invoke it through a specific interface, its API (Application Programming Interface) [3]. These approaches also allow the implementation of a very common design pattern, the Mashup, which consists in aggregating resources from several sources, mixing them to achieve something new, such as a more complete service [4]. Web-scraping techniques are alternative approaches of information retrieval, aiming to extract information directly from web pages, by manipulating its source code or even emulating web page interactions. Screen-scraping is one of these techniques, which allows parsing an interface's views such as a platform's web pages [5]. It is commonly used to retrieve information from platforms that do not provide APIs or platforms whose APIs does not give access to the whole information provided on the platform's web pages. This technique might be complex to implement, since it is very dependent on the page's source code. Also, changes to a page internal structure might lead to malfunctions on the implemented process and consequently require adjustments.

In this paper, we describe the development of a web platform that could mitigate both these problems, by focusing on information retrieval and centralization. Using this data, the platform is able to present upcoming events to the user, according to a specified location or artist. Apart from event's general details, additional contextual data is provided, covering artists (biographies, repertoire, lyrics, videos and merchandising), venues (hotspots, weather forecasts, itineraries and transports) and social networks hype, providing a very complete event divulging service.

This article presents the following structure: section 2 presents the analyzed platforms for concert divulging and the integrated data sources. Section 3 details the platform's development methodology and section 4 addresses the platform's validation. On section 5 the achieved results are discussed, and finally section 6 presents the conclusions and lines for future work.

2 Concerts Divulging Platforms and Data Sources

A comparative analysis of several existing platforms for divulging musical events is shown in Table 1. It covers several functionalities identified as relevant: (1) Concerts divulging (scheduled concerts); (2) Ability to purchase tickets; (3) Presentation of artist's biography and repertoire; (4) Artist's videos; (5) Artist's lyrics; (6) Purchase/download artist's repertoire material; (7) Artist's related merchandise; (8) Interactive map with concert location; (9) Directions to concert location; (10) Available transports for concert location; (11) Reference to hotspots near concert location; (12) Weather forecast for concert location; (13) User accounts management and social interactions. This analysis provided a knowledge of the state of the art of events divulging platforms, and identified the currently followed approaches.

Table 1. Comparison of analyzed concerts divulging platforms

Solution	1	2	3	4	5	6	7	8	9	10	11	12	13
Everything Is New	✓	✓		✓									
Yeaaaah	✓			✓				✓					✓
Festivais de Verão	✓							✓					✓
TicketLine	✓	✓						✓		✓			
TicketLine (UK)	✓	✓	✓					✓	✓	✓			
Bilheteira Online	✓	✓						✓					
Palco Principal	✓		✓										✓
Songkick	✓	✓		✓				✓					✓
Artistas & Espectáculos	✓												
Live Nation Entertainment	✓	✓					✓	✓			✓		✓
Last.fm	✓	✓	✓	✓		✓		✓					✓
MySpace	✓	✓	✓	✓		✓		✓					✓
BLITZ	✓		✓	✓	✓								✓
MTV	✓		✓			✓							
Fnac	✓	✓											
Myway	✓		✓			✓							✓
IOL Musica	✓							✓					
ReverbNation	✓	✓	✓	✓		✓		✓					✓
Cotonete	✓		✓		✓								✓
Guia da Cidade	✓							✓			✓		
SetList.fm	✓												✓

All these platforms share common characteristics on concerts divulging approaches and are similar on functionalities; however, it is noticeable that none of these matches all the criteria identified as relevant. The developed platform aims to surpass this limitation, by providing a more complete service and relying exclusively on information existing on the Web. With the goal of retrieving and centralizing external information, there was a need to define which informative contexts were relevant to integrate on the platform. Several contexts have been identified as relevant, covering information about events, associated artists or venues. After identifying these contexts, several platforms were analyzed in order to identify possible data sources for each one. Due to some of the characteristics inherent to the information retrieval approach used (APIs and their terms of use), the quantity and quality of data provided (analyzed through tests), or the platforms' relevance, not all analyzed platforms were integrated. Figure 1 presents the data sources identified as relevant for the platform, highlighting the ones integrated for each of the informative contexts identified.

3 Development Methodology

The proposed platform needs to support constant human interaction and logical operations, such as retrieving, processing and storing data. Therefore, its architecture, shown in Fig. 2, was defined based on the choice of developing the platform with Play![1], a framework that aims to ease Java web apps development. There are 3 core components: models, views and controllers, derived from the MVC[2] design pattern.

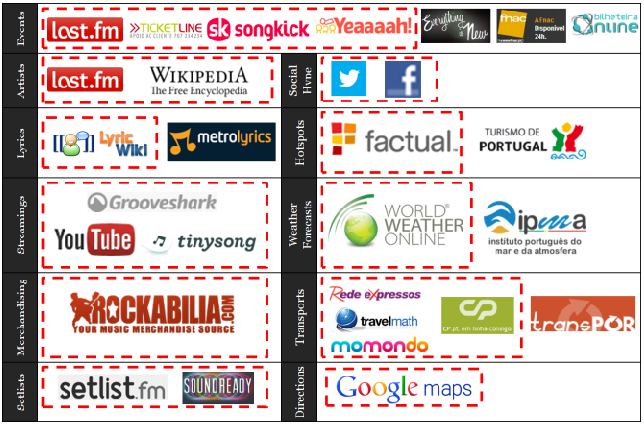


Fig. 1. Contextual information sources

Controllers are responsible for handling user requests, which may trigger responses at the usability or logical level. To ensure the robustness of the platform, an auxiliary data aggregator module was created for each addressed information context. It should be noted that there are two cron jobs, responsible for invoking the transports' data aggregator, as to pre-aggregate necessary data. Views comprehend the several interface screens, designed using Play!'s Scala template mechanism and tools such as jQuery and Bootstrap, attaining simple and intuitive interface, striving for consistency among screens, which is considered to be a golden rule of interfaces design [6]. Models are a relevant part of the platform's logical on managing and processing

[1] More information available online at: `http://www.playframework.com/`

[2] Model-View-Controller (see `http://msdn.microsoft.com/en-us/library/ff649643.aspx`)

information. Several data models, with their own attributes and inter-relations, were defined to handle all information retrieved from the several integrated platforms.

4 Validation Methodology

To validate the developed platform, two areas were covered: functionality and usability. As for the information retrieval performance, several parameters, shown in Table 2, were defined according to information retrieval IQ (Information Quality) metrics based on the Precision and Recall model. In this model, used for IR and NLP (Natural Language Processing), Precision measures the reliability of data extracted and that Recall measures the amount of relevant data that is correctly extracted [7].

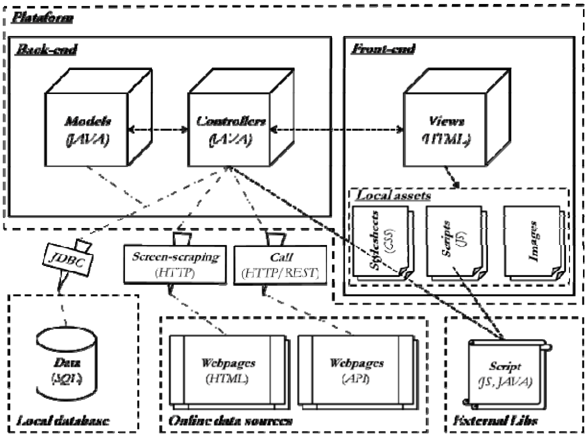


Fig. 2. Implemented architecture

As for the platform's responsiveness, it was measured in terms of the time (in seconds) that the controllers took to respond to the user's requests. In order to assess the platform's functionalities performance, several test scenarios were defined, covering all functionalities (each of these scenarios was tested three times). Regarding the usability validation, several approaches were followed, focusing on the conformity with design guidelines and user reviews and opinion. To understand potential user's opinion, two usability surveys were conducted, as shown in Table 3.

Table 2. Information retrieval parameters

Parameter	Description
Reliability	Percent of reliable entities (excludes duplicates) from aggregated entities;
Relevance	Percent of relevant aggregated entities from source existing entities;
Integrity	Percent of aggregated entities comprehending a minimum level of data;
Processing Time	Time (in seconds) that entities' aggregation process lasts.

Table 3. Usability surveys

Survey	Description
SUS (System Usability Scale) based	SUS is a well-defined scale and recognized as a valid metric for evaluating general systems' interface usability; It allows users to evaluate their conformity with 10 statements, according to a 5 (or 7) Likert-scale conformity levels, resulting in a score scale [8].
CSUQ (Computer System Usability Questionnaire) based	CSUQ is a standardized questionnaire used for assessing a system's interface usability; It allows users to evaluate their conformity with 19 statements, according to seven Likert-scale conformity levels; It shares some of the SUS features, by relying on conformity evaluation; however, it evaluates more statements, comprehending different points of view [9].

5 Results and Discussions

The several implemented functionalities were validated; however, two of them should be highlighted: searching for events scheduled for a location and viewing artist information. Table 4 presents the defined test scenarios for these two functionalities.

Location-Based Events Search. Apart from depending on retrieving external data, this functionality also requires a mash up of data aggregated from several platforms. Table 5 presents some statistical data regarding events information retrieval for the three test scenarios, as well as the results attained for the performance parameters.

Table 4. Events search data aggregation evaluated parameter results

Scenario	Description
S_LOC_EV-1	Event search for Coimbra, Portugal (expected reasonable amount of events).
S_LOC_EV-2	Event search for Porto, Portugal (expected high amount of events).
S_LOC_EV-3	Event search for Paris, France (expected high amount of events).
V_ARTINFO-1	View Dead Combo artist's information.
V_ARTINFO-2	View Depeche Mode artist's data.

Table 5. Events search data aggregation statistics and evaluated parameter results

Scenario (functionality)	Existing events	Events not retrieved	Different events retrieved	Unin-tegral events	Proces-sing time (seconds)	Relia-bility	Rele-vance	Integrity
S_LOC_EV-1	49	0	45	13	61.453	91.84%	100.00%	71.11%
S_LOC_EV-2	237	3	166	43	274.533	70.94%	98.73%	74.10%
S_LOC_EV-3	2633	0	2073	52	929.686	78.73%	100.00%	97.49%

As can be seen, the number of existing events varied according to the test scenario, as expected. Apart from one scenario, where data regarding three existing events was not aggregated (decreasing data relevance) and the existence of some duplicated events among the data sources (reliability reduction), overall, information on a good amount of distinct events was retrieved. It is worth noting the overall good integrity level of the retrieved events, meaning that most of them had information regarding its artists, venue and date, which was crucial to retrieve additional contextual data. As for processing time (which includes data retrieval), two of the scenarios (Porto and Paris) revealed high processing times. Porto's processing time is due to the fact that this search retrieves a large amount of events through screen-scraping, using jsoup and Selenium WebDriver (automates the Chrome Web Browser in order to retrieve information from TicketLine), which turned out to be a complex process. As for Paris, in spite of not needing screen-scraping, since it only retrieves data from two different APIs (Last.fm and Songkick), this scenario results in a huge amount of data, with over 2000 different events, which implies parsing large XML documents. Table 6 presents statistics regarding the event retrieval procedures from each integrated data source.

As can be seen, for some scenarios, not every existing event in each platform is aggregated, whether its retrieval could be made through an API or another approach, like screen-scraping. Since a mash up of the retrieved events from the several sources is required, there is the possibility of detecting duplicated events among two or more platforms. In these cases, only a single event is aggregated and presented to the user, to avoid data redundancy. It should be noted that in the Paris scenario information from TicketLine and Yeaaaah is not aggregated, since these platforms only divulge events scheduled for Portugal. As for the artist events search scenarios, no data is retrieved from these platforms, since none of them provides artist based search. Overall, by comparing the number of events divulged by each platform and the number of different events that the developed platform found and presented, it is noticeable how advantageous it can be to centralize existent information on the Web.

Table 6. Events retrieval statistics and average processing time per data source

Scenario (functionality)	Total different events aggregated / Existing events – Processing time (seconds)			
	Last.fm	Songkick	TicketLine	Yeaaaah
S_LOC_EV-1	11/11 – 1.63	8/8 – 3.34	2/2 – 48.35	24/28 – 8.14
S_LOC_EV-2	38/48 – 5.87	15/25 – 5.69	46/48 – 243.93	67/116 – 19.05
S_LOC_EV-3	732/757 – 448.11	1341/1876 – 481.57	N/A – N/A	N/A – N/A

Artist Information Consulting. From the information entities aggregated on this functionality, only the events associated to an artist involve data aggregation from several sources, and only on these entities has the existence of duplicated entities been evaluated. As for Top Album and Top Track, in spite of their information being aggregated from a single platform, sometimes an album has multiple editions or a track is issued on several albums, which could be considered duplicated entities. It was chosen not to treat these cases as duplicates, since the source platform itself handles these as different entities. Table 7 illustrates statistics on the information aggregation, with the defined performance parameter results being presented in Table 8. The Similar Event processing time was not measured because it is retrieved along the Artist entity data itself on a single external API invocation.

Table 7. Artist data aggregation statistics

Scenario (functionality)	Existing entities	Entities not retrieved	Different entities retrieved	Processing time (seconds)
V_ARTINFO-1 (Artist)	1	0	1	58.656
V_ARTINFO-2 (Artist)	1	0	1	128.609
V_ARTINFO-1 (Top Album)	9	0	9	3.665
V_ARTINFO-2 (Top Album)	50	0	50	23.987
V_ARTINFO-1 (Top Track)	50	0	50	3.061
V_ARTINFO-2 (Top Track)	50	0	50	2.963
V_ARTINFO-1 (Artist Bio)	1	0	1	1.100
V_ARTINFO-2 (Artist Bio)	1	0	1	0.957
V_ARTINFO-1 (Event)	151	136	13	21.303
V_ARTINFO-2 (Event)	768	618	82	58.731
V_ARTINFO-1 (Similar Artist)	5	0	5	N/A
V_ARTINFO-2 (Similar Artist)	5	0	5	N/A
V_ARTINFO-1 (Merch. Product)	0	0	0	48.363
V_ARTINFO-2 (Merch. Product)	53	44	9	98.854

Table 8. Artist album data aggregation evaluated parameter results

Scenario (functionality)	Relia-bility	Rele-vance	Scenario (functionality)	Relia-bility	Rele-vance
V_ARTINFO-1 (Artist)	100%	100%	V_ARTINFO-2 (Artist)	100%	100%
V_ARTINFO-1 (Top Album)	100%	100%	V_ARTINFO-2 (Top Album)	100%	100%
V_ARTINFO-1 (Top Track)	100%	100%	V_ARTINFO-2 (Top Track)	100%	100%
V_ARTINFO-1 (Artist Bio)	100%	100%	V_ARTINFO-2 (Artist Bio)	100%	100%
V_ARTINFO-1 (Event)	86.67%	9.9%	V_ARTINFO-2 (Event)	54.7%	19.5%
V_ARTINFO-1 (Similar Artist)	100%	100%	V_ARTINFO-2 (Similar Artist)	100%	100%
V_ARTINFO-1 (Merch. Prod.)	N/A	N/A	V_ARTINFO-2 (Merch. Prod.)	100%	17%

The only entities registering failures in retrieving information were Event and Merchandise Product. For events, the failures can be justified by the need of a data mash up, from which duplicate events might be identified. Apart from that, this procedure requires aggregating not only scheduled events, but also past ones, limiting these to the ten most recent. With this limit, several events might be discarded, which reflects in a lower aggregated data relevance. As for merchandising products, since its information retrieval has revealed to be a complex process, and to avoid overloading the user with too many merchandising options, it was chosen to aggregate only information on six products. This can lead to discarding existing product information (decrease of aggregated data relevance), but it should be noted that in spite of discarding these products, an external link is presented, where the user can see all existing products. All the remaining entities were successfully retrieved, which results in 100% levels of reliability and relevance. These levels are explained by the fact that the majority of these entities are aggregated using APIs, the only exception being the Artist Bio, where it is required to screen-scrape Wikipedia to retrieve the information.

Generally speaking, a good level of contextual data retrieval among the several implemented functionalities has been attained, also being verified reasonable rates of retrieved information integrity, reliability and relevance. Although, in some scenarios, specific information retrieval procedures turned out to be quite extensive which consequently lead to high controllers' response times; however, these response times were optimized with caching mechanisms that lead to significant improvements.

Usability Heuristic Evaluation. To show the results from following this kind of approach, a screenshot is presented in Fig. 3, depicting one of the most complex views in terms of human-computer interaction, as it presents a large amount of data. To avoid overloading the view with too much information, some of the presented data is structured in tabs, allowing the user to focus on one tab at a time. This also illustrates that the presented information is distinguishable, following an appropriate color combination based on good interface design guidelines [10].

25 users took part on the survey, averaging 25 years of age, with studies on diversified areas, and all with a good experience with the Web. Regarding the SUS, an average score of 89/100 was obtained, which according to some studies is an indicator that the platform has an above-average usability level, which is a very good result. Figure 4.a) presents the average conformity levels per statement. Since half of these are worded positively (odd statements) and the other half negatively (even statements), the expected oscillation from statement to statement is noticeable, revealing the tendency to agree with the positively worded statements and disagree with the others. The statements with higher deviation (6 and 8) refer to the presence of inconsistent information and to the platform being cumbersome to use. These can be explained by the fact that users were free to use the platform, with no defined test scenarios, which may have reflected on more specific or complex data searches. As for the CSUQ based survey, Fig. 4.b) presents the average conformity levels.

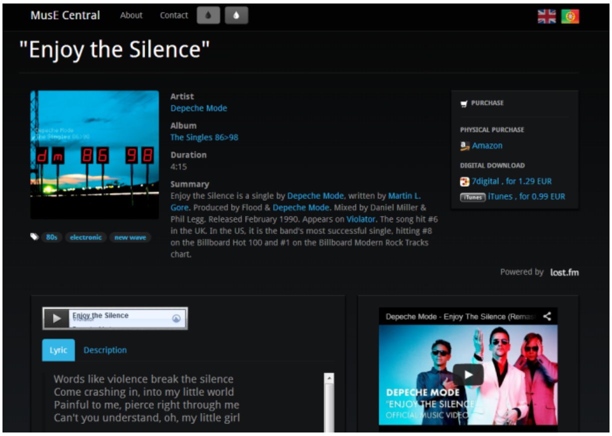


Fig. 3. Track information view

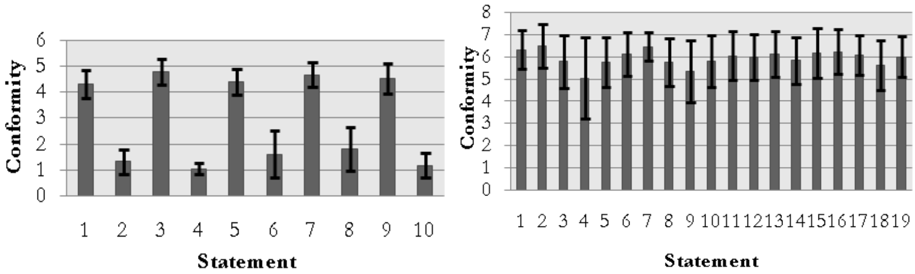


Fig. 4. a) SUS conformity levels b) CSUQ conformity levels

6 Conclusions

Along the project development, it only became more noticeable that concerts information is currently decentralized over the Web in several platforms with their own knowledge base for scheduled events. With the developed platform, it has been proven how advantageous it is to take benefit from existing data on the Web, centralizing it and enhancing the provided service, in this case, a concerts searching service complemented with event, artist and venue information. The developed platform has gathered and divulged, on the event-searching test scenarios, a higher number of events than its data source platforms. Apart from providing a more efficient search service, due to a knowledge base enriched with concert information from several providers, a much more complete service was also attained, by using additional contextual data, made available on a single service. In addition, it has been proven that the platform provides these services through a simple, intuitive, clean and pleasant interface. In some scenarios, the data aggregation process has revealed to be time-consuming, due to either screen-scraping procedures or high number of retrieved events. Nevertheless, the amount of retrieved information has revealed to be compensatory over its retrieval time. Still, some performance optimizations could be explored: concerning events, a time span of 2-3 months could be applied on the search procedures, focusing on closer scheduled events that can be more relevant to the user.

Aiming to enhance this kind of service, some future work should be explored, such as integrating new information sources, whether from already covered contexts or new ones. User account mechanisms could also be introduced to adapt existing functionalities (eg. keep search history) or explore recommendation approaches. Due to the increasing popularity of social networks, new integration approaches could be addressed, providing a more user-focused integration. According to recent web apps and interfaces development paradigms, another question that would be convenient to analyze is the design of a mobile interface for the platform. This would allow enhancing the platform usage from mobile devices, with smaller display sizes, which could be achieved by adjusting the current interface to have a more responsive design.

Acknowledgments. This work is partially supported by iCIS project (CENTRO-07-ST24-FEDER-002003) which is co-financed by QREN, in the scope of the Mais Centro Program and FEDER.

References

1. Silva, D., Mendes, P., Abreu, P., Vinhas, V.: Integrador automático de notícias. In: Actas da XATA 2007, Quinta Conferência, XML, pp. 223–234. Aplicações e Tecnologias Associadas, Lisboa (2007)
2. Fielding, R.: Architectural Styles and the Design of Network-Based Software Architectures. Ph.D. Dissertation, University of California, Irvine (2000)
3. Schimler, J.: Web Services Description Requirements, `http://www.w3.org/TR/ws-desc-reqs/`
4. Governor, J., Hinchcliffe, D., Nickull, D.: Web 2.0 Architectures, 1st edn. O'Reilly Media, CA (2009)
5. Hartmann, B., Doorley, S., Klemmer, S.: Hacking, Mashing, Gluing: Understanding Opportunistic Design. IEEE Pervasive Computing, 46–54 (2008)
6. Shneiderman, B., Plaisant, C.: Designing The User Interface - Strategies for Effective Human-Computer Interaction, 5th edn. Addison-Wesley (2010)
7. Cardie, C.: Empirical Methods in Information Extraction. AI Magazine 18, 65–79 (1997)
8. Brooke, J.: SUS - A quick and dirty usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L. (eds.) Usability Evaluation in Industry, Taylor & Francis, UK (1996)
9. Tullis, T., Albert, W.: Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics, 2nd edn. Morgan Kaufmann Publishers Inc. (2013)
10. Galitz, W.: The Essential Guide to User Interface Design: An Introduction to GUI Design Principles and Techniques, 3rd edn. John Wiley & Sons, Indianapolis (2007)

An Approach for Graphical User Interface External Bad Smells Detection

J.C. Silva[1], J.C. Campos[2], J. Saraiva[2], and J.L. Silva[3]

[1] Dep. Tecnologias, Instituto Politécnico do Cávado e do Ave, Barcelos, Portugal
[2] Dep. Informática, Universidade do Minho and HASLab / INESC TEC, Braga, Portugal
[3] Madeira-ITI, Universidade da Madeira, Funchal, Portugal
jcsilva@ipca.pt, {jose.campos,jas}@di.uminho.pt, jose.l.silva@m-iti.org

Abstract. In the context of an effort to develop methodologies to support the evaluation of interactive system, this paper investigates an approach to detect graphical user interface external bad smells. Our approach consists in detecting user interface external bad smells through model-based reverse engineering from source code. Models are used to define which widgets are present in the interface, when can particular graphical user interface (GUI) events occur, under which conditions, which system actions are executed, and which GUI state is generated next. From these models we obtain metrics that can later be used to identify the smells.

Keywords: GUI, Reverse Engineering, Bad Smells.

1 Introduction

In the Software Engineering area, the use of reverse engineering approaches has been explored in order to derive models directly from existing interactive systems using both static and dynamics analysis [13] [14] [15]. Static analysis is performed on the source code without executing the application. Static approaches are well suited for extracting information about the internal structure of the system, and about dependencies among structural elements. Classes, methods, and variables' information can be obtained from the analysis of the source code. On the contrary, dynamic analysis extracts information from the application by executing it. Within a dynamic approach the system is executed and its external behaviour is analysed.

Reverse engineering is a process that helps understand a computer system. Similarly, user interface modelling helps designers and software engineers understand an interactive application from a user interface perspective. This includes identifying data entities and actions that are present in the user interface, as well as relationships between user interface objects. This paper makes use of user interface models to detect bad smells of interactive systems. Models are derived, through a reverse engineering approach, from source code.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2,* Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_19,

The first step of our approach extracts models of the GUI directly from its implementation. Models support analysis and can be used to detect bad smells at reasonable cost [9]. These particular models specify which GUI components are present in the interface and their relationship, when a particular GUI event may occur and the associated conditions, which system actions are executed and which GUI state is generated next. The main goal of this paper is to describe an approach to detect bad smells presence from these models.

This paper is structured into six sections. The first one presents an introduction. Section 2 describes user interface modelling, and the aspects that are usually specified for graphical user interfaces. Section 3 presents background information on bad smells. Section 4 describes bad smells related to graphical user interfaces behaviour and the proposed approach for bad smells detection. Section 5 presents relevant preliminary results. Finally the last section presents conclusions.

2 User Interface Modelling

User interface models can describe the domain over which the user interface acts, the tasks that the user interface supports, and others aspects of the graphical view presented to the user. The use of interface models gives an abstract description of the user interface, allowing us to express user interfaces at different levels of abstraction. User interface models are useful to validate the user interface throughout its design, implementation and maintenance processes.

Previous research from the authors enables extraction of dialogue models [4]. Dialogue models are one of the more useful type of models to design or analyse the behaviour of the system. These models are the closest to the implementation, thus reducing the gap to be filled by reverse engineering. Dialogue models describe the behaviour of the user interface. Unlike task models, where the main emphasis is the users, dialogue model focus on the device, defining which actions are made available to users via the user interface, and how it responds to them. These models capture all possible dialogues between users and the user interface. Dialog models express the interaction between human and computer. To this end, they stipulate all the interactions the user can engage in (at a given level of abstraction — e.g. pressing buttons, issuing commands, etc.) and the results of those interactions on the system. Using dialogue models, we focus our attention on bad smells related to GUI behaviour and on automatic processes for their detection.

3 Bad Smells Background

Bad smells detection aims to address software quality. Bad smells can be divided into two groups: external and internal. External bad smells are defined in relation to running software. In what concerns GUIs, external bad smells can be used as usability indicators [21]. Internal bad smells are defined and statically searched for in source code.

3.1 Internal and External Bad Smells

The analysis of source code can provide a means to guide the evaluation of the application. Fowler introduced 22 different kinds of bad smells in code, i.e internal bad smells [11], which are useful to enhance software's internal quality through refactoring processes. Different smell types were specified by Fowler, such as:

– *Duplicated Code*: means that the same code structure appears in more than one place;
– *Feature Envy*: means that a method is in the wrong place, since it is more tightly coupled to a class other than to the one where it is currently located;
– *God Class*: refers to a class that tends to perform too much work;
– *Large Class*: refers to classes that have too many instance variables or methods;
– *Long Method*: a method that is too long, so it is difficult to understand, change, or extend.
– Long Parameter List: a parameters list that is too long and thus difficult to understand.

As an example, the detection of the *Feature Envy* bad smell can be achieved by measuring the connections that a method has to methods belonging to foreign classes. From a user interface perspective, a method related to a particular widget should belong to the form's class where the widget is defined.

As described above, internal bad smells detection may indicate a code or design problem and provide information to improve software development, being useful to enhance the software's internal quality through refactoring processes.

External bad smells are not obtainable from source code analysis. External bad smells can be defined in relation to the user interface as perceived when running software. Analysis of these smells must see the system as a black box and can be achieved, for example, through users feedback or the analysis of behavioural models. Obtaining user's feedback is a costly process. Hence, our approach is to to extract external bad smells through the analysis of behavioural models generated through reverse engineering from source code. With our approach, we aim to extract external bad smells obtained from dialogue models.

3.2 Related Work

In the sequence of Fowler's study, Mantyla *et al.* [12] presented a taxonomy to categorize similar bad smells. The taxonomy makes the smells more understandable and recognizes the relationships between smells. Mantyla *et al.* created five groups of smells, namely, the bloaters, the object-oriented abusers, the change preventers, the dispensables and the couplers. Both Fowler and Mantyla use metrics for bad smells detection. Stamelos et al. [6] used also metrics within the *Logiscope2* tool in order to study the quality of open source code. Ten different metrics were used. The results enable evaluation of each function against four basic criteria: testability, simplicity, readability and self-descriptiveness. While the GUI layer was not specifically targeted in the analysis, the results indicated a negative correlation between component size and user satisfaction with the software. Yoon and Yoon [7] developed quantitative metrics to support decision making during the GUI design process. Their goal was to quantify quality attributes of interaction design. Three internal metrics were proposed and defined as numerical values: complexity, inefficiency and incongruity.

While the above approach focus on calculating metrics over the code, Thimbleby and Gow [8] calculate them over a model capturing the behaviour of the application. Using graph theory they analyse metrics related to the users ability to use the interface (e.g., strong connectedness ensures no part of the interface ever becomes unreachable), the cost of erroneous actions (e.g., calculating the cost of undoing an action), or the knowledge needed to use the system. While Thimbleby and Gow manually develop their models from inspections of the running software/devices, an analogous approach can be carried out analysing the models generated by our approach. Indeed, by calculating metrics over the behavioural models produced, relevant knowledge may be acquired about the dialogue induced by the interface, and, as a consequence, about possible external bad smells related to GUI behaviour.

4 External Bad Smells Detection Approach

Considering different types of external bad smells, we aim to detect them and discuss some of the relevant problems that we have to face for their detection in interactive systems. To achieve that purpose adequate metrics must be specified and calculated. In what concerns graphical user interfaces, external bad smells' presence can be used as usability indicator. By calculating metrics over dialogue models, relevant knowledge may be acquired about the characteristics of interface that may indicate a code or design problem. The approach described in this paper makes use of a fully functional reverse engineering prototype tool developed by the authors [4]. The tool makes use of static analysis as in [5] and is able to derive user interface models of interactive applications from source code.

The interactive source code extraction process starts by defining/reusing a front-end for the programming language of the interactive applications source code. Modern parser generators automatically produce a parser and the construction of the Abstract Syntax Tree (AST), given the context-free grammar defining the programming language of the source code. Using this front-end, an AST is obtained from the source code of the system for which the user interface related code is to be analysed. Then, the process needs to identify all fragments in the AST that are members of the GUI layer. To achieve this a set of abstractions is used. In order to extract user interface relevant data from the AST, a slicing function was proposed which isolates the GUI sub-program from the entire program. Behavioural models may then be generated which capture graphical user interface behaviour by detecting components in the user interface through source code analysis. These components include user interface objects, events, actions and respective control flow.

Different metrics have already been applied to the generated models. These metrics are used to detect external bad smells related to the interaction between users and the system. Pagerank and betweenness are two metrics used for external bad smells detection.

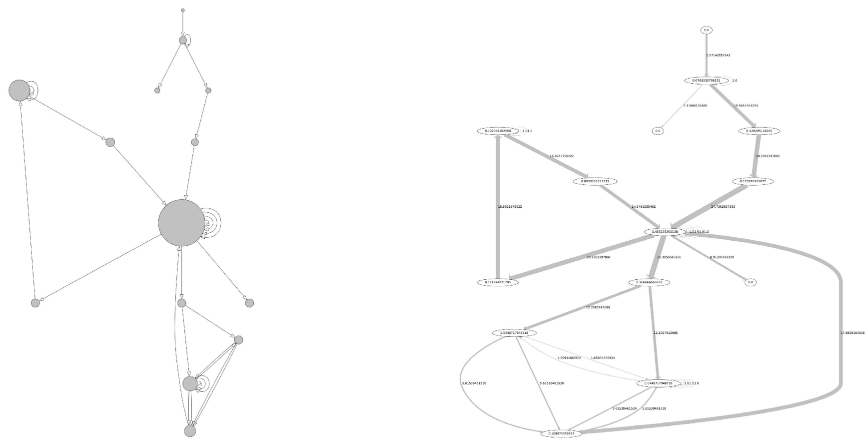
Pagerank is a link analysis algorithm, that assigns a numerical weighting to each node [19]. Pagerank is a distribution used to represent the probability that users randomly executing events will arrive at any particular state [19]. A probability is expressed as a numeric value between 0 and 1. The main objective is to measure the relative importance of the states. Larger nodes specifies window internal states with higher importance within the overall application behaviour. This metric is used to detect external bad smells, for example, to find erroneously distributed complexity along the application behaviour.

Betweenness is a centrality measure of a vertex or an edge within a graph [20]. Vertices that occur on many shortest paths between other vertices have higher betweenness than those that do not. Similar to vertices betweennes centrality, edge betweenness centrality is related to shortest path between two vertices. Edges that occur on many shortest paths between vertices have higher edge betweenness. Betweenness values are expressed numerically for each vertices and edges. Highest betweenness edges values are represented by larger edges. Some states and edges have the highest betweenness, meaning they act as a hub from where different parts of the interface can be reached, representing a central axis in the interaction between users and the system. Like pagerank, this metric can be used to detect external bad smells. For example, to find misplaced central axes in the interaction between users and the system.

5 Relevant Preliminary Results

The application of our approach was applied to a real interactive system written by a third party: a Healthcare management system (HMS). This application is a proof of concept for the approach. The HMS system is implemented in Java/Swing and supports patients, doctors and bills' management. The implementation contains 66 classes, 29 windows forms (message boxes included) and 3588 lines of code. The approach enabled the extraction of different behavioural models. Metrics such as pagerank and betweenness were also automatically calculated to support the activities involved in GUI model-based reasoning. These GUI behavioural metrics were used as a way to detect bad smells.

As an example, Figure 1 presents the, automatically generated, pagerank (Figure 1(a)) and betweeness (Figure 1(b)) results for the HMS user interface. These results enable us to visually detect external bad smell related to the overall HMS user interface complexity. For both models, each node specifies a particular window state and each edge specifies a particular transition between window states. The left model (Figure 1(a)) enables us to visually detect erroneous distributed complexity along the application behaviour. The right model (Figure 1(b)) enables us to visually find misplaced central axes in the interaction between users and the system. This case study demonstrates that the approach enables the generation of models of real interactive applications written by third parties, which can then be used for analysis.



(a) Pagerank result (b) Betweeness result

Fig. 1. HMS behavioral results

6 Conclusions and Future Work

Tools are currently available to developers that allow for the fast development of user interfaces with graphical components. However, the design of interactive systems does not seem to be much improved by the use of such tools. As described in this paper Fowler popularized the concept of program smells in the context of object-oriented programming. The presence of bad smells in software code can make programs harder to understand, maintain, and evolve. The detection of bad smells allows programmers to improve their programs by eliminating them.

In this paper we have discussed an approach for the detection of graphical user interface external bad smells through reverse engineering from source code. In the future, we intend to extend the approach presented in this paper enabling it to detect automatically a more significant amount of smells using our tool. Since the smells are usually associated with refactorings that can eliminate them, we plan also to improve applications' GUIs through refactoring and redesign. These are promising research directions that we are already exploring and whose results we plan to bring out in the near future.

References

1. Myers, B.A.: Separating Application Code from Toolkits: Eliminating the Spaghetti of Call-backs, School of Computer Science. Carnegie Mellon University (1999)
2. Memon, A.M.: A Comprehensive Framework for Testing Graphical User Interfaces. PhD thesis, Department of Computer Science, University of PittsBurgh (July 2001)
3. Berard, B.: Systems and Software Verification. Springer (2001)
4. Silva, J.C., Campos, J.C., Saraiva, J.: A generic library for GUI reasoning and testing. In: 24th Annual ACM Symposium on Applied Computing, USA (March 2009)

5. Moore, M.M.: Rule-based detection for reverse engineering user interfaces. In: Proceedings of the Third Working Conference on Reverse Engineering, Monterey, CA, pp. 42–48 (November 1996)
6. Stamelos, I., Angelis, L., Oikonomou, A., Bleris, G.L.: Code quality analysis in open source software development. Information Systems Journal 12, 43–60 (2002)
7. Yoon, Y.S., Yoon, W.C.: Development of quantitative metrics to support ui designer decision-making in the design process. In: Jacko, J.A. (ed.) Human-Computer Interaction, Part I, HCII 2007. LNCS, vol. 4550, pp. 316–324. Springer, Heidelberg (2007)
8. Thimbleby, H., Gow, J.: Applying graph theory to interaction design, pp. 501–519 (2008)
9. Miller, S.P., Tribble, A.C., Whalen, M.W., Heimdahl, M.P.E.: Proving the shalls early validation of requirements through formal methods (2004)
10. Nielsen, J.: Usability Engineering. Academic Press, San Diego (1993)
11. Fowler, M.: Refactoring: Improving the Design of Existing Code. Addison-Wesley, Boston (1999)
12. Mantyla, M., Vanhanen, J., Lassenius, C.: A taxonomy and an initial empirical study of bad smells in code. In: Proceedings of the International Conference on Software Maintenance, ICSM 2003, pp. 381–384. IEEE Computer Society, Washington, DC (2003)
13. Paiva, A.C.R., Faria, J.C.P., Mendes, P.M.C.: Reverse Engineered Formal Models for GUI Testing. In: Leue, S., Merino, P. (eds.) FMICS 2007. LNCS, vol. 4916, pp. 218–233. Springer, Heidelberg (2008)
14. Chen, J., Subramaniam, S.: A GUI environment for testing guibased applications in Java. In: Proceedings of the 34th Hawaii International Conferences on System Sciences (2001)
15. Systa, T.: Dynamic reverse engineering of Java software. Technical report, University of Tampere, Finland (2001)
16. Kataoka, Y., Imai, T., Andou, H., Fukaya, T.: A quantitative evaluation of maintainability enhancement by refactoring. In: Proceedings of International Conference on Software Maintenance (ICSM 2002), Montral, Canada, pp. 576–585. IEEE Computer Society (October 2002), doi:10.1109/ICSM.2002.1167822
17. Tahvildari, L., Kontogiannis, K.: A metric-based approach to enhance design quality through meta-pattern transformations. In: Proceedings of the Seventh European Conference on Software Maintenance and Reengineering, Benevento, Italy, pp. 183–192 (March 2003), doi:10.1109/CSMR.2003.1192426.
18. Fontana, F.A., Spinelli, S.: Impact of refactoring on quality code evaluation. In: Proceeding of the 4th Workshop on Refactoring Tools, WRT 2011 (Workshop Held in Conjunction with ICSE 2011), Waikiki, Honolulu, HI, USA, pp. 37–40. ACM (2011), doi:10.1145/1984732.1984741.
19. Berkhin, P.: A survey on pagerank computing. Internet Mathematics 2, 73–120 (2005)
20. Shu, Y.S., et al.: Fast centrality approximation in modular networks (2009)
21. Garrido, A., Rossi, G., Distante, D.: Refactoring for Usability in Web Applications. IEEE Software 28(3), 60 (2011)

Improving High Availability and Reliability of Health Interoperability Systems

Fernando Marins[1], Luciana Cardoso[1], Filipe Portela[2], Manuel F. Santos[2], António Abelha[1], and José Machado[1,*]

[1] University of Minho, Computer Science and Technology Center, Braga, Portugal
[2] University of Minho, ALGORITMI Research Center, Guimarães, Portugal
{f.abreu.marins,lucianacardoso05}@gmail.com, {cfp,mfs}@dsi.uminho.pt, {abelha,jmac}@di.uminho.pt

Abstract. The accessibility and availability of patient clinical information are a constant need. The Agency for Interoperation, Diffusion and Archive of Medical Information (AIDA) was then developed to ensure the interoperability among healthcare information systems successfully. AIDA has demonstrated over time the need for greater control over its agents and their activities as the need for monitoring and preventing its machines and agents.

This paper presents monitoring and prevention systems that were developed for machines and agents, which allow not only prevent faults, but also watch and evaluate the behaviour of these components through monitoring dashboards. The Biomedical Multiagent Platform for Interoperability (BMaPI) implemented in *Centro Hospitalar do Porto* (CHP) revealed provide the necessary data and functionalities capable to manage and to monitor agents' activities. It was found that the prevention systems identified critical situations successfully, contributing to an increase in the integrity and availability of AIDA implemented in CHP.

Keywords: Interoperability, Health Information Systems, Monitoring System, Fault Forecasting.

1 Introduction

The health information systems (HIS) have been a topic increasingly attractive in the field of scientific research, since provide to health professionals a better contact with the equipment, thus reducing the number and incidence of medical errors. Also contribute to a reduction of costs and may provide a means to improve the management hospital. Due to its specificity, the HIS constitutes a complex environment of heterogeneous systems, distributed and ubiquitous, communicating in different ways, integrating specific medical equipment that are developed by different entities, with different goals. Thus, increasingly emerges the need of HIS stay in constant communication [1, 2].

* Corresponding author.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2,*
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_20,

The possibility and the need to communicate is one of the main characteristics of human beings. Also HIS needs to communicate and cooperate in order to improve their overall performance, their usefulness, the quality of diagnoses, but mainly to improve the quality of patient care. Cooperation and exchange of data and information is in fact one of the most important characteristics, and essence for the optimization of existing resources and improving the decision making process through consolidation, verification and dissemination of information [3].

It is presented then the concept of interoperability, which defines the capacity for independent systems to exchange meaningful information and perform actions belonging to other systems in order to operate together towards a mutual goal [3, 4].

In order to aggregate and consolidate all relevant information, a solid and efficient process of integration and interoperability has to be developed. This process must take into account the scalability, flexibility, portability and security (confidentiality, integrity and availability) when applied in this environment. The Agency for Interoperation, Diffusion and Archive of Medical Information (AIDA) arises to be part of this process of integration of different sources of information through the use of different protocols and methodologies. Furthermore this platform also provides tools to implement and facilitate the communication with humans through web based services [1]. However, this platform contains some limitations of which highlights the difficulty in controlling the AIDA's agents and the need to monitor and prevent failures in their components. To address these two limitations and improve the availability and functionality of AIDA, several systems have been developed for monitoring and prevention and they are presented in Section 4. Section 2 presents and evaluates the AIDA platform. Section 3 presents MEWS, a fault forecasting model used in Intensive Care Units, the prevention systems developed are inspired in this model. Finally, Section 5 presents the results about the implementation of these systems and Section 6 presents conclusions and future work.

2 AIDA

The techniques based on Artificial Intelligence (AI) has been demonstrated a high potential when introduced into solutions applied to hospital environments. Many of these solutions are already in production in several portuguese healthcare institutions, centred in systems integration and decision support systems. Also in the interoperability area, these techniques have been highlighted, namely a solution developed by a group of artificial intelligence at University of Minho called AIDA [5] is already a central unit for interoperability in several large organizations such as *Centro Hospitalar do Porto* (CHP), *Centro Hospitalar do Tâmega e Sousa*, *Centro Hospitalar do Alto Ave* and *Unidade local de Saúde do Norte Alentejano*. This agent-based platform has demonstrated a great adaptability, modularity and effectiveness through the use of a basic Multi-Agent System (MAS) that has grown according to the particular needs of each institution.

AIDA was designed to assist medical applications in the form of a network of intelligent systems for the information processing, its major subsystems, their functional role, and for control the flow of information with a level of autonomy adjustable. It was developed to allow the diffusion and the integration of information generated in a healthcare environment. It is a complex system composed of specialized subsystems understood as intelligent and flexible entities, in this work designed as intelligent agents, responsible for tasks such as the communication among heterogeneous systems, the sending and the receiving of information (e.g. clinical reports, images, sets of data, prescriptions, etc.), management and storage of information and responding to requests in a timely and correct manner [5,6].

This platform uses many integration resources, namely technologies like Service Oriented Architectures (SOA) and MAS for the implementation of interoperability in a distributed, specific and conform to a standard manner, comprising all the service providers within the healthcare institution. The main objective of this system is, as the name implies, integrate, disseminate and archive large sets of data arising several sources (e.g. services, departments, units, computers and medical equipments, etc.). It also provides tools to implement and facilitate the communication with human beings through web-based services [3]. The AIDA platform has taken a vital role in the normal functioning of the SIH in the institutions where it is implemented. Thus it is very important ensure that it offers the best features and that the users directly related to its operation are satisfied with its performance. For this, were raised the main advantages and disadvantages of the AIDA platform, presented in Table 1.

Table 1. AIDA's advantages and disadvantages

Advantages	Disadvantages
– High power managing changes in the system; – Objets's customization ability; – High availability, accessibility and timely support; – High security of the information and the system; – Technologically modern; – Ease of maintenance and simplicity of use of the system.	– Lack of system documentation; – Need for a great education and training of the professionals; – Physical resources old, slow and limited; – Difficulty in the controlling agents and its activities; – Necessity of monitoring and preventing failures of AIDA's components.

Although it is possible to verify that the disadvantages of AIDA does not directly influence its operation. This work aims to circumvent and eliminate the last two disadvantages presented in the Table 1 by developing and implementing the BMaPI whose main goal addresses the need for a greater control over the agents and monitoring their pro-activity, as well as by developing and implementing monitoring and fault forecasting systems of AIDA's components, particularly its machines and its agents. To achieve the prediction of faults it is crucial to create fault forecasting models. For this project the models created (presented in the Section 4) are based on MEWS.

3 MEWS

MEWS stands for Modified Early Warning Score and this model has the purpose of predicting serious health problems. It is used especially in Intensive Care Units and it assumes that a serious health problem is often preceded by physiological deterioration. For the MEWS succeed, it is crucial to resort to monitoring of patient's vital signs. Based on the Table 2, scores are assigned to each parameter. The sum of the scores obtained means the patient's level of risk. In this way, physicians are able to prevent serious problems [7].

Table 2. MEWS Scores (adapted) [7]

MEWS Score	**3**	**2**	**1**	**0**	**1**	**2**	**3**
Temperature (^{o}C)		< 35.0	35.1-36.0	36.1-38.0	38.1-38.5	> 38.6	
Heart rate (min^{-1})		< 40	41-50	51-100	101-110	111-130	> 131
Systolic BP ($mmHg$)	< 70	71-80	81 - 100	101 - 199		> 200	
Respiratory rate (min^{-1})		< 8		8-14	15-20	21-29	> 30
Blood oxygen (%)	< 85	85-89	90-93	> 94			
Urine output ($ml/kg/h$)	Nil	< 0.5					
Neurological		New confusion		Alert	Reacting to voice	Reacting to pain	Unresponsive

In order to categorize the patient's state based on Table 2, the physicians follow several guidelines such as: if any of the parameters have a score equal to two, the patient must be in observation; if the sum of scores being equal to four or there being an increase of two values, the patient requires urgent medical attention; if a patient has a score higher than four, he is at risk of life [7].

The MEWS's implementation presents several advantages: it allows to analyse the patient's historical over time, through the continuous collection and storage of the patient's vital signs; it also allows to set priorities for the interventions to perform; the process of monitoring and observation of the patient is enriched because it provides information about organism's physiological trends; it assists in medical decision making, once it is based on quantitative criteria; it forecasts situations where the patient requires hospitalization in the Intensive Care Unit [7,8].

MEWS is used in hospitals in some countries such as Australia, the UK and Spain. In the case of Portugal, it is implemented in *Centro Hospitalar do Barlavento Algarvio* and it is used in researches developed in CHP [6,8,9]. Given the MEWS's advantages and applicability, this model proves to be a reliable option to adapt to AIDA's components.

4 Intelligent Systems for Monitoring and Preventing

Today it is vast the number of platforms developed to create agents and to establish their communication. However the monitoring of agents' activity has been a subject rarely addressed yet. With respect to the fault forecasting systems, some similar works have been done such as systems to predict failures in healthcare databases [9]. The monitoring of AIDA's components (such as its agents and machines) it is fundamental to control their activities, furthermore this process enables the administrators to balance the system's resources properly. Through a successful process of monitoring, it is possible to implement fault forecasting systems, in this case based on MEWS model presented in the Section 3, in order to prevent critical situations that might compromise the proper functioning of AIDA and consequently the quality of the healthcare delivery.

4.1 Monitoring Systems

In order to overcome one of the drawbacks of AIDA platform, was designed the Biomedical Multi-agent Platform for Interoperability (BMaPI), a system that allows the control and manage of a community of agents, ensuring their survival in a heterogeneous environment. More specifically, a platform that integrates all agents in the environment, regardless of the machine where they perform their activity. Beyond it ensure that all AIDA's agents communicate, it allows the creation of an interoperable environment. To facilitate the control of the agents, it was also developed a graphical user interface integrated into BMaPI for the user to manage in a simple and attractive way, the agents that constitute the system and their activities. In general, and taking into account the needs described above, the BMaPI was designed to be able to: ensure a greater control over the agents that constitute the AIDA; facilitate the user to the creation and registration of new agents locally or remotely; allow the user to enable and disable services through the launch or stop a specific agent; facilitate scheduling and rescheduling the activity of agents; monitoring the activity of agents.

Several agent platforms have been developed with particular attention to the thematic interoperability and compatibility. Accordingly, the reference model FIPA has emerged as a standard for agents oriented programming, which among other specifications have one to standardize the communication between the agents in order to ensure their social skills. This specification, called FIPA-ACL was used for the development of this platform, to ensure the uniformity of the messages exchanged among the agents. In general terms, BMaPI is based on a multi-thread application respecting the specifications of a client/server architecture making use of TCP/IP sockets to exchange the messages between the stakeholders, the agents. The TCP protocol used confers to BMaPI a great reliability because it guarantees that the information is transported and delivered without any data manipulation. The client/server multi-thread architecture and the property that the server has to communicate with multiple clients at the same time, gives to BMaPI large scalability because this network may gradually grow. In BMAPI, agents of various types can be created, thus being considered as a system with great adaptability. It is also a system with great portability once its server can run on any machine, adapting automatically to their properties. Its distributed architecture also benefits the BMaPI reducing your processing time and consequently its time of response. Moreover, these technologies confer on BMaPI a high level of interoperability.

Once the BMaPI ensures AIDA's agents monitoring, it emerges the necessity to control the AIDA's machines activity. After all, it is on these machines that the agents execute their tasks. Besides that, the monitoring of these machines enables the system administrators to evaluate their performance and to balance the resources properly, maximizing AIDA's efficiency.

The selection of performance indicators demands a high knowledge about the system wherein it is intended to implement a monitoring system. The selection should be based on the workload that each component is subjected and parameters that are important to achieve the objectives delineated [10].

In this way, a monitoring system for AIDA's machines was developed. This system collects periodically information from several machines about the CPU and memory usage and disk free space. In order to collect this information, it uses the WMI (Windows Management Instrumentation) technology, which contains a powerful set of tools, based on standards, that enables an easy exchange of data. It uses the credentials that are provided to access remotely objects that contains the desired information. This access is made through a simple query-based language: Windows Query Language (WQL) [11].

It was developed monitoring dashboards through a business intelligence tool named Pentaho Community [12]. These dashboards, which have a friendly graphical interface, allow system administrators to control the machines activities and aids them in making decisions relative to AIDA's resources management.

Once the monitoring process is achieved, it emerges the possibility to implement preventing systems that increases the high availability of AIDA.

4.2 Preventing Systems

It was developed two fault forecasting systems: one of the AIDA's machines and another of the AIDA's agents. For the machines, once selected the performance indicators and the monitoring process is in progress, the next step is to adapt MEWS to the AIDA's machines context. The Table 3 presents the score table for the machines. Initially there was an attempt to create a score table based on percentiles, more concretely inspired on the 95-percentile method used for billing in Internet Service Providers and websites [13]. However, the quantifying of the performance indicators by percentiles did not succeed because a reasonable number of warnings false positives were sent per day. The limits for a good performance of a machine varies a lot according to the objectives that the system administrators delineate. In this way, the solution was to create the Table 3 based on default fixed limits that can be modified anytime through an administrative page, for each parameter either generally or specifically for one machine. The limits established as default were discussed among the system administrators, which have a high experience and knowledge about AIDA.

Table 3. Score table for the machines

Score	**0**	**1**	**2**	**3**
Free processor (%)	≥ 50	$]50, 25]$	$]25, 10]$	< 10
Free memory (%)	≥ 15	$]15, 10]$	$]10, 5]$	< 5
Free disk's space (%)	≥ 15	$]15, 10]$	$]10, 5]$	< 5

As in MEWS, if the sum of the scores is equal to four or there is an increase of two values, it is considered that the situation is grave and a warning on the monitoring dashboard will appear. If the sum of all parameters' score is more than four, critical situations are detected and a warning (e-mail) is sent to the administrator, in order to take preventive actions. When one of the parameters in Table 3 reaches its maximum value, it is considered a critical situation, which triggers a warning (e-mail). This measure is justified due to the high dependence that machines have of all three parameters of Table 3 to complete successfully its tasks. In other words, if any of these parameters is overly consumed, the respective machine will have a poor performance.

In the case of AIDA's agents, the prevention system is based on the agent's activity frequency. This performance indicator is the interval of time that an agent takes to refresh its log file with its newest activity, in other words it is the frequency that the agent is executed. The Table 4 presents the score table created to forecast agents' faults also based on MEWS. The negative side was removed (as in the model related to the machines) and once it is used one variable, it was added the score four in order to the system becomes more accurate. As Table 4 indicates, the quantifying of the performance indicator is based on percentiles. Inspired in 95-percentile method [13], it was possible create this model. The usage of percentiles demands the necessity to have a large set of information about the indicator. So, during a reasonable period of time, data about agents' activity frequency was collected. In this way it is possible to evaluate the normal behaviour of agents and start the forecasting process.

Table 4. Score table for the agents based on its activity frequency

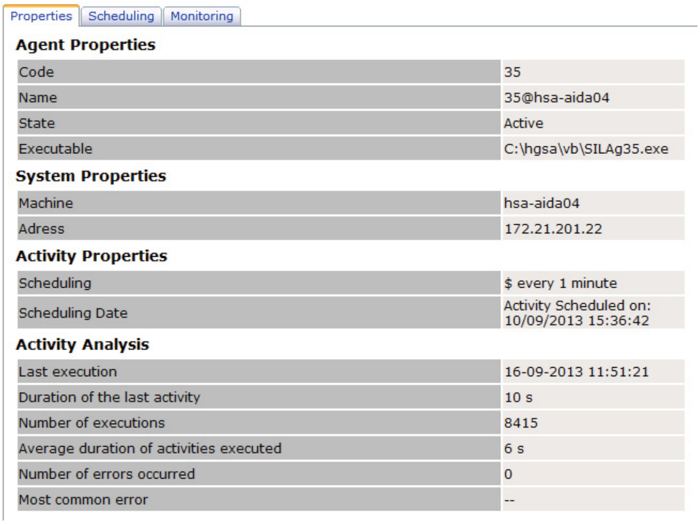
Score	**0**	**1**	**2**	**3**	**4**
Activity frequency (min)	$\leq p85$	$]p85, p90]$	$]p90, p95]$	$]p95, p97,5]$	$> p97,5$

If the score is four, it is considered an abnormal situation because the agent does not log its activity for a long time, which is unusual. An e-mail is sent to a system administrator, in order to take speedy action to restore the normal working flow of AIDA. In this way, the administrator prevents serious damages that might occur. This system constantly calculates new percentiles for each agent, improving the system's efficacy to detect abnormal situations.

5 Results

The BMaPI was implemented in a machine of the AIDA, which is implemented in CHP. The results presented are related to agent 35, responsible for validating laboratory data (e.g. clinical analysis) during the period 10 to 16 of September 2013.

After some time of execution, in this case about 6 days, the BMaPI user can already draw some conclusions about the agent through its properties page shown in figure 1. Beyond its individual properties and the way it performs its activity (every 1 minute), this page provides a general analysis of the agent activity. It presents the date and the time of the last run of the agent, and its duration in seconds, it shows the user that the activity started at 11 hours and 51 minutes of September 16 had a duration of 10 seconds. The Figure 1 also



Properties | Scheduling | Monitoring

Agent Properties

Code	35
Name	35@hsa-aida04
State	Active
Executable	C:\hgsa\vb\SILAg35.exe

System Properties

Machine	hsa-aida04
Adress	172.21.201.22

Activity Properties

Scheduling	$ every 1 minute
Scheduling Date	Activity Scheduled on: 10/09/2013 15:36:42

Activity Analysis

Last execution	16-09-2013 11:51:21
Duration of the last activity	10 s
Number of executions	8415
Average duration of activities executed	6 s
Number of errors occurred	0
Most common error	--

Fig. 1. BMaPI (Properties page from the agent 35)

shows the number of total runs that the agent had since its creation to the time of analysis and the average duration of these activities (6 seconds), this average duration was calculated using the 50th percentile in order to present a measure of central tendency, eliminated so some outlines. With these two data, besides the user can verify its values, may also calculate the number of times the agent has performed its activity for a day, i.e. if the agent is performed 8415 times its activity for nearly 6 days, per day performed approximately 1402 times and by hour 58 times.

To finish, this page makes possible even check the number of errors that occurred along the 8415 activities performed by the agent 35 and also what is the most common error.

The BMaPI also allows the user to control the agent's activity in real time, i.e. the monitoring of activities in a dynamic manner. The monitoring page then has two types of graphs, one graph for the duration of the activities and another for the number of activities performed. By Selecting the type of graph and the day for which the user wants to check the agent's activity, comes a set of three groups. The first refers to a daily analysis, followed by a weekly analysis and finally an analysis monthly.

For preventing systems, it was analysed the warning e-mails sent by both systems presented in the Section 4.2.

The preventing system for machines, based on Table 3, detected 6 critical situations during the period between 30th July to 25th September 2013. These 6 occurrences happened in three machines of AIDA (2 occurrences per machine). For example, the e-mail received on 30th July reported that a machine was consuming 97% of CPU and 91% of memory. The administrators were properly informed and preventive actions were taken, avoiding serious damages in AIDA's functioning. Similar situations have happened to the other five occurrences.

In the case of agents' forecasting system based on the Table 4, three maintenance situations were realized in different days during the tests phase (June and July 2013). The process of maintenance, naturally, caused a stop in several agents' activities. In all of these three situations, the agents' fault forecasting system detected that those agents stopped. This fact proves that the system is capable of quickly detect an agents' irregular situation preventing bigger damages. During the same period, two critical situations (score 4) were detected. In both situations the agents were stopped for more than 15 minutes and they usually register their activities each minute. These occurrences were successfully reported to administrators who treated the case.

Based on the results presented, it can be stated that the developed systems contribute to the improvement of AIDA's integrity and high availability and its functionalities and usability are increased through the BMaPI.

6 Conclusions and Future Work

After a detailed study about AIDA platform, it arises the need to eliminate some of its weaknesses. Thus were developed and implemented monitoring and prevention of failures systems, in order to improve not only the functionality and usability of the platform but also give it a greater availability.

Regarding the prediction of failures, it was possible to adapt the model MEWS to the AIDA's machines and agents' performance, to build tools with a high potential to assist system administrators and consequently improve the quality of services provided.The results show that the BMaPI platform allows the user to make a detailed analysis about the activity of each agent.

It is proposed as future work the relation between BMaPI and the machine monitoring system, in order to make a proper balancing of resources automatically. Namely the control of machines' CPU and memory so that whenever one machine reach a certain threshold the agents that are hosted on this machine can migrate to another machine. This way it is possible a managing the resources in an automatic and efficient way. Another future work related to the fault forecasting systems developed is the improvement of their alert module, adding alerts via SMS in extreme cases.

Acknowledgement. This work is financed with the support of the Portuguese Foundation for Science and Technology (FCT), with the grant SFRH/BD/70549/2010 and within project PEst-OE/EEI/UI0752/2011.

References

1. Duarte, J., Salazar, M., Quintas, C., Santos, M., Neves, J., Abelha, A., Machado, J.: Data quality evaluation of electronic health records in the hospital admission process. In: 2010 IEEE/ACIS 9th International Conference on Computer and Information Science (ICIS), pp. 201–206 (August 2010)
2. Weber-Jahnke, J., Peyton, L., Topaloglou, T.: ehealth system interoperability. Information Systems Frontiers 14(1), 1–3 (2012)
3. Miranda, M., Duarte, J., Abelha, A., Machado, J., Neves, J.: Interoperability in healthcare. In: Proceedings of the European Simulation and Modelling Conference (ESM 2010), Hasselt, Belgium (2010)
4. Rogers, R., Peres, Y., Müller, W.: Living longer independently–a healthcare interoperability perspective. E & I Elektrotechnik und Informationstechnik 127(7), 206–211 (2010)
5. Abelha, A., Machado, J., Santos, M., Allegro, S., Rua, F., Paiva, M., Neves, J.: Agency for integration, diffusion and archive of medical information. In: Proceedings of the Second IASTED International Conference-Artificial Intelligence and Applications, Benalmadena, Spain (2002)
6. Portela, C.F., Santos, M.F., Silva, Á., Machado, J., Abelha, A.: Enabling a pervasive approach for intelligent decision support in critical health care. In: Cruz-Cunha, M.M., Varajão, J., Powell, P., Martinho, R. (eds.) CENTERIS 2011, Part III. CCIS, vol. 221, pp. 233–243. Springer, Heidelberg (2011)
7. Devaney, G., Lead, W.: Guideline for the use of the modified early warning score (MEWS). Technical report, Outer North East London Community Services (2011)
8. Albino, A., Jacinto, V.: Implementação da escala de alerta precoce - EWS. Technical report, Centro Hospitalar do Barlavento Algarvio, EPE, Portimão (2010)
9. Silva, P., Quintas, C., Duarte, J., Santos, M., Neves, J., Abelha, A., Machado, J.: Hospital database workload and fault forecasting. In: 2012 IEEE EMBS Conference on Biomedical Engineering and Sciences, Malaysia (2012)
10. Boudec, J.: Performance Evaluation of Computer and Communication Systems. Computer and communication sciences. EPFL Press (2010)
11. Boshier, A.: Windows Management Instrumentation: A Simple, Powerful Tool for Scripting Windows Management. MSDN Magazine 4 (2000)
12. Tereso, M., Bernardino, J.: Open source business intelligence tools for smes. In: 2011 6th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–4. IEEE (2011)
13. Dimitropoulos, X., Hurley, P., Kind, A., Stoecklin, M.P.: On the 95-percentile billing method. In: Moon, S.B., Teixeira, R., Uhlig, S. (eds.) PAM 2009. LNCS, vol. 5448, pp. 207–216. Springer, Heidelberg (2009)

Developing a Learning Objects Recommender System Based on Competences to Education: Experience Report

Silvio César Cazella[1,2], Patrícia Alejandra Behar[3], Daisy Schneider[3], Ketia Kellen da Silva[1,3], and Rodrigo Freitas[3]

[1] Universidade do Vale do Rio dos Sinos (UNISINOS)
São Leopoldo - RS -Brazil
[2] Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA)
Porto Alegre - RS - Brazil
[3]Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre - RS - Brasil
{silvio.cazella,daisy.schneider,ketiakellen}@gmail.com,
pbehar@terra.com.br, rodrigo_frelei@yahoo.com.br

Abstract. This article describes a learning objects (LO) Recommender system based on competences to Education. The system is being developed by a multidisciplinary team, and it is in an evaluation and improvement stage. The present paper presents the experiences performed during the system application to post-graduation groups, and the improvements that were implemented from the feedback of this application. The research aims to collaborate to build didactic-pedagogical strategies with the objects' support in superior education, having the competences as a base. The results show a quality system related to LO filtering and recommendation.

Keywords: Recommender Systems, Learning Objects, Competences.

1 Introduction

This article describes an experience report of development and validation of Learning Objects Recommender System based on competences. This resource is being developed by a multidisciplinary team and has to stages until present time. The first one provided the system model building, prototyping, and its evaluation performed by two post-graduation students. The second stage consisted on the detailed adjustment and significant improvement implementation in the system.

It is comprehended that within the technologic advances, new teaching and learning conceptions had emerged, as well as different possibilities of developing content using learning objects (LOs). It is known that the LOs when incorporated to presence or distance education become important resources, for they provide easy access to the themes being studied, making possible to the subject to commit himself in an independent and autonomous way. Thus, one of the challenges faced by the professor is the selection of these digital materials, in order to accomplish his students' profiles and necessities. To collaborate in the aggregation and search to objects in the same system, learning objects repositories were created, which make

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_21, © Springer International Publishing Switzerland 2014

easier to select objects from fields, contents and types (video, simulation, pages, etc.). On the other hand, is observed that these repositories when consulted end up bringing irrelevant contents, causing information overload to the user.

As a solution proposal to this problem that were originated the Recommenders systems that, according to [2], are systems that aim to help the user on the search and selection of a content focused on his profile, working literally as information filters. Thus, the user will have as a search result only the closest and more relevant materials, as he uses and feeds the system with new pieces of information to his profile or searches performed.

The Recommender system presented in this article aims to filter the learning objects according to competences to be built which were signalized in the user's profile. It is important to mention that this article describes the developed system and does not make a discussion about metadata issues, assuming that metadata is filled at the moment of the LOs registering. Following these innovation processes, the concept of competence emerges in the integral subjects formation perspective. In this study it is comprehended that it is composed by Competences, Abilities and Attitudes, which might be built with digital competences support. The Recommender system has potential to collaborate to LOs suggestion, which may help the user to build one or more competences. Thus, it might be an important tool for both presence and distance education in different education contexts. These aspects will be discussed in the sequence.

The article is organized in five sections, including the Introduction and the Conclusions. Section two discusses the concept of competences, while section three describes the content Recommender Systems based on competences. Section four presents Recommender model conceived and the prototype developed, as well as the evaluation performed.

2 Competences: An Education View

The term "competence" has its origin in the legal scope being the "competence to judge something". Little by little this concept had been inserted in the Education field. Initially, it was utilized in relation to qualification and certification, especially in the professional education. The term started to get strength in the professional field, being used and explored in the area literature. In Education, by the end of the 90's, there was a great raise on interest and on the production of articles, books, events, and documents about the subject to different levels and modalities

However, this area still faces resistance due to the taylorist-fordist view that was associated to competences, mainly in the 70's. Though, the competences when related to a constructivist perspective have potential to contribute to the student's integral formation, for they go beyond simply content memorization. It does not mean that the contents are not necessary, but it is really important to build knowledge about them, as well as develop abilities and attitudes, making possible the competences expression. They might also depend on the context definition to which they will be applied, and thus they have a dynamic character. Le Boterf (2004) [5] assures that being competent in the 21st century is the same as other historical moments. This is because the technological, social and cultural changes modify what is necessary or relevant for that scene in face to new challenges.

It is important to mention that the perspective of this article has the term "competence" in the plural, because it is comprehended that there is not only one competence, but more than one that are required in a certain situation. Those competences are formed by a group of elements: Knowledge, Abilities and Attitudes (KAA). Le Boterf (2004) [5] proposes the competence to be conceived as the capacity of mobilizing this group of resources to deal with a complex situation. This way, "the competence […] lies [...] in the mobilization of these resources. The competence belongs to the act of knowing how to moblize". Thus, this is a metaphor used as an example to how the KAA competences relate when the subject faces a new situation. Perrenoud (2004) [8] adds that the mobilization is not only the "use" or the "application", but also adaptation, generalization or specification, orchestration, or elements coordination. To summarize, it is a set of complex mental operations applied to the situations, which is not limited to knowledge only [7].

The KKA may be related to the four pillars to the 21st centuray Education [4]: learn to know (knowledge), learn to do (ability), learn to live with the others and learn to be (attitudes). Thus, facing a scene of constant transformation, in which the contents are created and recreated with a great speed, emerges more and more the necessity of learning and building competences capacity development. The learning objects, in this context, may collaborate in the development of part of the KAA or of the competence in its totality. So, the Learning Objects Recommender System, prototyped from the defined model that is described in this article, intends to collaborate with the pedagogical practice in this different perspective.

3 Content Recommender Based on Competences

With the quantity of pieces of information available on the internet in an easy and fast way, people face a great diversity of options. Many times, an individual is not experienced enough to choose among the many alternatives of contents that are presented. In this universe, professors and students have the Learning Objects (OAs) available, which are modules or units of contents for learning with digital technology support (IEEE 2002, p. 5 apud [3]). They have as characteristics the possibility to be adapted, reused, accessible, durable, and can be used in different platform (Fabre et al 2003 apud [10]). Thus, the learning objects repositories are there to support the professor on the resources selection, because they are databases which store the objects aiming to facilitate the access and organization. However, the difficulty of selection persists due to the great amount of options.

Thus, the Recommender Systems come to help the user in the content selection process. In typical information filtering system, a person give recommendations as entrances and the system aggregates and directs the individuals who are considered to be interested in. One of the challenges of this type of system is to perform a proper combination between the users' expectations (their profiles) and the items to be recommended, in other words, define this interest's relationship. In the educational context, it is possible to glimpse a student being exposed to a great amount of LOs which might help in his formation. In this case, the Educational Recommender System (ERS) work as information filters, forwarding the object that best suits the students' learning necessities. Actually, recommender systems are extremely dependent to domain [9], being fundamental taking into consideration the educators participation in an ERS modeling.

In the present study, we assume that Recommender Systems based in Collaborative Filtering and Content Based Filtering and Hybridism variations may help in the indication of the most relevant material to the students' profile [6]. Thus, there is a direct participation of it in relation to the profile's indication and the search for materials to his studies. So, the users are very important actors in the recommender feeding with new objects and in metadata descriptions, creating a sharing community. Moreover, the model proposed in this study allows the LOs filtering according to the competences to be built by the user. On the following, the recommender system model will be presented in its two implementation stages.

4 RECoaComp Model

In this section, the system development stages from the model called RECoaComp (REComendador de objetos de aprendizagem baseado em Competências; in English, learning objects Recommender based on Competences) will be presented.

4.1 Learning Objects Recommender System – STAGE I

The RECoaComp model was conceived as shown on Fig. 1.

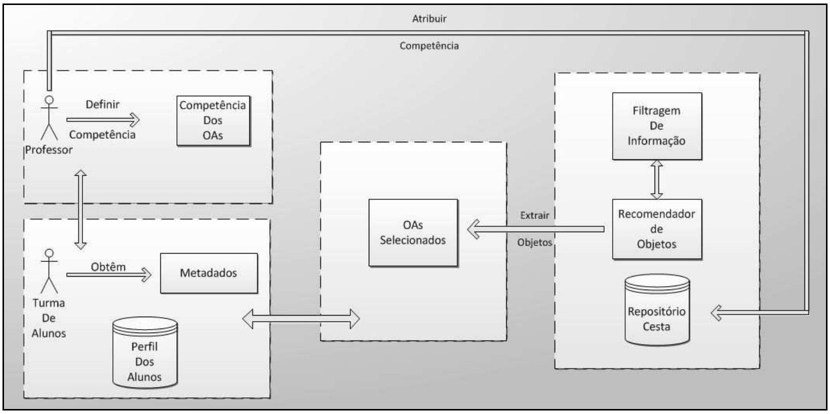


Fig. 1. Macro view of the RECoaComp model

The basic function proposed by the model consists on the following stages:

1) The professor of the subject selects which objects from the repository he will use with his group, aiming to build specific competences. He lists the competences these objects may help providing. It is important to mention that a LO can reach more than one specific competence.

2) The students answer a questionnaire about the competences that are relevant to the subject (already defined by the professor), performing a self-evaluation, making possible the LO recommendation.

3) Through the search the filtering is activated, selecting and recommending the Learning Objects to the students. These objects are selected directly from the chosen repository using the metadata records, linking the object and the profile's data, especially about competences.

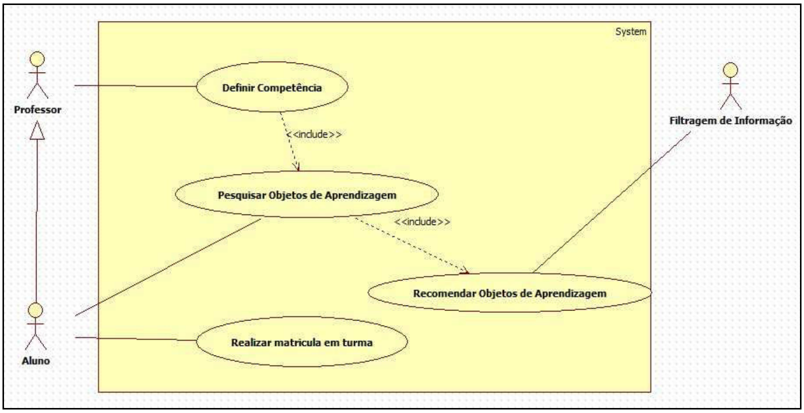


Fig. 2. Use Cases

Table 1. Use Cases Description

Use Cases
***Use Case*:** Define Competence **Actor:** Professor. **Fluxo de Eventos:** 1) Professor lists competence or competences to be built in class. 2) System searches for learning objects to reach the competences in the database. **After-Conditions:** 1) List of Learning Objects that reach the defined competences. **Use Cases Included:** 1) Search for learning objects.
***Use Case*:** Search for Learning Objects **Actors:** Student **Before-Conditions:** 1) Learning Objects Competence defined by the professor. **Events Flow:** 1) The student accesses the learning objects recommended, which probably reach his necessities towards the defined competences. 2) The student adds to his list the learning objects recommended for his study. **Use Cases Included:** 1) Recommend Learning Objects
***Use Case*:** Recommend Learning Objects **Actor:** Information Filtering **Before-Conditions:** 1) Verifies if the students' learning objects list is full. **Events Flow:** 1) Analyses the other students' learning objects. 2) Verifies the objects the students in the group have in common 3) Recommends learning objects that the student would be interested in.
***Use Case*:** Enroll in a group **Actor:** Student. **Before-conditions:** 1) Verify if the student is not already enrolled in the group. **Events Flow:** 1) Enroll the student in a group.

The Fig. 2 presents the functions that were prototyped based on the described model, while Table 1 gives details about these functions in use cases.

To describe the functions in the prototype, the UML (Unified Modeling Language) artifact was used, also known as use case.

Fig. 3 presents classes diagram of the proposed solution. The classes diagram is a very useful modeling to the system, and defines all classes (aesthetics view of the system) which participate of it.

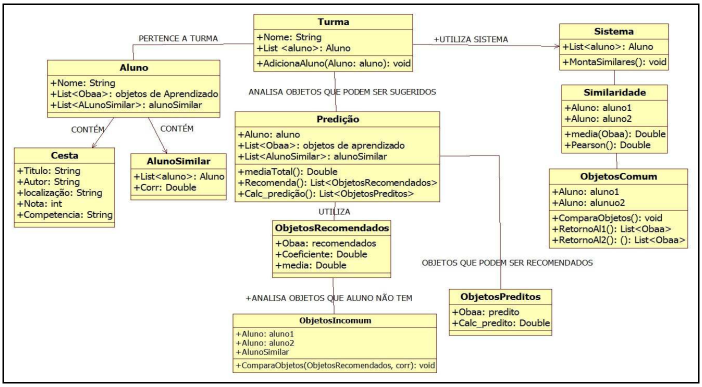


Fig. 3. Classes Diagram

In fig. 3, the group class has a list of students and each student has a list of LOs, and a list of student who have a similar way of evaluating content (Learning Objects). In ERSs it is comprehended that these students have similar "interests". In the actual implemented solution we work with Collaborative Filtering [1], although the final solution will also include the Filtering Based on Content [6]. The "system" class uses the "group" class and verifies the students who have similar "interests" related to LOs. Based on this similarity calculus it is possible to predict if a student should receive the object recommendation or not. The complete similarity and prediction calculus is in [6]. The prediction class recommends the LOs to the students.

In this stage, the following technology was used in the prototyping: Java Server Pages - JSP (view layer), and MySQL (database layer). Fig. 4 and 5 present the prototype developed based on the defined model. The first one presents the form that allows defining the competences reached by a certain Learning Object. The second one offers the student a list of recommended Learning Objects.

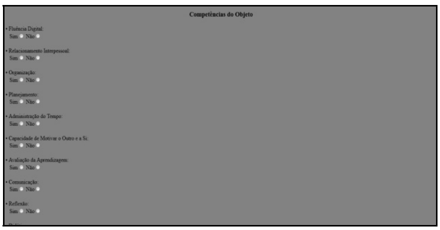


Fig. 4. Competences Recording Interface

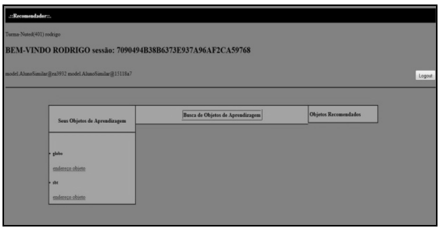


Fig. 5. LOs Recommender Interface

4.1.1 Experiment – STAGE I

To perform the recommender system prototype developed from the RECoaComp model of STAGE I verification and evaluation, it had to be applied in a group composed by students of two post-graduation programs during the second semester of 2011. The group composed by 25 students (sample for convenience) was invited to take part in the experiment that aimed to validate and evaluate the system on this first stage. A tutorial about how to use the tool was available. The methodological steps were the following:

1) Record in the system by filling in a form about profile definition related to competences. The questionnaire involves questions about "Teaching Experience on E-learning" (options "yes" or "no"), competences evaluation based on knowledge, abilities and/or attitudes developed or not to one of them. The piece of information about the competences that the student believed he had not developed made possible the information filtering procedure, because it was possible to link the student's profile to the learning object relevant to his learning.

2) Once the student's record was done, the students had to insert the objects developed by the core research in a form based on the metadata repository developed at the University. This insertion was a team work. When inserting the selected LO, the team should inform to the system its General Category (ID information), Life Category (creation description), Technical Category (information to enable the use), Educational Category (educational description), and least Law Category (use restrictions or not). After that, the students from the groups would evaluate if the recorded LO allowed (yes or no) to develop competences. To each LO the group would analyze a number of 14 competences, which are also presented on the profile recording form. Thus, it was possible to have a good LO basis for recommendation.

3) After the recording process, the groups had to review the classification performed among them. This step was fundamental to identify some little distortions, and to adapt the filling in procedure in the system.

4) The next step was to start evaluating the recommender. To each effective LO recommendation made by the system, the students would access it and evaluate the object towards effectiveness of it in relation to the student's recorded profile, thus, giving a feedback in a numerical scale (5-point Likert scale: being "Terrible" when the LO was not reaching a competence development necessity, and "Excelent" when the LO reached a competence development necessity directly).

As a result of this experiment, it was verified that the conceived model and instantiated as a prototype completely reached the requirements defined in the system modeling (the requirements presented on the use cases in section 4.1). In relation to evaluation, it was verified that the system attended nicely the students' necessities, filtering information correctly, but in a small quantity because of the small number recorded LOs in the beginning. A qualitative analysis of the System use was performed, including the students, enabling the second improvement stage.

4.2 Learning Objects Recommender System – STAGE II

In the system evaluation performed by the post-graduation students in Stage I suggestions were made to improve the RECoaComp Model. In the first stage, the project team defined the functions, guaranteeing the prototype operation. There was not a concern about the man-machine interface, which happened in the second stage of the project. The model functions were not changed, but prototype architecture was improved by new technologies. The LOs repository remained the same.

The following Technologies were used in the prototyping procedure: JavaServer Faces - JSF (in the view layer), PrimeFaces, Java Persistence API - JPA com Hibernate (persistence layer), and MySQL.

Fig. 6. Students' Competences recording interface

Fig. 7. LOs' Competences recording interface

In fig. 6, it is possible to observe the a better readable and clean interface, because it was a request made by the validation participants in Stage I. This figure presents to the student, based on his profile and competences to be developed, the Learning Objects that he has to access, aiming to develop specific competences that the students had informed he had not reached when he performed his records in the system.

Fig. 7 shows the interface that allows the student defining the competences reached by certain Learning Object. It is observed a change in the LOs' recording Interface of Stage I, because it no longer points if there is or not a criteria in the LO (Fig. 4), but also started to point the potential the object has to develop Knowledge, Abilities or Attitudes in the analyzed criteria.

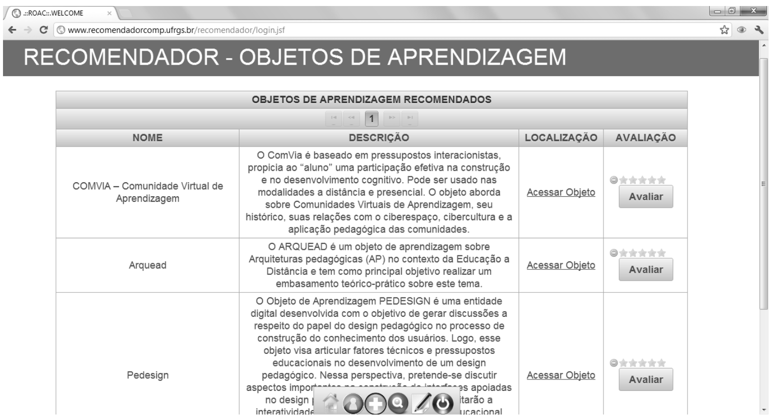


Fig. 8. LOs recommendation Interface

Fig. 8 presents to the student, based on his profile and competences to be developed, the Learning Objects that he has to access, aiming to develop specific competences that the students had informed he had not reached when he performed his records in the system. It is important to observe that the interface presents to the student the result of Recommender System based on competences, obtained by information filtering, presenting automatically the LOs which might be interesting to the target student. Using this interface, the student can access the object, and evaluate (fig. 8, word in portuguese “Avaliar” on the fourth column) it later on using a 5-point Likert scale (in fig. 8, we can see five stars above the word in portuguese “Avaliar” on the fourth column).

4.2.1 Experiment - STAGE II

To perform the recommender system prototype from the RECoaComp Stage I model verification and the evaluation, which was used in a post-graduation program group during the first semester of 2012. A total of 32 students (convenience sample) was invited was invited to participate of the experiment that aimed to evaluate the system in its development second stage. The same methodological procedure (section 4.1) of Stage I was performed. Aiming to make the understanding and use of the system easier, a tutorial video was available to the students in http://www.youtube.com/ watch?v=B6spX5NzN3k& feature=player_embedded, about the remodeled system. The students recorded 21 LOs from the selected repository. As a result of this experiment, it was verified that the model conceived and instantiated as a prototype reached the students’ expectations in Stage II, in other words, the interface was more friendly and readable. Related to evaluation, it was verified that the LOs’ recommender system reached nicely the students’ necessities, which means the information filtering was performed properly, showing that the LOs basis was sufficiently constituted. Over again, the students performed a qualitative analysis of the system, and, from this analysis, new possibilities to its improvement were glimpsed and will be future works.

5 Conclusion

This article presented a model for learning objects recommendation based on competences to all modalities of Superior Education. The use of recommender systems makes possible to forward content to the student, in this case, the Learning Objects (LOs) which better suits his necessity of competence development. This filtering is based on the link between pieces of information from the student’s profile, and the competences the Learning Objects may help developing. From the experiments performed up to the moment (Stages I and II of the research), it was possible to verify that the system is efficient and substantial in relation to the LOs recommendation quality according the students’ necessities. Other points are worth mentioning, such as it is an easy use system, conceived in its Stage II with friendly interfaces and easy understanding, as well as the possibility of using it in other domains. It is important to highlight the importance of this study to the professor’s pedagogical praxis, aiming to support new pedagogical strategies creation from the use of this tool for students’ competences development.

In future works, it is intended to include research of techniques for “Automatic machine learning” of the system for user profile update, because it remains the same until the student decides to update it to the current version. It will also be implemented a numerical scale of 0 to 10 (discrete intervals) about the competences levels; hierarchy definition (administrator, student and professor); help buttons; and other details aiming to improve results and the system usage. In relation to the technology used, it is intended to give special attention to the software security issue, probably by applying the spring-security 3.1 technology to the system.

References

1. Cazella, S.C., Silva, K.K.A., Behar, P., Schneider, D., Freitas, R.: Recomendando objetos de aprendizagem baseado em competências em EAD. Revista Novas Tecnologias na Educação 9, 1–10 (2012)
2. Cazella, S.C., Reategui, E., Machado, M., Barbosa, J.: Recomendação de Objetos de Aprendizagem Empregando Filtragem Colaborativa e Competências. In: Simpósio Brasileiro de Informática na Educação (SBIE), Florianópolis (2009)
3. Coll, C., Monereo, C.: Educação e aprendizagem no séc XXI: novas ferramentas, novos cenários, novas finalidades. In: Coll, C.E., Monereo, C. (eds.) Psicologia da Educação Virtual: Aprender e Ensinar com as Tecnologias da Informação e da Comunicação, Porto Alegre (2010)
4. Delors, J., et al.: Educação um tesouro a descobrir: Relatório para a UNESCO da Comissão Internacional sobre Educação para o século XXI, Cortez, 1ª edição (1996)
5. Le Boterf, G.: Construir as competências - individuais e coletivas, Éditions d’Organization, 3ª edição (2004)
6. Nunes, M.A., Cazella, S.C.: O que sua personalidade revela? Fidelizando clientes web através de Sistemas de Recomendação e traços de personalidade. In: Vilain, P., Roesler, V. (Org.) Tópicos em banco de dados, multimídia e Web. Tópicos em banco de dados, multimídia e Web, 1st edn., vol. 1, pp. 91–122. Sociedade Brasileira de Computação, Florianópolis (2011)
7. Perrenoud, P.: Construir as competências desde a escola, Artmed, 1ª edição (1999)
8. Perrenoud, P.: De uma metáfora a outra: transferir ou mobilizar conhecimentos? In: O enigma da competência em educação, editado por DOLZ, J. e OLLAGNIER, E (Org.), Artmed, Porto Alegre (2004)
9. Santos, O.C., Boticário, J.G.: Modeling recommendations for the educational domain. In: 1st Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL 2010), vol. 1, pp. 2793–2800. Computer Science (2010)
10. Tarouco, L.M.R., et al.: Objetos de Aprendizagem para M-Learning, In: Anais do Congresso Nacional de Tecnologia da Informação e Comunicação (SUCESU), Florianópolis (2004), http://www.cinted.ufrgs.br/CESTA/objetosdeaprendizagem_sucesu.pdf

Developing a Multi-purpose E-Learning Application for Speech and Language Therapy Students and Professionals Using a Design Science Research Approach

Miguel Baptista Nunes[1], Linhao Fang[1], Guochao Peng[1], Jingchao Zhou[1], Judy Clegg[2], Carla Ce Rohde[2], and Sarah Spencer[2]

[1] Information School, University of Sheffield,
Regent Court, 211 Portobello, Sheffield, UK
{J.M.Nunes,L.Fang,G.C.Peng,J.C.Zhou}@sheffield.ac.uk
[2] Department of Human Communication Sciences, University of Sheffield,
362 Mushroom Lane, Sheffield, UK
{J.Clegg,C.C.Rohde,Sarah.Spencer}@sheffield.ac.uk

Abstract. This paper presents Teen Talk as a multi-purpose e-learning application that was designed and developed using Design Science Research (DSR) approach. Teen Talk offers training for professionals and students who work with adolescents and young adults with significant speech, language and communication needs (SLCN). It is a multi-purpose tool, offering traditional education content to students/trainees (for example, student Speech and Language Therapists) and a professional development opportunity for professionals working in the field (for example, professionals working with adolescents with SLCN in Youth Offending teams). The resource uses video material of adolescents with SLCN to structure interactive learning tasks. Finally, the structure as well as the different aspects and components of Teen Talk are presented and discussed.

Keywords: E-learning, Speech and Language Therapy, SLCN, Design Science Research, IT artefact, Blended Learning.

1 Introduction

1.1 The Need for the Application

Speech, Language and Communication Needs (SLCN) is a term used in the UK to describe children and young people who experience significant difficulties in their speech, language and communication development. In the UK, approximately 7 to 10% of children, adolescents and young people are reported to have SLCN [1-3]. These may be children with delayed language development or children and young people who have more specific and pervasive learning disabilities involving speech, language and communication impairments, such as Autism Spectrum Disorders (ASD), Specific Language Impairment (SLI) and other developmental disorders.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_22, © Springer International Publishing Switzerland 2014

Children and young people with SLCN are a heterogeneous population and often experience a range of difficulties including speaking clearly, understanding spoken language, learning and using vocabulary, acquiring a grammatical system in order to communicate verbally with others, and using effective and appropriate social communication skills [4]. Studies reporting on the longer term trajectories and outcomes of children with SLCN in adolescence and adulthood confirm that these children compared to children without SLCN are at a much greater risk of low educational attainment, difficulties with psycho-social adaptation including mental health problems, offending behaviour, and impoverished life chances [5-11]. Identifying and understanding how to effectively meet the needs of these young people are identified as a high priority [2].

The profession of Speech and Language Therapy has traditionally supported this group of vulnerable young people [12]. However, speech and language therapy provision for adolescents with SLCN in the UK is sparse and generally speech and language therapy services in England offer more of a consultative/training approach rather than offering speech and language therapy on an individual basis [12]. As a result, there is currently much emphasis on training professionals such as teachers and teaching assistants in secondary schools and colleges, and professionals in other services such as youth offending services to understand how to meet the needs of adolescents with SLCN. This training may be delivered at the level of the secondary school or a whole service, or individually to relevant staff. Generally, such training aims to help the professional to understand the needs of these young people and then for the professional to work with the young person to facilitate their access and engagement in their learning or provision offered. For example, the charity ICAN (Invalid Children's Aid Nationwide) has developed Secondary Talk, a whole school language programme implemented in secondary schools which aims to improve pupils' behaviour and raise academic attainment by focusing on language and communication (ICAN www.ican.org.uk/secondarytalk). The Communications Trust designed and delivered a training package to professionals working in youth offending services in the UK to facilitate their knowledge and understanding of how to work with young offenders with SLCN. There are some limited evaluations of these programmes [3, 13]. These approaches though rely on face to face training which in the current economic climate is costly and has a limited reach.

E-learning is an attractive medium to deliver alternative approaches for this type of training. It is easily accessible and can have a wide reach across a range of professionals while ensuring cost-effectiveness. To date, there are few online training resources aimed at professionals working with adolescents with SLCN. Other UK based relevant online training resources around children with special educational needs including SLCN have been developed by organizations such as the Department for Education (DfE) (www.education.gov.uk/lamb), The Communications Trust (www.talkingpoint.org.uk/slcf) and the National Association for Special Educational Needs (NASEN) (www.sentrain.net). Collectively these e-learning resources target teachers, teaching assistants and other education professionals working with primary age children rather than secondary age, and range from offering free access to a financial payment.

This research discusses the design and development of such an e-learning application, entitled Teen Talk, which aims to be used both at University teaching to support a blended learning approach and to support individual practitioners wanting to engage in professional development. This paper presents and discusses the design approach adopted and the resulting application.

2 Research Methodology

2.1 A DSR Model

Teen Talk was developed using a Design Science Research (DSR) approach. This approach emerged from the Simon's [14] propositions and can be conceptualised as an iterative process of generating alternatives and testing them against predefined utility requirements and constraints. This research adopted the model proposition by Takeda et al. [15] as illustrated in Fig. 1.

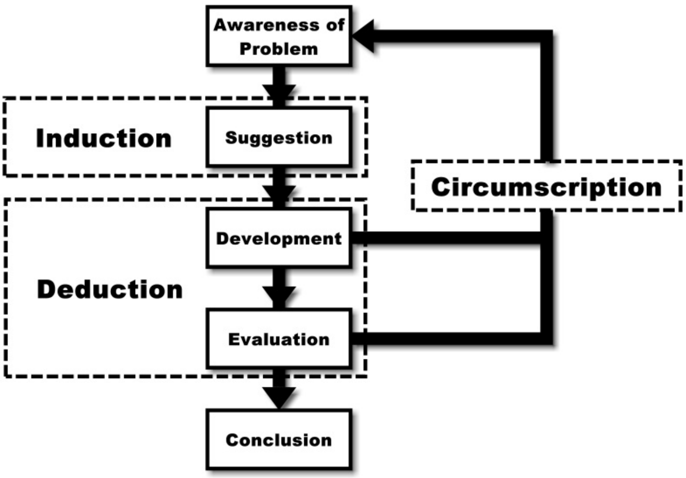


Fig. 1. The DSR Cycle model adapted from Takeda et al. [15]

The research activities proposed by Takeda et al. [15] are not significantly different from many other propositions in the literature as identified by Peffers et al. [16] in a systematic review of the most of the representative DSR studies up to that date. However, this remain the clearer and easier to implement model on which the vast majority of DSR variants are based on. Therefore, it was selected as base framework for this research.

Awareness of Problem. For DSR, a clearly identified research problem provides the focus throughout the development process [17]. Establishing clear and unambiguous research problem, questions and objectives is therefore seen as the starting step in DSR. Furthermore, the ideal problem definition requires awareness of both the state of the problem and the importance of its solution [16].

Suggestion. Takeda et al. [15] termed this stage as “suggestion” in which the key ideas needed to solve the identified problem are suggested. Peffers et al. [16] see this stage more of a setting of objectives and anticipation of a solution, similar to what is commonly defined as “requirements” in engineering fields [17, 18]. The significance of requirements specification is widely recognised in IT development. Nevertheless, the step of defining requirements (or objectives/suggestions) is neglected or only implicit in some models [19, 20]. Requirements are defined based on knowledge of the problem and functional feasibility [16]. Additionally, requirement definition should include both awareness and understanding of the environmental in which the artefact is to be used [21].

Design and Development. The design and development of the artefact are the core activities in DSR. Design determines the usability and functionality of the artefact as well as its architecture. Development activity manages the practical methods which produces the artefact [16]. These processes of design and development are explicitly separated in some elaborate models [17, 18]. In some others, the design step is merged with the requirements specification as a combined effort which delivers feasible design alternatives [15]. However, more modern views of IT artefact development, such as the one proposed by Hevner [22], integrate design and development by using prototyping and evaluation within the same process. The process of prototyping may result in cognitive conflict and changes in understanding of both initial assumptions and requirements, and consequent rethinking of suggestions.

Evaluation. Evaluation is seen as the process of comparing predefined design objectives to practically observed results from the demonstration activity [15-18]. The method to evaluate the IT artefact varies among researches. Functional comparison, satisfaction surveys, client feedback, or simulations are the instances of the investigation along with quantifiable measures such as response time or availability [16]. The result of the evaluation stage may also produce new problems which may lead to new design cycles and involve further development and refinement [15] or even the discovery of a new theory to explain observed new phenomena [17].

Conclusion or Communication. Researches aim to produce transferable knowledge which is to be shared and generalised beyond the study field. Therefore, DSR should “be presented both to technology-oriented as well as management oriented audiences” [21: 90]. This benefits both the practitioners on providing utility of the artefact and the researchers on knowledge accumulation.

2.2 DSR Research Outputs

Four types of products can be derived from DSR, namely:

- Constructs – basic vocabulary formed by the single concepts of a domain.
- Models – the description of relationships among certain constructs.

- Methods – procedures operated to perform a task basing on underlying constructs and possibly their models.
- Instantiations – the realisation of proposed artefact which operationalizes constructs, models and methods.

According to March and Smith's [23] there are potential dependencies between these different research outputs as illustrated in Fig. 2.

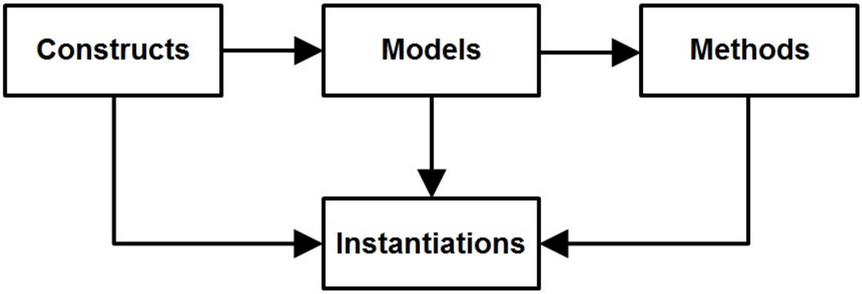


Fig. 2. Relation between the outputs of DSR

3 The Application: Teen Talk

Teen Talk. The online learning resource reported in this current study is the result of interdisciplinary collaboration between two departments of the University of Sheffield (Information School and Human Communication Sciences) and an independent post 16 years college for young people with severe and complex SLCN. The independent college offers vocational education to adolescents and young people aged between 16 and 25 years. Teen Talk offers training for professionals and student who work with adolescents and young adults with significant SLCN. It is a multi-purpose tool, offering traditional 'education' content to students/trainees (for example, student Speech and Language Therapists), and a professional development opportunity for professionals working in the field (for example, professionals working with adolescents with SLCN in Youth Offending teams). The resource uses video material of adolescents with SLCN to structure interactive learning tasks.

3.1 Using DSR to Develop Teen Talk

Throughout designing and developing the application, the general DSR framework has been adapted to suit the specific circumstances of this project.

Identification of the Needs. The need for this application emerged from a previous research project within the HCS department. The initial and very loose conceptualization of the application to be designed was then brought to the Information School. The initial identification of needs and early discussions defined the stakeholders from HCS as users of the artefact to be developed and the

Information School staff as developers. Design aspects were discussed and agreed upon together. At first this was not always an easy task and different cultures and reasoning frameworks made the dialogue at times challenging, since speech and language therapy and information systems are not necessarily close disciplines. In the end a somewhat artisanal structure chart (see Fig. 3) proved extremely useful to bridge different academic cultures. This first modelling tool was then complemented by use cases that were generated to illustrate the application's underpinning requirements. The requirements and associated use cases were reached through successive meetings where negotiation of meanings and clarity of purpose slowly emerged.

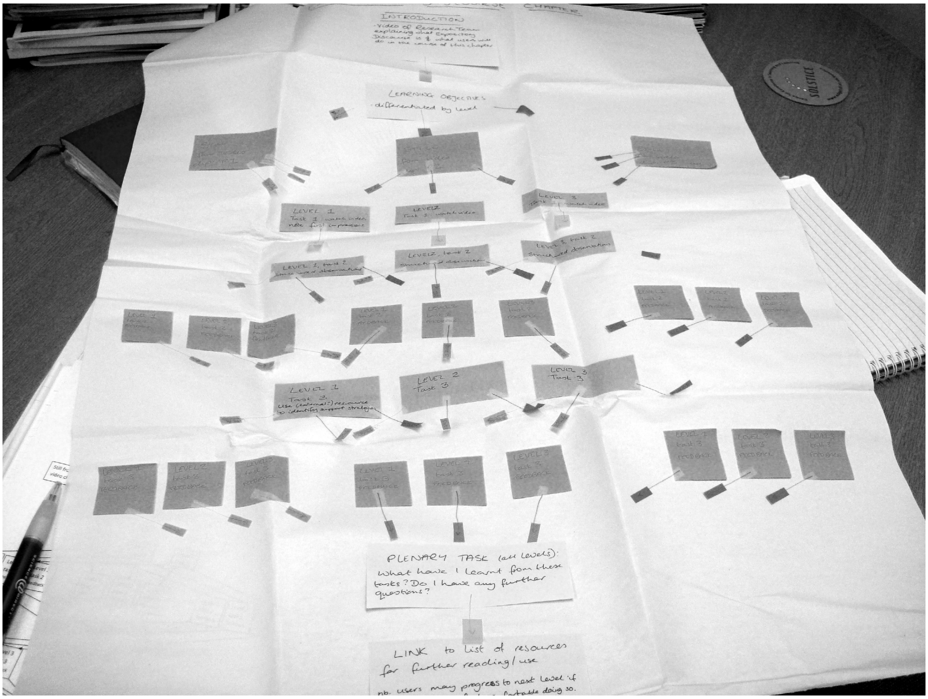


Fig. 3. Negotiation tool that enabled the interdisciplinary design of the artifact

Initial Design – Suggest Solutions. The initial design phase consisted first in the fine tuning of the use cases and then in the use of an evolutionary prototype. The use of the prototype proved to be extremely efficient. HCS stakeholders started to interpret the requirements in the light of the IT artefact that they now could visualise and manipulate. This initial design phase enabled further discussion and revision of early designs, the users' initial requirements and the establishing of a basic structure for the e-learning application being developed (see Fig. 4). The interaction and order of elemental components of the different learning activities were then described by activity diagrams.

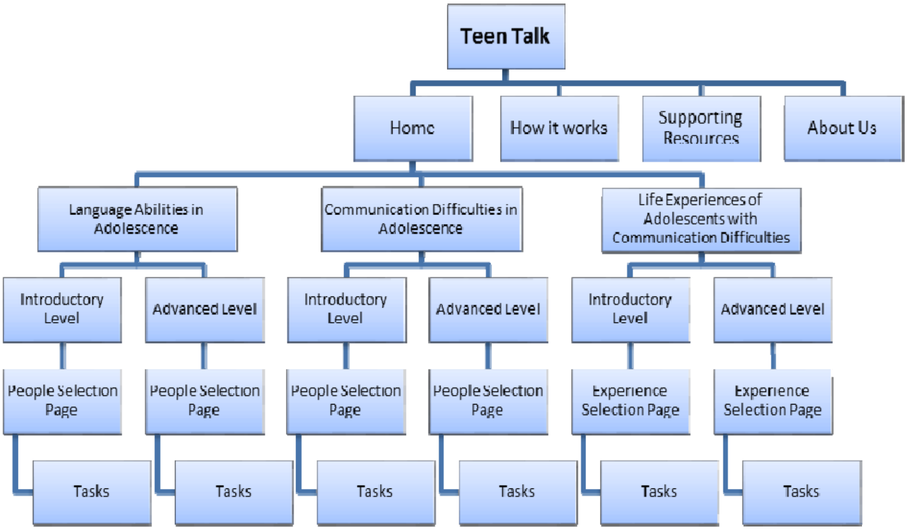


Fig. 4. Basic structure of Teen Talk

Development through Rapid Prototyping. Criticism of both functionality and human computer interface was achieved though iterative prototyping cycles. Content, multimedia objects and learning activities were developed at this stage. The evolutionary prototype strategy adopted allowed HCS stakeholders to better understand both potential and limitations of the artefact being produced. Prototype meetings were held between the Information School and the clients from HCS. A first period of great enthusiasm and amazement by HCS, soon gave way to realistic criticism and increasing demands. While each rapid prototype is delivered, further requirements and the internal test results of last prototype are discussed. The prototype itself was built using a combination of PHP, HTML and JavaScript, specific multimedia plug-ins.

3.2 Structure of the Resulting Artefact

The developed application Teen Talk is now being prepared for the formal evaluation stage. As illustrated in Fig. 4, the Teen Talk system is a multi-dimensional tool which has all its training resources divided into two levels:

- Introductory level suitable for those who want to know more about SLCN, for example student/trainee and/or professionals with little or no prior knowledge of adolescents with SLCN;
- Advanced level designed for professionals who want to develop more specialised knowledge and skills, such as specialist teachers and speech and language therapists already working in this area.

Teen Talk includes video clips of young people with SLCN talking about their experiences. Users are guided through the interactive video material to learn to conduct detailed analysis of real life speech, language and communication behaviours in adolescents. These first-hand accounts are the starting point for learning activities about person-centred approaches to support and intervention. The importance of such approaches has been highlighted by findings of the Better Communication Research Programme (BCRP) [24], established in the light of the Bercow Review of services for children and young people (0-19) with speech, language and communication needs [2]. Throughout the resource, the user is encouraged to consider how their learning can be used to inform and develop their own professional practice.

Learning materials are provided through tasks that the system would ask user to accomplish. All tasks are categorised into 3 chapters (the term preferred by the HCS stakeholders and referring to specific areas of the artefact learning materials coverage) with a different focus on adolescents with SLCN, namely:

1. Communication Difficulties in Adolescence
2. Life Experiences of Adolescents with Communication Difficulties
3. Language Abilities in Adolescence

The first chapter Communication Difficulties in Adolescence develops understanding of the nature of difficulties associated with SLCN. It uses video clips and activities based on current research to illustrate particular communication difficulties, such as word-finding difficulties and comprehension difficulties. Users are given specific tasks which will enable them to recognise particular communication difficulties, and are encouraged to consider ways in which to support people, either through general support strategies or with more detailed, specific SLCN interventions, depending on whether they are working at the introductory or the advanced level.

Chapter two Life Experiences of Adolescents with Communication Difficulties develops knowledge and understanding of the impact of speech, language and communication difficulties for adolescents and young adults. Young adults were interviewed about the impact of living with SLCN in relation to their experiences of school, college, friendships, their own SLCN and their plans and hopes for the future. Users complete tasks designed to enhance their awareness and understanding of the challenges which young people with SLCN face or may have faced previously, and are encouraged to consider ways in which to support them. The videos are also discussed in relation to research literature at the advanced level.

The third chapter Language Abilities in Adolescence aims to develop skills in analysing real-life examples of communication and language skills. This enables the user to examine the complexity of young people's language skills and language difficulties. At the advanced level, users are supported to develop transcription and analysis skills in more detail, in relation to recent research methods.

In addition to the practical task chapters, the Teen Talk system also contains a supporting resources section where further learning materials are provided for underpinning the users' learning. It includes a glossary, aimed at introductory-level users who may have had limited exposure to terms used in this field. Referencing links are also available for further readings where users can access in-depth supportive information.

The website ensures users' access of explicit instructions for operating the application. In "how it works" page, clear introduction of the structure of the system is presented to the users. With the purpose of each level and chapter elaborated, users can set their own learning procedures based on their specific needs.

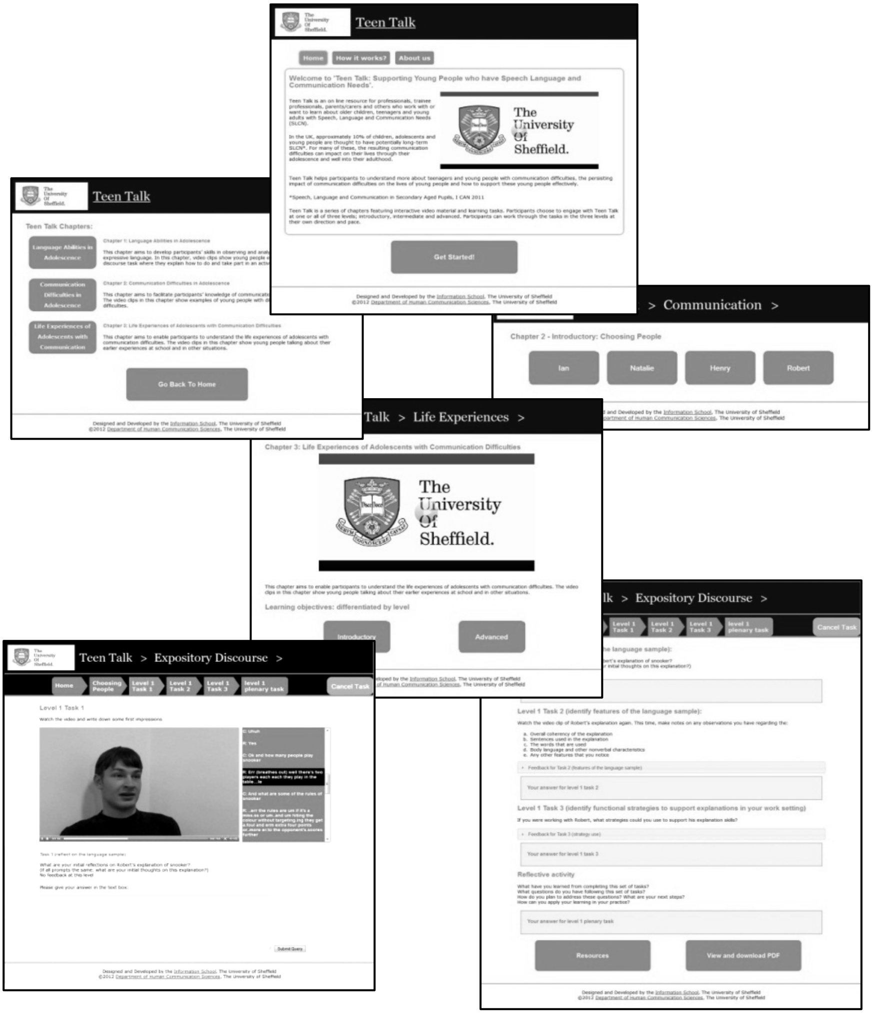


Fig. 5. User interfaces screenshots of Teen Talk

Fig. 5 illustrates Teen Talk as an e-learning artefact including examples of its different menus and functionalities.

4 Conclusion

In this paper, Teen Talk is presented as an e-learning artefact that was developed using a DSR approach. The process model of DSR was described in general and instantiated for this specific design and development research. It enabled the bridging between two very different academic cultures and resulted in a cohesive and, at times, enthusiastic collaboration. The use of rapid prototyping cycles enabled a better comprehension of potential and limitations of this type of artefact as well as a better understanding of its application as an e-learning tool. Teen Talk itself is currently being used both at HE and FE levels.

References

1. Shriberg, L.D., Tomblin, J.B., McSweeny, J.L.: Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. Journal of Speech, Language and Hearing Research 42, 1461 (1999)
2. Bercow, J.: The Bercow Report: A review of services for children and young people (0-19) with speech, language and communication needs (2008)
3. Clegg, J., Leyden, J., Stackhouse, J.: An Evaluation of Secondary Talk. Department of Human Communication Sciences. University of Sheffield (2011)
4. Paul, R.: Language Disorders from Infancy through Adolescence: Assessment & Intervention. Mosby Elsevier, USA (2007)
5. Gregory, J., Bryan, K.: Speech and language therapy intervention with a group of persistent and prolific young offenders in a non-custodial setting with previously undiagnosed speech, language and communication difficulties. International Journal of Language & Communication Disorders 46, 202–215 (2011)
6. Bryan, K., Freer, J., Furlong, C.: Language and communication difficulties in juvenile offenders. International Journal of Language & Communication Disorders 42, 505–520 (2007)
7. Clegg, J., Hollis, C., Mawhood, L., Rutter, M.: Developmental language disorders – a follow-up in later adult life. Cognitive, language and psychosocial outcomes. Journal of Child Psychology and Psychiatry 46, 128–149 (2005)
8. Clegg, J., Ansorge, L., Stackhouse, J., Donlan, C.: Developmental Communication Impairments in Adults: Outcomes and Life Experiences of Adults and Their Parents. Language, Speech, and Hearing Services in Schools 43, 521–535 (2012)
9. Whitehouse, A.J., Watt, H.J., Line, E., Bishop, D.V.: Adult psychosocial outcomes of children with specific language impairment, pragmatic language impairment and autism. International Journal of Language & Communication Disorders 44, 511–528 (2009)
10. Snowling, M.J., Bishop, D., Stothard, S.E., Chipchase, B., Kaplan, C.: Psychosocial outcomes at 15 years of children with a preschool history of speech‐language impairment. Journal of Child Psychology and Psychiatry 47, 759–765 (2006)
11. Dockrell, J., Lindsay, G., Palikara, O., Cullen, M.A.: Raising the achievements of children and young people with specific speech and language difficulties and other special educational needs through school to work and college (Research Report 837). Department of Education and Skills, London (2007)

12. Lindsay, G., Soloff, N., Law, J., Band, S., Peacey, N., Gascoigne, M., Radford, J.: Speech and language therapy services to education in England and Wales. International Journal of Language & Communication Disorders 37, 273–288 (2002)
13. Spencer, S., Hartshorne, M., Clegg, J.: Communication Difficulties and Youth Offending Teams: Working collaboratively to design, deliver and evaluate staff training. National Association of Professionals Concerned with Language Impairment in Children (NAPLIC) (2011)
14. Simon, H.A.: The Sciences of the Artificial. MIT Press, Cambridge (1996)
15. Takeda, H., Veerkamp, P., Tomiyama, T., Yoshikawa, H.: Modeling design processes. AI Magazine 11, 37–48 (1990)
16. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. Journal of Management Information Systems 24, 45–77 (2008)
17. Nunamaker, J.F., Chen, M., Purdin, T.D.M.: Systems development in information systems research. Journal of Management Information Systems 7, 89–106 (1991)
18. Eekels, J., Roozenburg, N.F.M.: A methodological comparison of the structures of scientific research and engineering design: their similarities and differences. Design Studies 12, 197–203 (1991)
19. Rossi, M., Sein, M.K.: Design Research Workshop: a Proactive Research Approach, `http://www.cis.gsu.edu/~emonod/epistemology/Sein%20and%20Rossi%20-%20design%20research%20-%20IRIS.pdf`
20. Cole, R., Purao, S., Rossi, M., Sein, M.K.: Being proactive: where action research meets design research. In: Conference Being Proactive: Where Action Research Meets Design Research, pp. 325–336. Association for Information Systems, Atlanta (Year)
21. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Quarterly 28, 75–105 (2004)
22. Hevner, A.R.: A three cycle view of design science research. Scandinavian Journal of Information Systems 19, 87–92 (2007)
23. March, S.T., Smith, G.F.: Design and natural science research on information technology. Decision Support Systems 15, 251–266 (1995)
24. Dockrell, J., Ricketts, J., Palikara, O., Charman, T., Lindsay, G.: Profiles of need and provision for children with language impairments and autism spectrum disorders in mainstream schools: A prospective study. DfE (2012)

Current Issues on Enterprise Architecture Implementation Methodology

Babak Darvish Rouhani[1], Mohd Nazri Mahrin[1], Fatemeh Nikpay[1], and Bita Darvish Rouhani[2]

[1] Advanced Informatics School, Universiti Teknologi Malaysia, Jalan Semarak, 54100, Kuala Lumpur, Malaysia
[2] ECE Department, Rice University 6100 Main St., MS 366, Houston, TX 77005, USA
{drbabak2,nfatemeh6}@live.utm.my, mdnazrim@utm.my, bd14@rice.edu

Abstract. Enterprise Architecture (EA) becomes a strategy plan for enterprises to align their business and Information Technology (IT). EA is developed, managed, and maintained through EA implementation methodology (EAIM) processes. There are some problems in current EAIMs that lead to ineffectiveness implementation of EA. This paper represents current issues on EAIM. In this regard, we set the framework in order to represent EAIM's issues based on the processes of EAIM including: modeling, developing, and maintaining. The results of this research not only increase the knowledge of EAIM, but also could be useful for both scientific and practitioner in order to realize the current situation of EAIMs.

Keywords: Enterprise Architecture, Enterprise Architecture Implementation Methodology, Current issues on EA implementation methodology, EA Implementation.

1 Introduction

Enterprise Architecture (EA) is employed by enterprises for providing integrated Information Systems (ISs) in order to support alignment of their business and Information Technology (IT) [1]. In EA the framework represents the structure to model enterprise's business and IT entities [2]. There are different models for various perspectives in EA Framework (EAF), each with different scope and activities [3]. The outcomes of EAF are EA's artifacts that consist of models, diagrams, documents and reports. Since EA artifacts are not sufficient for enterprises by they own, enterprises are looking to find a method to address theirs challenges on competiveness by implementing those artifacts. In addition, enterprises utilize EA to provide appropriate solution for their business's demands and it needs to implement EA models [4].

EA Implementation Methodology (EAIM) can describe the structured approach in order to solve some or all of the problems related to EA implementation [5].

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_23,

Moreover, it covers all aspects of the EA lifecycle - the planning for enterprise understanding projects, the analysis of business requirements, the design of systems, the evolution of systems, and the ongoing enhancements of all of the above [6]. The methodology is both complete and concise, serving as a coherent guide for practitioner professionals. It allows paths and pieces of content to be selected and extracted for application on specific projects [7].

The reminders of this paper are divided as following parts: EAIMs is describe in section 2, the research framework are represented in section 3, the results and discussions are stated in section 4 and 5 respectively, finally the conclusion of this study is expressed as section 6.

2 EA Implementation Methodologies

Basically, EA implementation contains three principal phases including: As-Is architecture, To-Be architecture, and migration plan. In As-Is architecture (also known as baseline, current, and initial architecture), EA will be defined current situation of business and IT of enterprise by means of set of definitions which illustrate the current state of the enterprise's mission, business processes and technology's infrastructure. The key role of this stage is vision of enterprise. In To-Be architecture (also known as desired, future, target architecture) EA will be represented the desired architecture including future of business and IT based on vision of enterprise. This type of architecture is the result of enterprise's long-term strategies and plans. The key role of this stage is to identify appropriate ISs. In EA migration plan (also known as transition plan) is the essential strategy that will be employed for Transition from the As-Is to the To-Be one. The key role of this stage is using the proper implementation method [6, 8, 9].

Well implemented EA helps a company innovate and change by providing both stability and flexibility [10]. There are several EAIMs. Although they are different in processes and phases of implementing, they are common in the concepts, principles, and reaching to To-Be architecture. The last phase in EA project phases is “Migration Plan”. It includes a strategic information base with a clear definition of business objectives and strategy. The strategy is needed for the transitional processes in order to implement new technologies in response to the changing business needs. That means the enterprise architecture includes also the process to create, update and manage the evolution of the architecture domains in line with business strategy. EA implementation methodologies focus on migration plan and provide some techniques and tactics to reaching to the To-Be architecture [8].

EA implementation method can be independent or dependent to a framework. While EA framework tries to capture information from enterprise's business and IT and model them, EA method tries to utilize models for developing appropriate ISs and IT Infrastructure for enterprise [11].

3 Research Framework

In order to obtain intended result we set the specific research framework. This research framework focuses on EAIM process including modeling, developing, and maintaining. These processes contain both generic EA attributes and features that are uniquely found in EAIMs.

Modeling Process: refers to those processes that are relating to portray of enterprise in various perspective and within different EA lifecycle. Typically, modeling comprises of three major components: symbols, syntax and semantics. Modeling different perspectives of enterprise are significant part of modeling that need to utilize in EAIM. Consequently, by using an appropriate modeling the EAIM could reduce the complexities of current and desired architecture, and transition plan effectively.

Developing Process: refers to those processes that are relating to developing needed ISs and IT infrastructure in order to cope with business challenges. It is importance for enterprises generally and for EAIMs particularly. A useful EAIM should cover the following stages, enterprise modeling, current architecture analysis, desired architecture analysis, managing and providing detailed design of projects, describing controlled transition plan, and implementation. EAIM that covers all parts of the EA development by considering EA concepts is a consistent and complete methodology.

Maintaining Process: refers to those processes that are relating to keeping and governing EA artifacts in order to increase implementing success. As mentioned above, the modeling is considered as a compulsory part of any EAIM. However, EAIM also emphasizes the series of activities and steps performed as part of the EA life cycle. Utilizing appropriate repository and governance process contribute to complete implementation of EA within enterprise by considering future changes.

Moreover, target studies that we investigated them are selected by searching on reliable databases, including: science direct, springer, IEEE, and ACM. The intended papers were selected by using the related search keywords, including "Enterprise Architecture Implementation", "Enterprise Architecture Implementation Methodology", "Enterprise Architecture Implementation Process", "Enterprise Architecture Modelling" and reading the whole parts of them in order to obtain appropriate studies. In addition, we consider the studies, which published after 2010 in order to collect latest issues on EAIM. This study aim is to collect current issues on EAIM based on defined research framework especially in current decade.

4 Results

This section represents collected information about EAIMs based on defined research framework.

4.1 Modeling Process

Since modeling process is employed within EA project in order to capture the essential components of enterprises and their relationships to serve documentation, analysis and developing purposes and the components and their relationships change over time, EA implementation becomes increasingly complex [12]. Current EAIMs do not deal with complexity of dynamic aspect sufficiently and this leads to inappropriate understanding of the future changes and difficult deployment of ISs [13, 14].

In addition, although several EA modeling process have been proposed by EAIMs that they are common in the idea of comprehensive supporting of EA management and development, it is difficult to select the right issues from them for a restricted development effort [15]. EAIMs also do not suggest complete tool for employing in EA implementation which it covers all required models and diagrams [12].

Furthermore, today's modeling practices are proprietary, time-consuming, and generally ineffective as tools for communicating strategic-level planning across and down all levels of enterprise. As a consequence it is difficult to learn and use [16, 17].

Finally, there are deficiencies in current EAIM which do not perfectly cover the modeling extraction process. Some of them are: lack of covering all basic components in EA plan, lack of separation in phases and extraction process, and lack of possibility to identify basic components and give priority over enterprises activities [18, 19].

4.2 Developing Process

Since developing process is employed within EA project in order to analysis and develop appropriate ISs and IT infrastructure to serve enterprise's business demands, the importance of developing process is to provide new IT solution or upgrade existing ISs in order to integrity between ISs according to enterprise's business [20]. In EA, enterprise business means all aspect of activities of enterprise including: marketing, human resource, IT, finance, and others, so EA is complex. Therefore EA project needs a comprehensive implementation methodology in order to support all needed process of mentioned domain from analyze (modeling) up to develop (implementation) [18, 21].

In addition, current EAIMs were developed in the context of enterprise integration rather than interoperability [22]. As a consequence it is difficult to make appropriate communication and decision. It is need to EAIM consider interoperability in a step-by-step manner [21, 23].

Moreover, most of EAIMs do not use recent IT and software engineering approaches for developing EA, like cloud computing, service-orientation, agent-oriented and broad security reference architectures, this leads to problems in designing and developing wide-ranging EA [24-26]. EAIMs also do not sufficiently cover dynamic aspects of EA projects including: managing, changes, and deployment [27, 28].

Finally, although, current EAIMs answer the demand of enterprises for alignment their Business and IT and how to adapting with ongoing changes and numerous

EAIM are developed, they are remain bloated, time-consuming and lacking in precision [20, 29].

4.3 Maintaining Process

Since maintaining process is employed within EA project in order to maintain and govern EA artifacts and project to serve better and ready for future business's changes, the problem of large companies is how to organize the processes of baseline EA development and change management under conditions of steady change that are introduced by implemented and changed IT solutions [30]. It could be a common situation when several IT projects are developing concurrently conducted by different enterprise’s part. A problem is how to keep up-to-date EA repository in such circumstances? Current EAIMs do not represent holistic plan and process in order to maintain and reuse the EA artifacts based on changes demand [15, 24].

In addition, current EAIMs do not represent the appropriate maintaining process; some of those represents implicitly process and not in detailed and others do not consider this process as whole [23]. As a result Table 1 shows the summary of current issues on EAIMs.

Table 1. Summary of finding

MODELING	DEVELOPING	MAINTAINING
Modeling dynamic aspect	Considering interoperability	Utilizing maintain process
Modeling tools	Using new IT and software engineering approaches	Keeping up to date EA repository
Time-consuming	Time-consuming	Detailed process
Detailed process	Control the complexity of development	Complexities of maintaining artifacts
Modeling the complexities		

5 Discussion

As mention in section four the complexity play the main role on all three defined processes of EAIMs. This section discusses more about these complexities in detail.

Some complexities have root on EA implementation methods. IT and business are even more complex in theirs processes, therefore bringing these two sides together in EA is a critical problem. Current EAIMs do not address complexity in any meaningful manner by providing models for complexity against which architectures can be validated [16, 25].

Since the involvement of heterogeneous stakeholder groups such as application owners, business developers, software developer, system analyzer, enterprise architect, and the others may create complexity requirements in an enterprise, an appropriate documentation and communication of the enterprise models are vital [15, 31].

Some complexities have root on EAIM's activities. Some EAIMs provide a number of analyses activities that may be employed in a transition plan or a list of EA application scenarios for which methods may be developed. However, this list is neither complete nor are its items disjunctive. So the complexities of understanding the way of transition remain [29]. On the other hand, an EAIM may be defined a method as a systematic aid that guides the transformation of a system from an As-Is architecture to a To-Be architecture, but it is unlikely that there is an EA method, which fits to every problem situation in the field. Instead it is advisable to adapt a current method or to use dedicated parts, like method components or method fragments [32].

There are several EAIMs that created by EA architects based on theirs experiences from preceding projects. The main reason is practitioner and Consultant Company are trying to use concise and efficient method rather current EAIMs to cope with complexities and project time. Another significant reason is current EAIMs are either too abstract therefore difficult to implement, or too extensive to be use in all project. Consequently, there is no scientific foundation behind these types of EAIMs and it is difficult to teach and govern them [29, 33].

Based on this research investigation on current EAIMs the effective EAIM should include the following features that shown in Table 2.

Table 2. The requirement process of complete EAIM

Alignment	Repository	Strategy Plan	Step by step guideline
Detailed design	Maintenance	Governance	Requirement
Modeling	Iterative	Continuum	Management
Business strategy	Business plan		

6 Conclusion

This paper represented current issues on EAIM. In this regards, we set the criteria in order to represent EAIM's issues based on the processes of EAIM including: modeling, developing, and maintaining. In order to obtain the primary studies the defined keywords were searched in selected databases. The significant issue on three process of defined framework criteria is complexities that come from different aspects of EA implementation. The following outcomes could be extracted from this research and they are useful for both scientific and practitioner in order to understand current situation of EAIMs:

- Controlling the complexity of EAIM contribute to effective EAIM
- Better managing the EA project provide appropriate development and deployment process
- Using effective process for each project (neither abstract nor detailed)
- Considering the time for each type of process (modeling, developing, and maintaining)
- Considering interoperability for making good decision on developing

References

1. Chen, D., Doumeingts, G., Vernadat, F.: Architectures for enterprise integration and interoperability: Past, present and future. Computers in Industry 59(7), 647–659 (2008)
2. Darvish Rouhani, B., et al.: Presenting a framework for agile enterprise architecture. IEEE (2008)
3. Ahlemann, F., et al.: EA frameworks, modelling and tools. In: Ahlemann, F., et al. (eds.) Strategic Enterprise Architecture Management, pp. 201–227. Springer, Heidelberg (2012)
4. Bente, S., Bombosch, U., Langade, S.: Collaborative enterprise architecture: enriching EA with lean, agile, and enterprise 2.0 practices 2013. Morgan Kaufmann, Waltham (2013)
5. Chung, H.M., McLeod, G.: Enterprise Architecture, Implementation, and Infrastructure Management. In: Proceedings of the 35th Hawaii International Conference on System Sciences 2002. IEEE, Hawaii (2002)
6. Aier, S., Saat, J.: Understanding processes for model-based enterprise transformation planning. International Journal of Internet and Enterprise Management 7(1), 84–103 (2011)
7. Ahlemann, F.: Strategic enterprise architecture management: challenges, best practices, and future developments. Management for professionals 2012, p. xvi, 298. Springer, Berlin (2012)
8. Engelsman, W., et al.: Extending enterprise architecture modelling with business goals and requirements. Enterprise Information Systems 5(1) (2011)
9. Jonkers, H., et al.: Enterprise architecture: Management tool and blueprint for the organisation. Information Systems Frontiers 8(2), 63–66 (2006)
10. Darvish Rouhani, B., Kharazmi, S.: Presenting new solution based on Business Architecture for Enterprise Architecture. International Journal of Computer Science 9(3), 5 (2012)
11. Goethals, F., et al.: An overview of enterprise architecture framework deliverables. In: Banda, R.K.J. (ed.) Enterprise Architecture-An Introduction (2006)
12. Engelsman, W., et al.: Extending enterprise architecture modelling with business goals and requirements. Enterprise Information Systems 5(1), 9–36 (2011)
13. Minli, J., Decai, K., WuLiang, P.: Research of information system technology architecture. In: 2010 2nd International Conference on Industrial and Information Systems (IIS) (2010)
14. Aier, S., Gleichauf, B., Saat, J., Winter, R.: Complexity Levels of Representing Dynamics in EA Planning. In: Albani, A., Barjis, J., Dietz, J.L.G. (eds.) CIAO! 2009. LNBIP, vol. 34, pp. 55–69. Springer, Heidelberg (2009)
15. Winter, K., et al.: Investigating the State-of-the-Art in Enterprise Architecture Management Methods in literature and Practice. In: MCIS, p. 90 (2010)
16. Davidrajuh, R., Davidrajuh, R.: Developing enterprise archiecture with "model-evaluate-extend" approach. In: 2010 International Conference on Educational and Information Technology (ICEIT). IEEE (2010)
17. Clark, T., Barn, B.S., Oussena, S.: A method for enterprise architecture alignment. In: Proper, E., Gaaloul, K., Harmsen, F., Wrycza, S. (eds.) PRET 2012. LNBIP, vol. 120, pp. 48–76. Springer, Heidelberg (2012)
18. Šaša, A., Krisper, M.: Enterprise architecture patterns for business process support analysis. Journal of Systems and Software 84(9), 1480–1506 (2011)
19. Rapacz, N., Pacyna, P., Sowa, G.: Building Domain-Specific Architecture Framework for Critical Information Infrastructure. In: Dziech, A., Czyżewski, A. (eds.) MCSS 2011. CCIS, vol. 149, pp. 261–267. Springer, Heidelberg (2011)

20. Noran, O.: Building a support framework for enterprise integration. Computers in Industry (2012)
21. Niemi, E.: Quality Attributes for Enterprise Architecture Processes. Enterprise Architecture, 8 (2013)
22. Hernández, J.E., Poler, R., Mula, J.: An Interoperable Enterprise Architecture to Support Decentralized Collaborative Planning Processes in Supply Chain Networks. In: Popplewell, K., et al. (eds.) Enterprise Interoperability IV, pp. 213–224. Springer, London (2010)
23. Nakakawa, A., van Bommel, P., Proper, E.: Towards a theory on collaborative decision making in enterprise architecture. In: Winter, R., Zhao, J.L., Aier, S. (eds.) DESRIST 2010. LNCS, vol. 6105, pp. 538–541. Springer, Heidelberg (2010)
24. Malta, P.M., Sousa, R.D.: Dialogical action research for better enterprise architecture implementation (2012)
25. Brunner, T., Zimmermann, A.: Pattern-oriented Enterprise Architecture Management. In: The Fourth International Conferences on Pervasive Patterns and Applications, PATTERNS 2012 (2012)
26. Darvish Rouhani, B., Nikpay, F.: Agent-Oriented Enterprise Architecture: new approach for Enterprise Architecture. IJCSI International Journal of Computer Science Issues 9(6), 4 (2012)
27. Medini, K., Bourey, J.P.: SCOR-based enterprise architecture methodology. International Journal of Computer Integrated Manufacturing 25(7), 594–607 (2012)
28. Coetsee, R., Brown, I.: Towards a Theory for Enterprise Architecture Governance in Organisations. In: Proceedings of the International Conference on Information Management & Evaluation, pp. 82–90 (2010)
29. Aier, S., Schelp, J.: A reassessment of enterprise architecture implementation. In: Dan, A., Gittler, F., Toumani, F. (eds.) ICSOC/ServiceWave 2009. LNCS, vol. 6275, pp. 35–47. Springer, Heidelberg (2010)
30. Song, H., Song, Y.-T.: Enterprise architecture institutionalization and assessment. In: 2010 IEEE/ACIS 9th International Conference on Computer and Information Science (ICIS). IEEE (2010)
31. Hoyland, C.A.: RQ-Tech, A Strategic-Level Approach for Conceptualizing Enterprise Architectures. Procedia Computer Science 12, 37–42 (2012)
32. Cuenca, L., Boza, A., Ortiz, A.: Enterprise Architecture Framework with Early Business/ICT Alignment for Extended Enterprises. In: Ortiz, Á., Franco, R.D., Gasquet, P.G. (eds.) BASYS 2010. IFIP AICT, vol. 322, pp. 11–18. Springer, Heidelberg (2010)
33. Agievich, V., et al.: A new approach for collaborative Enterprise Architecture development. In: 2012 7th International Forum on Strategic Technology (IFOST) (2012)

User Behavior Detection Based on Statistical Traffic Analysis for Thin Client Services

Mirko Suznjevic[1], Lea Skorin-Kapov[1], and Iztok Humar[2]

[1] University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia
{mirko.suznjevic,lea.skorin-kapov}@fer.hr
[2] University of Ljubljana, Faculty of Electrical Engineering, Tržaška 25, 1000 Ljubljana, Slovenia
iztok.humar@fe.uni-lj.si

Abstract. Remote desktop connection (RDC) services offer clients access to remote content and services, commonly used to access their working environment. With the advent of cloud-based services, an example use case is that of delivering virtual PCs to users in WAN environments. In this paper, we aim to analyze common user behavior when accessing RDC services. We first identify different behavioral categories, and conduct traffic analysis to determine a feature set to be used for classification purposes. We then propose a machine learning approach to be used for classifying behavior, and use this approach to classify a large number of real-world RDCs. Obtained results may be applied in the context of network resource planning, as well as in making Quality of Experience-driven resource allocation decisions.

Keywords: user behaviour, remote desktop connection, traffic classification, machine learning.

1 Introduction

Today numerous solutions exist on the market supporting remote desktop connections (RDC), including those provided by Microsoft, Oracle, and Citrix. With the advent of cloud computing and data centres offering virtual desktop solutions to end users, thin-client solutions are becoming an increasingly popular mechanism for end users to access and interact with remote content and services in an easy-to-maintain and cost effective manner.

A key challenge in making such solutions viable from an end user point of view is meeting the stringent network performance requirements dictating low delays and high response times [1]. As opposed to a standard, "local" desktop, whereby user inputs are locally processed and rendered nearly immediately, RDCs require inputs to be transmitted to a remote computer, processed, and returned to the thin-client [2]. Consequently, the screen updates and response times become a critical issue, in particular in WAN environments.

RDC traffic generally corresponds to encrypted video bitmaps exchanged between a remote virtual PC and a thin-client using a single server port, imposing

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_24,

challenges in terms of detecting end user tasks and applications that are being remotely run (e.g., a user editing a document, browsing the Web, or viewing audio/video content)[3]. RD protocols (e.g., Microsoft Remote Desktop Protocol, RDP) run over TCP and offer a reliable connection. Different user behaviours while using RDC (in terms of conducted task or application being used) result in different traffic characteristics and different impacts of network performance on user perceived quality [4]. Therefore, a prerequisite in effectively managing the end user Quality of Experience (QoE) is related to determining the user behavior in the context of RDC for the purpose of accurately mapping between QoE and performance indicators (i.e., delay, bandwidth).

In this paper, we focus on detecting user behavior while using RDC based on employing statistical traffic analysis and machine learning. We compare achieved results with those reported in related work [3]. Further, going beyond related work, we apply our proposed approach to real-world RDC traffic traces collected at the University of Ljubljana to study actual user behavior. Knowledge regarding such traffic characteristics may prove beneficial in making QoE estimations and resource planning decisions.

The paper is structured as follows. In section 3 we present a traffic classification algorithm for RDC employing a machine learning approach. In section 4, we use our algorithm to analyze actual user behavior captured in 18.5 GB of real RDC traces. Section 5 presents concluding remarks and future research.

2 Related Work

Traffic Classification

While approaches such as those based on Deep Packet Inspection have been commonly used to identify and classify network traffic [5], a key drawback lies in the fact that traffic may be encrypted, preventing payload analysis. Additional drawbacks include necessary knowledge regarding payload formats, and potential government imposed privacy regulations. Approaches based on statistical traffic analysis present a different approach, whereby relevant traffic features (e.g., packet length and inter-arrival times) are extracted and analysed to identify a particular application. Emmert et al. [6] previously reported on traffic patterns for different types of thin client users when working with popular office applications such as Microsoft Word, Excel, or PowerPoint. A thorough workload characterization and modeling (session arrival process, interarrival times, duration) of RDC traffic was done by Humar *et al.* [7]. The authors further confirmed self-similar characteristics of RDC traffic at packet level in [8].

In the past, Machine Learning (ML) algorithms have also been used for traffic classification purposes [9,10]. Dusi *et al.* [3] propose a novel approach to RD application identification based on IP-level statistical traffic analysis and machine learning techniques, their goal being to make QoE estimations targeted towards verifying that Service Level Agreements are met. In their tests, the authors ran several applications over an RDP connection, categorized as audio (e.g., VLC, WMP, skype), video (e.g., Flash video, WMP), and data (e.g., Web browsing, Adobe reader). The authors extracted traffic features (bit rate and packet rate at IP and TCP levels, TCP payload length, and number of observed packets) and evaluated different statistical classification techniques in terms of accuracy. Their results show that on-the-fly application identification for thin-client flows can achieve over 90 percent accuracy (when applying Support Vector Machines and Decision Tree algorithms, and considering a window size of 10 s) even when the testing set includes applications for which their classifier was not trained.

Relation to QoE Studies

Previous research has studied RDC traffic and its relation to user perceived QoE. The user experience on RDC was quantified by Tolia *et al.* [11] through operation response times on different bandwidths (10 and 100 Mbps) and for different types of office applications (typing, tracking changes, creating slides for presentation and manipulating photos). Staehle *et al.* [4] studied how different network parameters (packet loss, jitter and delay) influence subjective and objective QoE in a controlled testbed environment. Further, while Dusi *et al.* proposed a mapping between QoE scores and round trip time (RTT) thresholds for different types of RD applications, a more extensive study addressing the QoE of remote desktop users was conducted by Casas *et al.* [2]. In a laboratory study involving 52 participants, the authors looked to identify relationships between network performance parameters (i.e., delay, bandwidth) and end user subjective quality scores for four typical remote desktop tasks found in enterprise scenarios: text typing (e.g., email and document editing), screen scrolling (e.g., web browsing), drag and drop of images, and menu browsing (e.g., menu selection). The authors further provide a traffic characterization whereby they constitute that the four aforementioned tasks present nearly identical traffic patterns in the uplink, while in the downlink typing sends much less data than the other three tasks, which share a common packet size behavior. Differences in throughput are linked with different requirements with regards to sending screen updates. Given that RDC traffic uses TCP as a transport protocol, network delay has been found to have the most significant impact on end user QoE.

In this paper, we focus on user behavior prediction for RDC and build on the work reported by Dusi *et al.* [3]. Given that their results showed that the Decision Tree ML algorithm (tested using the freely available WEKA software [12]) provided very good results in terms of traffic classification accuracy, we have opted to use this algorithm in our study.

3 User Behavior Detection Based on Machine Learning

3.1 Methodology

The first step in our work was to train a ML algorithm to classify different types of RDC behavioral activities, to be subsequently used for the analysis of real RDC traffic traces. The behavior labelled traces which were used for training the Decision Tree algorithm were collected at the Department of Telecommunications, Faculty of Electrical Engineering and Computing, University of

Zagreb. The capture was performed on Dell Optiplex 390 computer (configuration: i3@3,3 GHz, 4GB RAM, ATI Radeon HD 6450) connected to another Dell Optiplex 390 via a 100 Mbit LAN. The behavior categories we have addressed are as follows:

- idle - no actions for 10 seconds, static screen (desktop image);
- document editing - editing of a word document (including text writing, picture pasting, text copying etc.);
- browsing - searching for accommodation on the www.booking.com web page;
- audio - listening to a 128bit/s online radio station from www.radio365.com;
- video - watching a 10 minute full-screen movie on www.youtube.com.

The measurements for each of the behavior categories lasted 10 minutes, i.e., during those 10 minutes, only a given action was being performed. We note that previous cited efforts [3,2] have not considered the category of *idle*, even though this is commonly observed in RDC traffic. The reason is that previous efforts focused either on identifying certain applications being run over an RDC or studying end user QoE, while our goal is to study overall user behavior exhibited when using RDCs. Following identification of a number of traffic features to be used for classification, we conducted a traffic analysis and specified a decision tree algorithm. We then used our collected traces to train the ML algorithm. Finally, we validated the algorithm using a validation dataset.

3.2 Traffic Feature Extraction

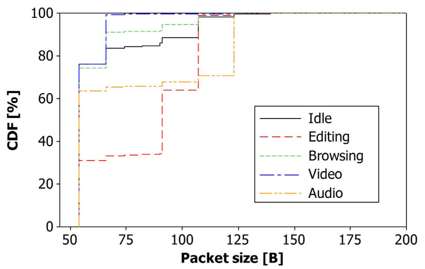
Features are attributes of flows calculated over multiple packets, used to train a ML classifier in associating sets of features with given application types [9]. Behavior labelled traces were processed such that for an epoch (time window) of 10 seconds, six features were extracted:

- up-link packet number;
- up-link average packet size;
- up-link average bandwidth usage;
- down-link packet number;
- down-link average packet size;
- down-link average bandwidth usage;

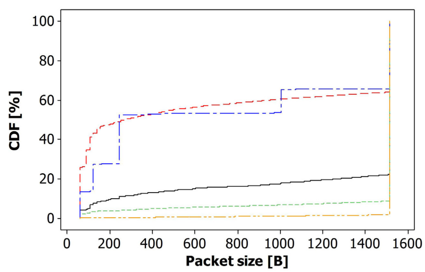
By *up-link* we refer to packets originating from the RDC client, while *down-link* refers to packets originating from the RDC server. The reason for choosing an epoch of 10 seconds is due to the fact that previous research has shown this approach as providing good results in terms of accuracy [3]. Traffic features were extracted from the network traffic traces using a custom built Java-based parser.

3.3 Machine Learning Approach

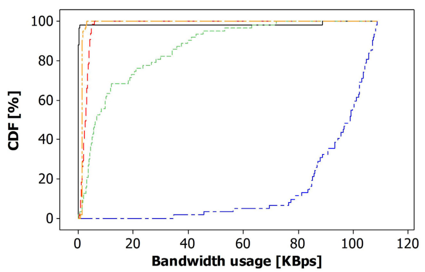
Following traffic capture, we analyzed the traces to determine different traffic patterns for the identified behavioral categories. With regards to the traffic characteristics, we note that the most distinguishing feature was found to be the number of down-link packets. Figure 1 portrays CDFs for packet size in both uplink and downlink directions, as well as CDFs for bandwidth usage across the different behavioral categories. In the case of traffic corresponding to the client being idle, only keep-alive packets were sent from server to client. As a result, we subsequently classified all traffic with <5 packet/s as idle. Video clearly represented the most resource demanding traffic, with average bit rates over 7 MBit/s. We once again note that we used full-screen video. Such high resource demands resulted in poor video quality in terms of image freezes and jerkiness, linked also to the fact that TCP usage is not optimized for real-time media. Browsing and editing exhibited highly variable traffic characteristics, with browsing resulting in higher down-link bandwidth utilization. Nevertheless, we presume that this is highly dependent on the end user interaction behavior resulting in screen updates. With regards to audio, we used a 128-bit radio channel and a static screen, i.e., there was no need for screen updates. We further note that all behavioral categories exhibit similar up-link characteristics (mostly composed of ACK packets), as also reported by Casas et al [2]. In the case of video, the higher bandwidth usage is due to the large number of ACKs being transmitted.



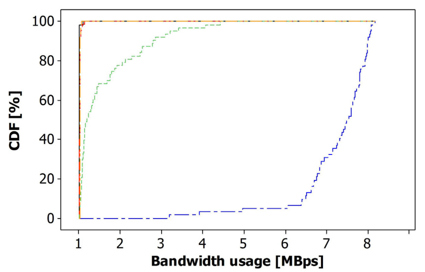
(a) Up-link packet size



(b) Down-link packet size



(c) Up-link bandwidth usage



(d) Down-link bandwidth usage

Fig. 1. Traffic analysis of RDC traces

Having statistically analyzed captured traffic, we used the freely available WEKA Java library (developed at the University of Waikato, New Zealand) [12]. While the WEKA software contains a large collection of tools for data processing, machine learning, and data analysis, we focused on support for implementation of the Decision Tree algorithm (noted in WEKA as J.48). The algorithm we specified in our study is portrayed in Figure 2, relying on the identified feature set for classification purposes. During the training phase, the algorithm uses extracted features from the traffic samples to build the mapping functions from the samples to the behavior category[1]. We used the approach proposed in [3] to split the samples into fixed time epochs (i.e., time windows) of 10 s and extract the features for each epoch. However, while in [3] the authors use only down-link traffic for algorithm training, we incorporate also up-link samples.

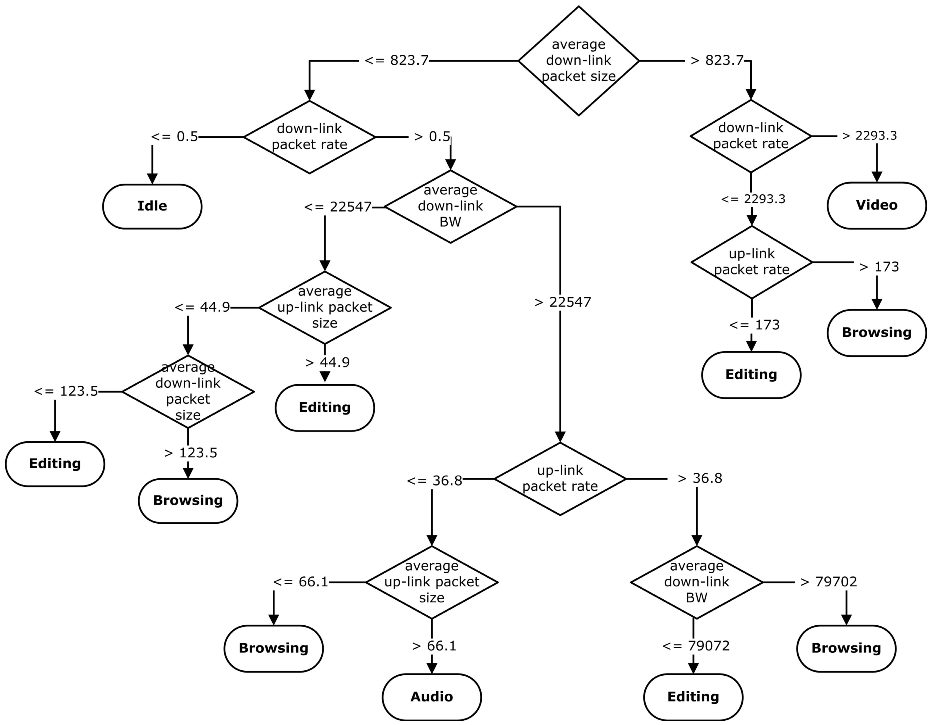


Fig. 2. A view of the Decision Tree algorithm used for traffic classification

Having completed the training phase, we generated an annotated dataset with the addressed behavioral categories for validation purposes. The accuracy in terms of correct classification (in terms of epochs) was found to be 78%, while the accuracy in terms of correctly classified bytes was found to be 89.7%.

[1] Interested readers may contact the authors for access to the data samples used for training purposes.

While this is somewhat lower than the accuracy reported by [3], we note that this may be attributable to the fact that we consider document editing and web browsing as different categories, whereas the cited authors considered a general “data” category (including web browsing, Adobe Reader, and Microsoft Powerpoint). Furthermore, the authors do not consider *idle* periods in their traffic classification. Finally, while the authors considered only down-link traffic, they used an extended feature set for classification purposes as compared to our set.

4 Analysis of Real RDC Traces

Following training and validation of the machine learning algorithm, we apply the algorithm to real RDC traces to study actual user behavior in a real world scenario (outside of a laboratory environment). To obtain empirical traffic traces, we collected packet-level traces on a 100 Mbit/s Ethernet link that connects the Faculty of Electrical Engineering, University of Ljubljana (FE) to the external Internet, as shown in Figure 3. The traces were collected from Sept. 24^{th}-Nov. 16^{th} 2012. We used a traffic capturing system on a personal computer running Linux OS and with a DAG network interface card. The traces were collected, filtered and stored in .cap format. We focused on the most frequently used RDC traffic, Microsoft Remote Desktop Protocol, which was recognized by transport layer (TCP) port numbers, configured to 3389 by default. To manage long traces, the individual files were limited to 0.5 GB. The traces contain RDC flows generated by students and staff of the Laboratory of Telecommunications, FE. The measurement procedure resulted in a total of 37 traffic traces which comprised 18.5 GB of RDC traffic. Each trace consisted of multiple RDCs and corresponding flows. A total of 1364 sessions were established when summarized across all traces (excluding sessions with a session length less than 10 s).

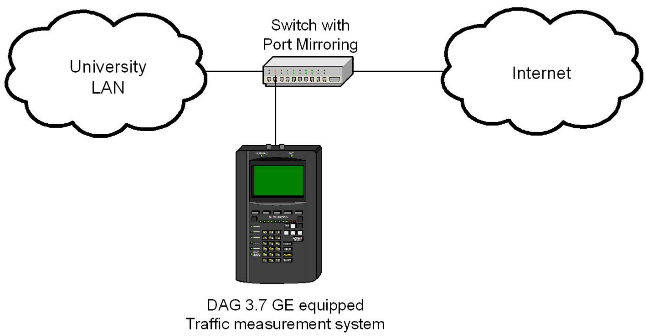


Fig. 3. Testbed environment for collecting RDC traces at FE, Ljubljana

Figure 4 shows the CDF for the different session lengths that were collected in the traces. Results show that while the majority of sessions were of short length,

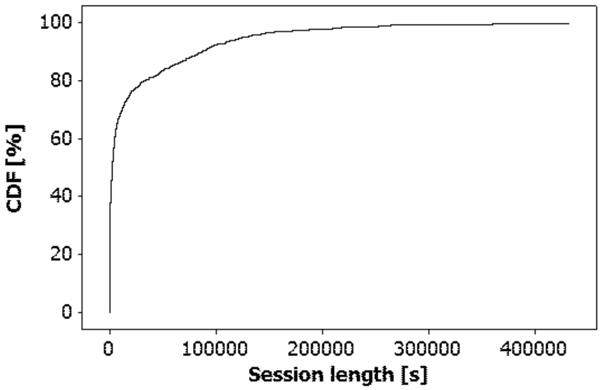


Fig. 4. CDF portraying session lengths for collected traces

there is a significant portion of longer sessions, some lasting up to multiple days. This is indicative of the fact that users established RDCs and left them open for extended periods of time, even though the connections were for the most part idle.

Figure 5 summarizes the duration of total time that users spent engaged in the given behaviors across all considered traffic traces. Results clearly show that for the majority of time (92.91 %), the RDC was idle, indicating that no end user activity was detected. Document editing (6.14 %) proved to be the most common activity, with small percentages of time spent browsing (0.89 %) and listening to audio (0.06 %). Throughout the entire collection of traces, there was no detection of behavior indicating that end users were running video applications. This supports the assumption that the users involved in this study

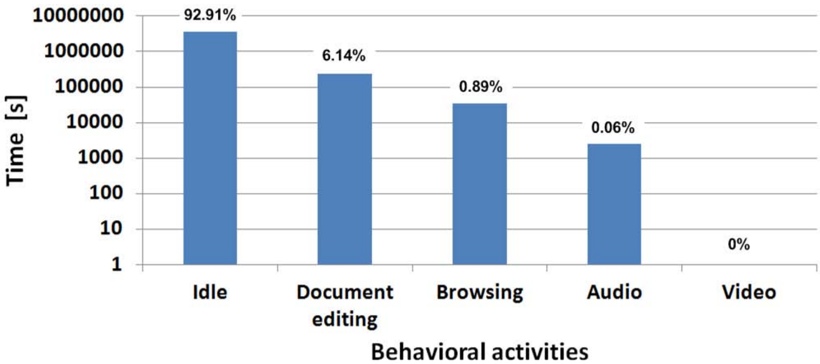


Fig. 5. Total time users spent engaged in a given behavioral category (logarithmic scale)

generally did not watch video content via a remote desktop connection, most likely due to both the nature of the tasks that users perform when using RDC, and also due to poor video quality resulting from network delays. While delay is not a determining issue in LAN scenarios, the use of RDC services in WAN scenarios clearly leads to the increased impact of RTT.

5 Conclusions and Future Work

The conducted study has served to provide insight into the actual behavior users engage in when using thin client services. The applicability of such results can be considered in the context of efficient planning and allocation of resources in a cloud environment, whereby passive monitoring of network traffic can be used to determine user behavior and make estimations with regards to network resource requirements. Further, by relying on studies that have explored QoE for RDC services [4,2,3], behavior detection can be used together with utility functions modeling QoE as a function of network delay and bandwidth for different types of remote desktop services to make QoE estimations. An example of recent work focused on exploiting RDC application identification using statistical mechanisms for the purpose of QoE-driven scheduling is discussed in [13].

In our work, we have employed a machine learning approach to study user behavior, with actual traces collected in an academic setting. Future work will focus on improving our classification accuracy, and further analyzing additional real world scenarios, such as those focused on use of thin client services in an enterprise/business setting.

Acknowledgements. This work was partially supported by the Ministry of Science, Education and Sports of the Republic of Croatia and the Ministry of Higher Education and Technology of the Republic of Slovenia under the joint Croatia-Slovenia Cooperation in Science and Technology Project "Services and Applications in the Future Internet: Ensuring Quality and Energy Efficiency".

References

1. Lai, A.M., Nieh, J.: On the Performance of Wide-Area Thin-Client Computing. ACM Transactions on Computer Systems (TOCS) 24(2), 175–209 (2006)
2. Casas, P., Seufert, M., Egger, S., Schatz, R.: Quality of Experience in Remote Virtual Desktop Services. In: 2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), pp. 1352–1357. IEEE (2013)
3. Dusi, M., Napolitano, S., Niccolini, S., Longo, S.: A Closer Look at Thin-Client Connections: Statistical Application Identification for QoE Detection. IEEE Communications Magazine 50(11), 195–202 (2012)
4. Staehle, B., Binzenhöfer, A., Schlosser, D., Boder, B.: Quantifying the Influence of Network Conditions on the Service Quality Experienced by a Thin Client User. In: 2008 14th GI/ITG Conference Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB), VDE, pp. 1–15 (2008)

5. Sen, S., Spatscheck, O., Wang, D.: Accurate, Scalable In-Network Identification of P2P Traffic Using Application Signatures. In: Proceedings of the 13th International Conference on World Wide Web, pp. 512–521. ACM (2004)
6. Emmert, B., Binzenhöfer, A., Schlosser, D., Weiß, M.: Source Traffic Characterization for Thin Client Based Office Applications. In: Pras, A., van Sinderen, M. (eds.) EUNICE 2007. LNCS, vol. 4606, pp. 86–94. Springer, Heidelberg (2007)
7. Humar, I., Bester, J., Tomazic, S.: Characterizing Graphical Desktop Sharing System's Workload in Collaborative Virtual Environments. In: Consumer Communications and Networking Conference, pp. 1–5. IEEE (2009)
8. Humar, I., Pustisek, M., Bester, J.: Evaluating Self-Similar Processes for Modeling Graphical Remote Desktop Systems' Network Traffic. In: 10th International Conference on Telecommunications, pp. 243–248. IEEE (2009)
9. Nguyen, T.T., Armitage, G.: A Survey of Techniques for Internet Traffic Classification Using Machine Learning. IEEE Communications Surveys & Tutorials 10(4), 56–76 (2008)
10. Park, B., Won, Y.J., Choi, M.J., Kim, M.S., Hong, J.W.: Empirical Analysis of Application-Level Traffic Classification Using Supervised Machine Learning. In: Ma, Y., Choi, D., Ata, S. (eds.) APNOMS 2008. LNCS, vol. 5297, pp. 474–477. Springer, Heidelberg (2008)
11. Tolia, N., Andersen, D.G., Satyanarayanan, M.: Quantifying Interactive User Experience on Thin Clients. Computer 39(3), 46–52 (2006)
12. University of Waikato: WEKA - Waikato Environment for Knowledge Analysis, `http://www.cs.waikato.ac.nz/ml/weka/`
13. Arumaithurai, M., Seedorf, J., Dusi, M., Monticelli, E., Lo Cigno, R.: Quality-of-Experience driven Acceleration of Thin Client Connections. In: 12th IEEE International Symposium on Network Computing and Applications (NCA), pp. 203–210. IEEE (2013)

Using Individual Interactions to Infer Group Interests and to Recommend Content for Groups in Public Spaces

Fernando Reinaldo Ribeiro[1,2], Patrick Santos[3], and José Metrôlho[1]

[1] Informatics Department, Polytechnic Institute of Castelo Branco, Castelo Branco, Portugal
[2] ALGORITMI Research Centre, University of Minho, Guimarães, Portugal
[3] TIMwe Lab, Covilhã, Portugal

Abstract. This work explores user's individual interactions combined with contextual information rules to support socially-aware content selection to best fit users' interests in public spaces. The system uses information from individual users to infer about the general interests of the place visitors and thus allowing the system to present the content that is representative of the interests of people who attend the same space. Two experiments were performed to evaluate the socially-aware proposed system and to analyze the users' opinions in different scenarios. Results from both experiments indicate that the proposed approach can be used to deliver content for individuals and for groups in public spaces and the users recognize its advantages over traditional information systems usually used in public spaces.

Keywords: adaptive systems, context-awareness, public displays, public recommender systems, socially-aware computing, ubiquitous computing.

1 Introduction

Nowadays, people in public spaces are constantly reached by a huge quantity of information (e.g. announcements, advices, news, etc.) which often do not pay enough attention because it is not in accordance with their interests, or it is not presented in the appropriate context or even because the content is not delivered using appropriate devices. Overpass this problem is not a simple task. Although we may think the presentation of the same content very frequently can overpass the problem of presenting the information to people that are interested in, it will also be annoying, probably presenting unexpected or misplaced content, leading to disaffection of users and consequently the system flop. To succeed, systems that broadcast content to people in public spaces need to be aware of the people interests and their environment because just so it is possible to present appropriate content and adequately answer to the needs of both content publishers and audience potentially interested in the content.

The proliferation of mobile computing devices and the even more common presence of large displays in public spaces has created new opportunities to develop a

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_25,

large range of novel systems to deliver content to people in public spaces, providing users with relevant information and giving to content publishers and advertisers new opportunities to promote their products or services. This has motivated advertisers and content publishers to develop systems that try to find the best advertising strategies to capture users' attention and thus to reach the target population that is potentially interested in their products or services. However, developing systems for delivering content to individuals and groups in public spaces poses many challenges that are not usually found when developing information systems for traditional media or online media, namely:

- The difficulty in obtaining information on the interests of the target audience. This is particularly important when recommendations are made for groups of people instead of individuals;
- The way content publishers specify the preferable context to present each content;
- How to prevent the presentation of serendipitous content;
- How to get user feedback on each presented content;
- The complexity of processing contextual information and how to use this information in an appropriate way;
- Which contextual information should be used and how it should influence the selection of each content.

In this work we describe a system that explores user's individual interactions combined with contextual information rules to support socially-aware content presentation to best fit users' interests in public spaces. Results, from two experiments, suggest that this approach presents some advantages over traditional information systems usually used in public spaces and that it can be used to deliver content for individuals and for groups in public spaces.

The remainder of this paper will be as follows. Section 2 presents a brief review of related work. Section 3 provides an overview of some guidelines about how to use personal interactions to infer about group interests in public spaces and describes the proposed approach for a public information system. In section 4 we describe the prototype we develop with the guidelines presented in section 3. Section 5 presents experimental work and the system evaluation. In section 6 we present and discuss the results and finally, in Section 7, we present some conclusions and we outline some of the future work.

2 Related Work

Several information systems, that aim to provide personalized content for individuals or groups in public spaces, have been presented and discussed in recent years. However, in this section we put special attention to systems that try to deliver appropriate content for groups of people instead of a person.

Many of them resort to personal mobile devices as part of the interaction with public displays. The *BluScreen* [1] collects audience statistics by exploiting Bluetooth to detect people's presence. It includes a public display which selects and displays ads in response to users detected in the audience. It uses Bluetooth-enable devices for identifying users and it explores history information of past users' exposure to certain sets of ads. Its main goal is to select the best content to maximize exposure to the current audience and thus ads are preferentially shown to those users that have not seen them yet. *PervAds* [2] tries to maximize the relevance of the ads sent to potential customers bases on a user profile. The relevance is determined by semantically matching the profile and the context of the customer against those carried by the ads. In this system ads are received in user's mobile devices and they are able to control the amount and the pertinence of the ads they receive. Muller [3, 4] describes a mechanism to adapt advertisements to the interests of the audience in public spaces. Here, each advertisement has a set of keywords and the history of all advertisements a user was interested in is kept. It uses voucher collection information as a form of feedback for the system. The advertisement is shown as a coupon which users can take a photograph with their mobile phone and show this photo to the cashier. The coupon contains a code that encodes where and when the user has seen the advertisement. This code is fed back into the system and it is used to influence the decision of which content to present.

The combination of personal mobile devices and public displays was also used in [1, 5]. They present a concept to allow people to individually influence public content such as songs played in shopping malls, news displayed on big displays, etc. based on the Bluetooth functionality in their mobile device. Users define once their preferences and store them encoded in the Bluetooth friendly name of their mobile phone. A system that also explores the use of mobile phones Bluetooth naming is presented in [6]. In this work they explore the creation of a public display system that evaluates the relevance of content from web sources and selects the most relevant content according to a dynamic tag cloud that incorporates static place definitions, but is also sensitive to the people around the display. In this system users interact with the display using keywords in the Bluetooth device name of their mobile phones.

The NFC technology is also used in many systems. Usually, NFC tags are placed in or on objects and are used to provide a link between the advertisers/publishers and the target audience through mobile devices [7, 8]. An advertising system using NFC is described in [7]. It combines public screens and mobile devices and explores personal user profiles and explicit input from users in order to provide appropriate advertisement content. In this work static displays such as paper posters and dynamic displays have been augmented with grids of NFC tags, which allow direct interaction with a mobile phone. They resort to the users' mobile phones to store information about their owners, track their activities, and let them interact directly with information on the displays. The content depends on the personal user profile stored on the phone and the current activities of the user.

In fact, there are several systems that try to provide personalized content to users' specific needs in public spaces. Some of them provide personalized content to individuals, others are trying to recommend content for groups. When content is targeted to individuals, many systems make use of user profiles to define the best content to deliver. However, people in public spaces is very heterogeneous and many times they are a single visit persons, and built individual user profiles is not feasible. When the content is targeted for groups, instead of individuals, approaches that are based on users' identification or on individual profiles are not appropriated and in these scenarios support their decisions in the social context that represents the interests of the place visitors can be a promising approach.

3 From Individual Interest to Group Interests

There are many systems that try to deliver personalized information to individuals in public spaces. The most common strategies make use of the users' personal mobile devices to identify their preferences. In these approaches users' preferences and interests are usually easy to identify because they result from the explicit interactions of the users or from the user identification. A very common way to specify user preferences is using personal mobile devices and, in these cases, there are several communication mechanisms to support them (e.g. SMS, Bluetooth, Wi-Fi, NFC, *qr-codes*). Using these devices users are able to send specific commands, to specify keywords denoting interest in some specific topic (e.g. SMS, Bluetooth) or read tags in products (e.g. NFC tags, *qr-codes*) and with this information systems are able to respond accordingly to their interests and thus deliver appropriate content or services. In this situations it is possible to make associations between services and/or products and user interactions, or build user profiles based on user past interactions. In fact, when the intention is to use information about users' individual interests it can be easily obtained through simple interactions and just based on this information the system is able to make individual recommendations. However, if the intention is to make recommendations for groups of people, instead of individuals, this ask is much more complex. In these scenarios the identification of their interests is not easy or obvious. Many of these people may not be able to interact with the system so as to expose their interests and because that it is more difficult to obtain information about their interests. But, without being aware of the activities and characteristics of the place and without knowing the preferences and interests of the users that frequent that space it is not possible to make appropriate decisions about what content present at each moment.

3.1 Place Social Interests

The main idea is to present in the display the items in which people are most interested in. To do this, we need some information about the general interests of the users that frequent the place. Thus, we propose a model to represent the place social interests that represents the interests of people that visit the same place. Thus, we use the individual users' interactions to obtain knowledge about the interests of people that frequent the same space, a place profile, and then use this knowledge to support decisions about which content is more relevant for the target people at each specific moment. This means that we use the individual explicit interactions of some users to infer about the general interests of the place visitors and thus allowing the system delivering items that are representative of the interests of people who attend the same space, to other people that visit the space and thus perhaps with similar interests, even without having interacted with the system. Using information related to the history of users' past interactions, topics they interact, number of times each topic is accessed, when they interact, recent items presented, etc. the system infers about place visit interests. This information is combined in a dynamic model to represent the place social interests that characterizes the interests of people that visit the same place. Additionally, the time period, that validates de place social interests, can be adjusted to turn the model less or more dynamic. A, i.e. a short time period makes the model more dynamic, because it includes in its behavior only recent interactions, and long time period, that makes the model less dynamic, incorporating in its behavior user interactions that happen for a long period and, consequently, less affected by recent interactions.

3.2 Contextual Rules

Obviously, knowing the context of the place where the content is presented and the interests of people that visit that place, are central features to decide which content is more appropriate to present for groups in a specific place. However, it is also important to take into consideration the characteristics of each content and the view point of the content publisher. Just so, it will be possible to frame each content in the environment where it will be presented. Thus, each content is characterized with a set of contextual rules that represent the preferable conditions under which the content gets more added value. Target places, date and time (the period of time where the item is valid), the preferable period of time to display, the main topic the item is associated to, the priority of each item, preferable characteristics of the place, or even weather, may help to identify the best moments to present each item.

4 Prototype

There are five main components in the prototype architecture: web portal, content server, large display to present content, a smart poster and a mobile application that allows users to visualize contents they request.

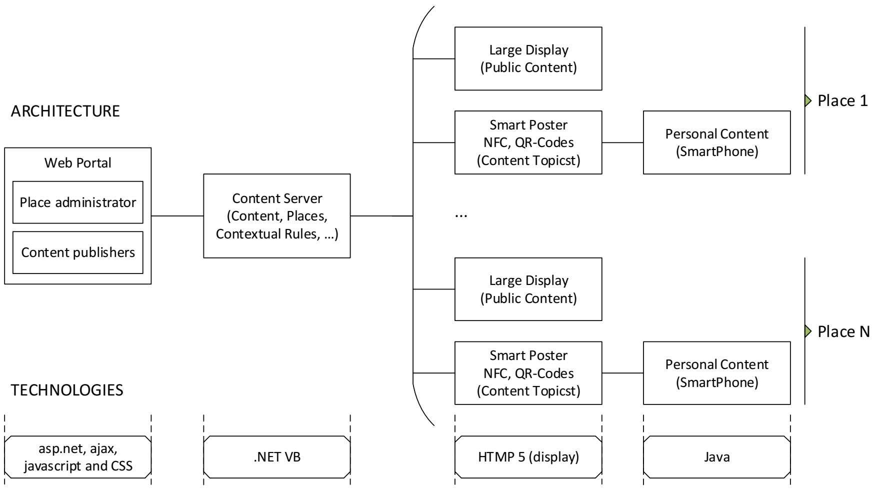


Fig. 1. System architecture and technologies

In the next section we describe each one of the main components of the system.

Web Portal: The Web portal is the back-office for content publishers and place administrators. For the place administrator it provides support to manage locations, configure displays and associate smart poster NFC tags and *qr-codes* to content topics. For content publishers it provides support to manage their content, namely: insert items (each item includes: title, description, image for the mobile phone, image for the display, start date, end date, target places, preferable days and periods, item main topic and priority level).

Content Server: The content server contains the logic associated to content, places and it is also responsible for the selection of the items to present in the display and the selection of the items to deliver at each mobile device.

Smart Poster: The smart poster is a poster that contains figures related to content main topics and a NFC tag and/or a *qr-code* associated to each topic. The NFC tag and/or the *qr-code* acts as a link between the figure of the topic and the content in the content server.

Large Display (Public Content): The large display presents the items that are targeted to people that frequents the public place. It includes three main areas: the left panel presents the name of the space and time; the main panel presents the current item and the bottom panel presents a general sentence about the system. The display presents the items in which people are most interested in. Thus, the selection of the items to present in the public display is supported by the place social profile according to the pool of available items to be presented in that specific place and their contextual rules. Thus, we calculate the relevance of each item, considering the recent place profile, and we verify its contextual rules and the history of presented items to avoid the presentation of the same item too often. Items with low level of relevance and consequently not presented in the display are included in the pool of items if there aren't enough items that satisfy the required contextual conditions to complete the queue. The combination of these conditions allow making a pool of items that is continuously updated according to the changes in context, place social interests and display history.

Mobile Phone (Personal Content): This application is responsible for reading the NFC tag or the *qr-code* (user selects the way they want interact) and it communicates with the content server to obtain the associated items and present them in a list according to their relevance. When one user selects a specific item from the list, a new screen with details about the selected item is presented and the user is able to sharing it through email and social networks (e.g. Twitter, Facebook).

First versions of our prototype only supported interactions through NFC. This decision was supported in the idea that NFC interactions are more spontaneous and less intrusive to the users. In this case, the users only need to install the app and it runs in background. Every time they want to interact with the smart poster, they only need to tap the intended topic. However, nowadays NFC technology is not present in many smartphones and because that a *qr-code* based interaction was included in the system.

The list of items is sorted according to the relevance of the items considering the context where they are displayed. This means that the selection of the items to be displayed in the mobile phone depends on the topic the user read (associated to the NFC tag or to the *qr-code*), the recent history of selected items in the users' smartphones, and also on the set of contextual rules specified by the content publisher for each item. These rules include: location, history of recent interaction (topics, items, and users), the period of time where the item is valid, the preferable period of time to display the item and the priority level of each item.

The communication between the different components of the system (public displays, mobile devices and Content server) is realized using SOAP web services and data is returned in JSON format.

5 Experimental Work and Evaluation

To evaluate different aspects of the proposed model, two experiments were performed in different scenarios. In both experiments, two different data gathering techniques to collect qualitative and quantitative data were used: a questionnaire [Q] to obtain quantitative ratings on a set of questions about the system usage and direct observations [O] taken during the experiments to obtain qualitative data. In both cases usage logs were also collected during the experiment time.

In addition, to the individual study of each one of these experiments, it is intended to analyze comparatively the findings that are obtained in different scenarios in order to uncover if the environments where experiments are performed have influence in the findings.

5.1 Experiment 1

The first experiment was performed in a common area of a technology company which is frequented by their workers and occasionally visited by other people outside the company. The system includes a smart poster showing three topics of content (Food, Technology and Telecommunications) and a public display presenting the content. In this case the system only supports interactions through NFC. The experiment was carried out for a week between 9am and 6pm of the working days and it was performed using a previous version of the prototype that only provided support to NFC based interactions (doesn't support *qr-codes* based interactions). A questionnaire was realized to evaluate users' perception about the sensitivity of the system to their interactions and to uncovering meaningful opinions about the relevance of the items presented in the public display (see table 1).

5.2 Experiment 2

The second experiment was performed in the common area of a school of technology. This prototype is installed in a transient place where students, teachers and staff walk by on their way to classes and dining hall. The system was configured with 4 topics (Academic issues, Events, Food and Jobs) and the experiment was carried out for three working days between 9am and 6pm. Users are able to interact with the smart poster using NFC or *qr-codes*.

The system includes a smart poster and a public display (see Fig. 2a and 2b).

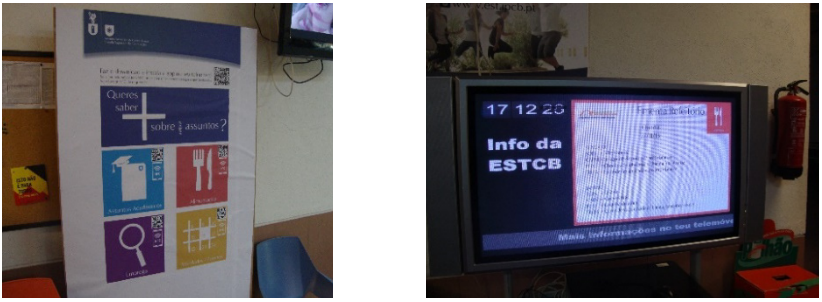


Fig. 2. a) Smart Poster. b) Public display.

When a user interacts with the smart poster, using his smartphone camera to read the *qr-code* or tap the intended topic to read the NFC tag, a list of items is presented in his smartphone (see Fig. 3a) and if he selects a specific item a new menu with detailed information about the item is presented.

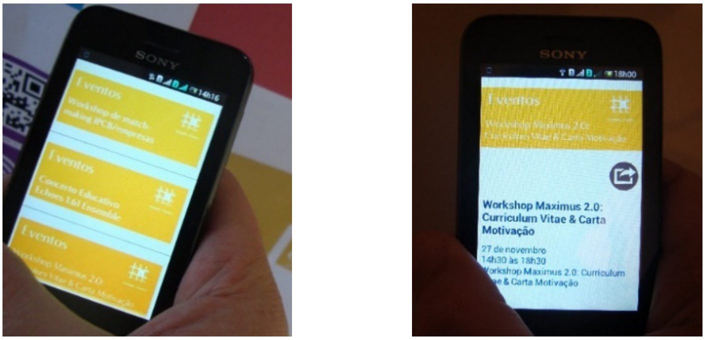


Fig. 3. a) List of items from a specific topic. b) Detailed information about a specific item.

5.3 Collected Data

During the experiments time statistical data about the system usage is collected (see table 1). In both cases the display was configured to present a new item every 15 seconds.

Table 1. Statistical data (Experiments 1 and 2)

Parameter	Experiment 1	Experiment 2
Number of distinct users	4	10
Number of interactions	89	48
Number of items clicked in mobile devices	155	68
Most viewed topic	Technology	Food

During the experiment time we asked to a set of users to fill a questionnaire about their opinion about the system usage. Each question was evaluated in a four levels scale (Good, Sufficient, Poor and Bad) and a *don't know* option. Table 2 shows the summary of the questionnaire. For each question, we present the total number of times the option is selected and the total number of answers for that question (excluding "*don't know*").

Table 2. Users' opinion about their experience (Experiments 1 and 2)

Parameter	**Experiment 1**	**Experiment 2**
Number of respondents	8	5
Users interact spontaneously with the system	Good (8/8)	Good (4/5) Sufficient (1/5)
Tasks complexity/laborious to configure mobile device to interact with the system	Good (6/8) Sufficient (2/8)	Good (3/5) Sufficient (1/5)
Clarity and readability of the items presented in the mobile phone	Good (6/8) Sufficient (2/8)	Good (5/5)
Advantages of this system over traditional systems (paper posters or cyclic content in public displays)	Good (8/8)	Good (3/3)
Relevance of the content in the public display	Good (8/8)	Good (4/4)

6 Results and Discussion

The first aspect concerns to the relatively low number of users that interact with the system in both experiments. In spite of, in both experiments the opinion of the users is similar, which suggests the system acceptance is not dependent of the place nature. In general, users had no problems understanding how the system works and how to configure their smartphones (how to download and install the app) to interact with the system [Q][O]. All the respondents quite or totally agreed that they did not have problems understanding the system and they all interacted with the system spontaneously and without help. With respect to the content delivered in the personal mobile phones, 75% in the first experiment and 100% in the second experiment of the respondents, refer that the content is presented with good clarity and readability. It is also clear that users find advantages of this system over traditional information systems usually used in public spaces. However, the relationship between individual user's interactions and the public display behavior is not clear perceived. Users recognize the relevance of the items presented in the public display but it is not clear if this is only a consequence of the set of available items or if different configurations of the place's social interests may affect the perceived relevance of the items. In this case it is important to refer that the system can be configured to adjust its behavior to be more reactive, with a more immediate influence of recent interactions, but consequently a more perceptible relation between users' interactions and the content in the display.

7 Conclusions and Future Work

In this work, we explored the usage of personal mobile devices and large public displays to present content for individuals and for groups in public spaces. Proposed system explores user's individual interactions combined with contextual information rules to support socially-aware content selection to best fit users' interests in public spaces. It allows the users to get content they are interested in, through simple interactions with a smart poster, and it uses information from individual users' interactions to characterize the place social interests. Place social interests are then used to support decisions about which content is more appropriate to present to the place visitors at each moment. In this approach the system uses information from individual users to infer about the general interests of the visitors of the same place and thus allowing the system to deliver items that are representative of the interests of people who attend the same space. To evaluate the proposed approach we perform two experiments in different scenarios. Users' opinions about the system are similar in both experiments and they suggest that the proposed approach can be used to deliver content for individuals and for groups in public spaces. Advantages of this system over traditional information systems usually used in public spaces are also recognized by users. Results, in both experiments, indicate that users easily interact with the system to get content related to their interests and that the content is presented with clarity in mobile phones. Users also recognize the relevance of the items presented in the public display. Despite of this, a more in depth study is needed to uncover if the perceived relevance of each item is only a consequence of the set of available items or if different configurations of the place's social interests may affect the perceived relevance of the items. Additionally, we need to perform more experiments with different types of public and to get more significant statistical results.

References

1. Payne, T., et al.: Auction Mechanisms for Efficient Advertisement Selection on Public Displays. In: European Conference on Artificial Intelligence. IOS Press, Riva del Garda (2006)
2. Orsi, L.C.A.G.: A New Perspective in Pervasive Advertising, p. 21. Oxford University (2011)
3. Müller, J., Kruger, A.: How much to bid in digital signage advertising auctions. In: Advances in Pervasive Computing. Adjunct Proceedings of the 5th International Conference on Pervasive Computing, Toronto (2007)
4. Müller, J., Kruger, A.: User Profiling for Generating Bids in Digital Signage Advertising Auctions. In: International Workshop on Ubiquitous and Decentralized User Modeling, 11th International Conference on User Modeling, Corfu (2007)
5. Mahato, H., et al.: Implicit personalization of public environments using bluetooth. In: Extended Abstracts on Human Factors in Computing Systems, Florence (2008)
6. Ribeiro, F., José, R.: Smart content selection for public displays in ambient intelligence environments. International Journal of Ambient Computing and Intelligence 5(2) (2013)

7. Holleis, P., Broll, G., Böhm, S.: Advertising with NFC. In: Workshop on Pervasive Advertising and Shopping, in Conjunction with the 8th International Conference on Pervasive Computing, Helsinki (2010)
8. O'Neill, E., Thompson, P., Garzonis, S., Warr, A.: Reach out and touch: using NFC and 2D barcodes for service discovery and interaction with mobile devices. In: LaMarca, A., Langheinrich, M., Truong, K.N. (eds.) Pervasive 2007. LNCS, vol. 4480, pp. 19–36. Springer, Heidelberg (2007)

A Distributed Architecture for Remote Validation of Software Licenses Using USB/IP Protocol

Mário J. Antunes[1,2], Albano Afonso[1], and Filipe Mota Pinto[1,3]

[1] School of Technology and Management, Polytechnic Institute of Leiria
Morro do Lena, Alto do Vieiro, 2411-901 Leiria, Portugal
{mario.antunes,fpinto}@ipleiria.pt, 2091357@my.ipleiria.pt
[2] Center for Research in Advanced Computing Systems (CRACS)
Faculty of Science, University of Porto, R Campo Alegre, 1021/1055 4169-007 Porto Portugal
mantunes@dcc.fc.up.pt
[3] Centro Algoritmi - University of Minho, Campus Azurém, 4800-058 Guimarães, Portugal

Abstract. USB dongles have been used by a wide range of software manufacturers to store a copy-protected of their application's license. The licenses validation procedure through USB dongles faces several concerns, as the risks of theft or losing dongle. Also, in scenarios where the number of dongles is reduced, users may have to wait for dongle access, which may lead to loss of productivity. In this paper we propose a client/server distributed architecture for remote software licenses validation, through USB/IP protocol. The proposed approach aims to take advantage of USB/IP for distributed access to a set of USB dongles physically connected to a remote USB server, over a TCP/IP network. We describe the deployment and enhancements made to an existing open source USB/IP implementation and also present the results obtained with this architecture in a real world scenario, for validation of computer forensics applications licenses that uses USB dongles.

Keywords: USB/IP, dongle USB, software license, distributed systems.

1 Introduction

A dongle is a device with an USB interface, used to store a copy-protected software license. There is a wide range of such applications that use exclusively USB dongles for their license validation. Those software applications are usually very expensive and commonly used in vertical markets, in which vendors offer specific services to an industry or group of customers with particular needs. Some examples are CAD/CAM applications (e.g. AutoCAD©), some business retail software, prepress and printing software and some imaging and scanning equipments and applications used in healthcare business.

The main concerns regarding the use of USB dongles are related to the risks of theft and loss, which may grant access to the applications by unauthorized users. In dynamic organizations, with several users geographically dispersed and accessing to

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_26, © Springer International Publishing Switzerland 2014

the same application, by having less USB dongles than the number of users that need to use the application, may lead to a decreasing of productivity, since each user has to receive the dongle from another user that previously used it and no longer need it.

In this paper we propose an open source distributed architecture for remote access to USB dongles physically connected to a server, through a local TCP/IP network [1]. In our work we take advantage of an open source implementation of the USB/IP protocol [1], commonly used to access to USB devices, like external disk or pen devices, that are attached to a remote computer. The underpinning ideal is to physically attach the USB dongles in a server running USB/IP protocol and accepting requests from the USB/IP clients. The USB/IP client applications run on the computers where applications are installed and to whose we need to validate the corresponding license. The overall remote license validation process is transparent to the users and they did not have to carry the USB dongles with them. Moreover, this procedure keep the legal commitments previously agreed, regarding software licensing (e.g. number of allowed users), since only one user is allowed to access each USB dongle (license) in a given moment. In our work we have also developed a graphical user interface for the USB/IP client, which improved the way each USB/IP client manages the virtual connections to USB devices through USB/IP protocol. The distributed architecture proposed in this paper aims to overcome the major limitations of commercial applications, as it is open source and the client application was designed to be customized for specific companies and application domains.

The existing commercial applications for remote access to USB devices through USB/IP reveal some drawbacks. That is, their licenses are expensive and charged in a per dongle basis. As the number of attached USB dongles increase so the value of the license. Those applications also offer a general interface to access USB device contents and cannot be customized for specific applications, like licenses validation.

Computer forensics applications have been used as a case study in our work. These applications operate in a very specific field and are used to validate and analyze digital evidences collected by criminal police searches [2]. The major challenge in police departments dedicated to analyze the collected evidences is to assure the feasibility of digital evidences under investigation. In order to help police examiners on analyzing digital evidences collection, there are two well known and commercial software computers forensics applications widely used by police departments, namely Forensic Toolkit (FTK)© [3] and EnCase© [4]. For both applications, the license is stored in an USB dongle and when the user intends to gain access to the software application, attach it to an USB port and wait for the corresponding permission. If there is no valid license in USB dongle, the user will not have access to the application [5].

We have made several acceptance tests in the computers forensics laboratory of Portuguese criminal Police, namely on validating the remote access to FTK© and EnCase© applications licenses. The deployed system is now being used experimentally by the computers forensics laboratory in their daily investigation routine tasks.

The rest of the paper is organized as follows. In Section 2 we introduce the necessary background to our work, namely the existing USB/IP applications and the protocol USB/IP. Section 3 presents the proposed architecture, followed by the development made and experimental results in Sections 4 and 5 respectively. Finally, Section 6 presents some conclusions and future work.

2 Background

In this section we present the background on USB/IP protocol and implementations, which constitutes the generic knowledge for understanding the approach proposed in this paper.

2.1 USB/IP Applications

There are a wide range of applications that tries to overcome the physical limitation to access to USB ports, by using a remote connection through client/server USB/IP protocol. Generally, the USB devices are connected to the USB/IP server and exported for further network access by USB/IP clients.

Hirofuchi, et al. [6] firstly proposed an open source application to have a remote access to USB ports. In this first USB/IP implementation, USB devices were shared in a TCP/IP local network, in which both client and server applications ran in Linux operating systems. In 2007, Kwon et al. [7] developed a new implementation of USB/IP server with an extension for interoperability with Microsoft Windows USB/IP client. Later on, in 2009 the first USB/IP in Widows was developed, based on the commands execution in a command line interface [8]. In 2011, an enhanced version of USB/IP client for Microsoft Windows was released, with the support of ReactOS. The new features include the development of digitally signed drivers and the compatibility with 32 and 64 bits Microsoft Windows operating systems [8].

The Linux kernel versions have evolved and existing USB/IP implementations became obsolete and incompatible with early Linux versions. However, in 2011 USB/IP source code was integrated natively in Linux kernel 2.6 drivers folders and was improved in its performance and with new features [8]. At the moment there is an USB/IP implementation for Linux kernel 2.6, available in the software repositories, which is however obsolete and incompatible with more recent Linux versions. Moreover, the actual version available has one important limitation related with the fully release of the remote USB device. When the client release the virtual connection with the remote USB device, the server unbind the device and make it unavailable and not ready to be used after that by a different user [8]. There is also some implementations of USB/IB for mobile networks. In [9] the authors propose a Android mobile implementation using a mobile phone as a wireless USB storage device through USB/IP connection.

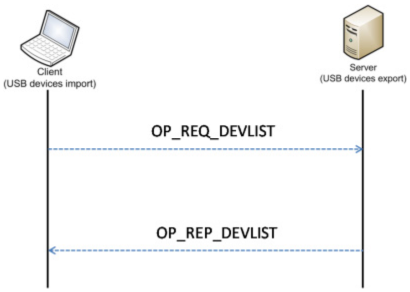
There are some implementations for Microsoft Windows based systems, which enable virtual emulation of the USB ports between two or more computers [10]. Regarding commercial applications, the licenses are increasingly expensive and are related with the number of dongles provided by the server and with the clients accessing to it. Some examples are the USB Network Gate for Windows, USB Redirector and USB over Network. USB Redirector is developed by Incentives Pro [11] and allows the use of shared USB devices remotely on the local network, wide area network or Internet as if the devices were physically connected to the remote computer. It uses a TCP/IP connection and can act both as a server and as a client. USB over Network is developed by FabulaTech [12]. This application assures the access to local USB devices by other computers on the network that have the client application. It supports a wide range of USB devices, encrypted data communications, compatibility with 32-bit and 64-bit systems and also with Hyper-V.

2.2 USB/IP Protocol

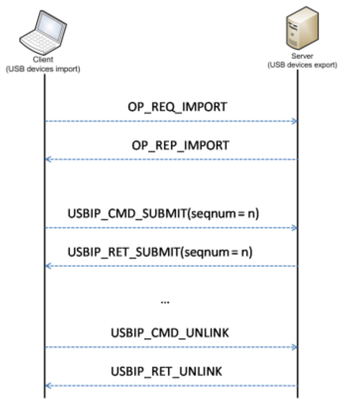
USB/IP protocol operates in a client/server model. The server exports the USB devices physically connected that are going to be used by the USB/IP clients. Moreover, the exported USB device drivers run on the USB/IP client. Figure 1(a) illustrates the request/respond USB/IP protocol messages. The client requests to the server a list of exported USB devices, by establishing a TCP/IP connection and by sending an `OP_REQ_DEVLIST` message. The server sends back a response with an `OP_REP_DEVLIST` which contains the list of its exported USB devices. After that, USB/IP client finishes the TCP/IP connection.

After obtaining the list of exported USB devices, the client may choose to use one of them, by establishing another TCP/IP session with the server. In this new session, client sends an `OP_REQ_IMPORT` message and the server answers with a corresponding `OP_REP_IMPORT` message. If the connection was succeeded, the TCP/IP connection remains established and will be used to transfer "USB Request Block (URBs)" [13] between USB/IP client and server.

USB/IP client may then send two types of messages: `USBIP_CMD_SUBMIT` to submit an URB; `USBIP_CMD_UNLINK` to unbind the session defined with the remote USB device. USB/IP server answers with an `USBIP_RET_SUBMIT` or `USBIP_RET_UNLINK` respectively. The process of importing an USB device is depicted in Figure 1(b). While the USB/IP server does not receive an `USBIP_CMD_UNLINK` message, USB messages are encapsulated in IP packets and sent by the TCP/IP network.



(a)



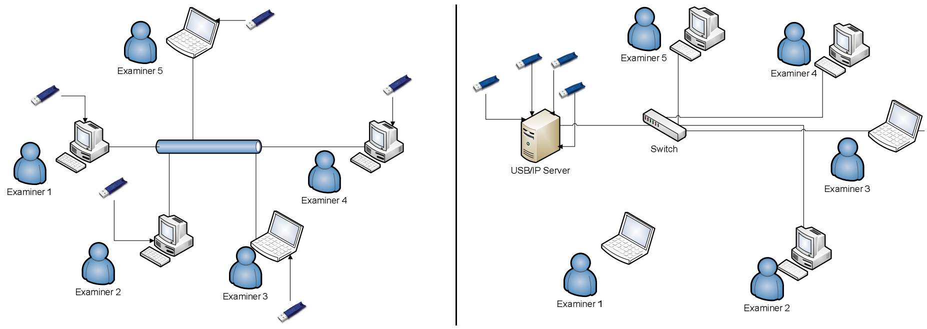
(b)

Fig. 1. (a) Request/Respond messages for obtaining a list of exported USB. (b) Messages involved on importing an USB device.

3 Proposed USB/IP Distributed Architecture

In this section we present the proposed open source architecture for remote validation of software license, by using USB dongles. This architecture is based on a distributed solution with remote access over USP/IP in a local TCP/IP network.

Our test scenario focused on the remote software licenses validation of computers forensics applications in use at Portuguese criminal Police. Figure 2(a) depicts the previous situation, in which the forensics applications were installed in each of examiner's computers and their license validation process was made locally through the USB dongle that stores a copy-protected license.



(a) (b)

Fig. 2. (a) Actual scenario with USB dongles physically connected to examiner's computers; (b) Proposed distributed architecture based on USB/IP client/server protocol

In this scenario, whenever the number of required active licenses is greater than the available dongles, it is introduced an "*investigation wait-state*" or "*shares-resource*" status between examiners. The proposed and developed scenario is depicted in Figure 2(b). It has a central server that physically holds the USB bus root, accessible from elsewhere in the local network. USB/IP clients are those computers running forensics applications. Those machines use Microsoft Windows operating system with previously installed USB/IP drivers and the USB/IP client application that will interact with USB/IP server and it will manage all applications access requests throughout TCP/IP network. The USB/IP server application remains waiting for requests at TCP port 3240. The main security concern is to define a firewall rule that may allow USB/IP clients machines to have a service access to this TCP port.

Within this distributed architecture, the numbers of required dongles are exactly the same as in previous situation, in respect to the license agreement contract. However, in this distributed scenario we may not have an "*investigation wait-state*" or "*shares-resource*" status between the examiners.

4 Client and Server Development

In this section we describe the development issues involved in the proposed distributed architecture, described in section 3. We present the major steps followed and actions taken on the development of client and server applications.

The development has been based on an already existing open source USB/IP project [6], as described on section 2. This project was initially conceived for a Linux kernel version 2.6, for both client and server applications. This approach faced some problems regarding the virtual USB connection management, which turned it in a non reliable version [8]. That is, whenever a client had ended his virtual connection to a USB device, the corresponding binding in the server was automatically removed, as well as from the exported devices list. In order to overcome this problem, it was necessary to share again the same device on the server and export it to the network. In our development we solved these issues and made available a new version of USB/IP server for Linux kernel version 3.2 [1].

The existing USB/IP client for Microsoft Windows was based on command line and thus not user friendly. Aiming to improve the application usability and user interaction with server application, we have developed a general and customizable graphical user interface (GUI) for USB/IP client, depicted in Figure 3.

Fig. 3. Generic graphical user interface (GUI) developed for USB/IP client

The development was made in Microsoft Windows .Net 4.5 Framework. We have used C and Visual C++ programming languages. The former was used for previous existing code re-use and the later for GUI development. We have used a Dynamic-Link Library (DLL) for encapsulating the existing C language code. For the graphical user interface development we have used the Platform Invoke [14] functionalities of C++ programming language for access to encapsulated C code.

5 Experimental Tests and Results Analysis

Figure 4 depicts the setup used in our tests made at computers forensics laboratory of the Portuguese criminal Police. It is composed by a client and a server interconnected by a switch, in the same IP network. In our experiments we have used the following USB dongles: CodeMeter [15] that stores FTK© software license and Aladdin HASP HL [16], which has the EnCase© software license. USB/IP server runs in an Ubuntu 12.04 LTS system, while the USB/IP client is installed in a Microsoft Windows 7 system, in 64 bits. USB/IP client system has EnCase© and CodeMeter Control Center© applications, being the later responsible for FTK© users license management.

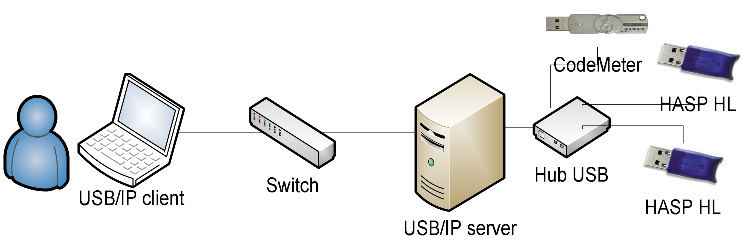


Fig. 4. Physical topology

In this test users were able to establish a TCP/IP network connection to the USB/IP server in which the USB dongles were previously exported. The graphical user interface customized to this particular scenario allowed a simple and easy to use USB/IP dongles management, as depicted in Figure 5.

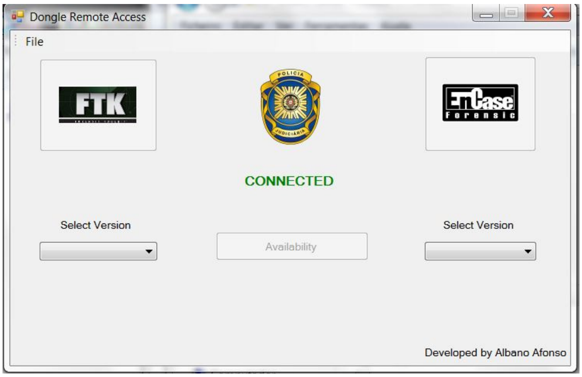
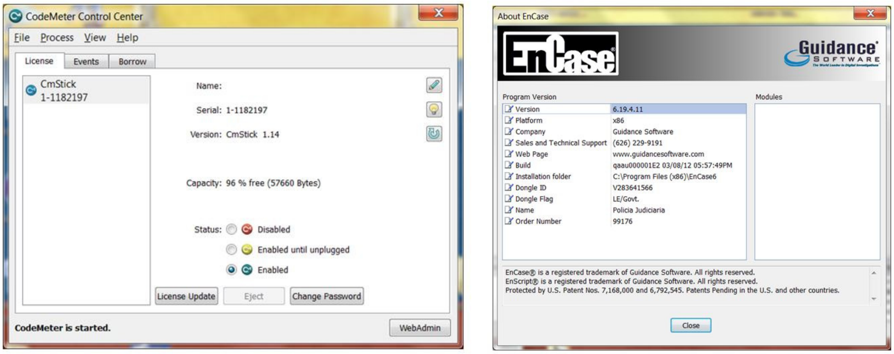


Fig. 5. Graphical user interface (GUI) adapted for the USB dongles management related with computers forensics applications used in criminal Police department

After the TCP/IP connection has been established between the USB/IP client and the remote USB dongles, the applications FTK© and EnCase© have successfully validated the user license and allowed the user access. Figure 6(a) depicts the report obtained by CodeMeter Control Center©, after enabling the user license for FTK©. In Figure 6(b) we present the details of user license obtained by EnCase©. The overall process was transparent to the police examiner, since he is able to validate its license stored in the dongle and gain access to the computers forensic application.



(a) (b)

Fig. 6. (a) CodeMeter Control Center© report after detected license stored in remote USB dongle; (b) General information about the EnCase© user license found in the remote USB *dongle*.

Table 1 summarizes the comparison between the three types of access to USB dongles: local access and distributed access with USB/IP protocol, which may be assured by commercial or open source applications.

Table 1. Comparison of local, commercial and open source distributed versions

	Local access	**Commercial applications**	**OpenSource applications**
Access Performance	USB resources sharing. Physical access per user	Multi-user environment with transparent access to different devices	
Productivity	Must wait for physical access to dongles.	Distributed access with an increase of productivity	
Risks	Theft, loss and unauthorized access.	Possible interception of USB/IP packets.	
License	Operating system license	Charged by the number of remote dongles and accesses.	GPL license for USB/IP application with free access
Client application	No specific client application required	Generic front-end USB/IP client	Customizable front-end USB/IP client

An important advantage of using a distributed approach for remote license validation is that we may mitigate the risk of loss and theft, as the user did not carry the USB dongle with him and thus did not have to connect it in the PC running the application. There are also some scenarios where we may have a higher number of users when comparing with the number of available USB dongles. In these situations, a user has to wait for a free dongle, which will be handed to him by another user who does not need it anymore. In some companies, these constraints may cause flaws in management and loss of productivity, which are mitigated by a distributed USB/IP implementation. The approach proposed in this paper, as it is open source, has a GPL license and thus with no associated cost. It is worth noting that in USB/IP distributed approaches, all the legal requirements to use the licenses are fully attained, as we only have one user using a license per user at each given moment.

6 Conclusions and Future Work

In this paper we have proposed a client/server distributed architecture based on USB/IP protocol for remote validation of copy-protected licenses stored in USB dongles. We have tested this distributed architecture in a real world scenario that uses two distinct computers forensics software applications for digital analysis of evidences by the Portuguese criminal Polices, namely FTK© and EnCase©. From the literature review, we did not find any open source USB/IP implementation to deal with remote validation of software licenses stores in USB dongles.

For that purpose we have deployed and developed a distributed USB/IP architecture for remote access to USB dongles available in an USB/IP server. The developments were made to an existing and outdated open source implementation of USB/IP protocol. In the server side we upgraded the existing USB/IP version for newly Linux 3.2 kernel version. Regarding the USB/IP client side, our development was two-fold. On one hand we have updated the existing version of the client for Linux kernel in version 3.2. On the other hand we have made available a USB/IP client version for Microsoft Windows 7 (64 bits), with a friendly and easy to use graphical user interface that allows the management of both the several USB dongles available in the server and also the virtual connections of those bounded in the PC.

The experiments were made in the computers forensics laboratory of Portuguese criminal Police, both with external devices (e.g. disks, pens, cameras) and with USB dongles which store copy-protected software licenses of computer forensics applications. The results were very promising, since we have succeeded on validating the user access to computer forensics applications by accessing to a remote USB dongle with the license. In an operational point of view, the proposed architecture have had a great impact in the examiners' daily routine, as they did not have to carry the USB dongles and may easily access to the licenses. In organizational terms, this distributed architecture means more flexibility and more physical security in the management of USB dongles, avoiding the risk of loss or theft. In particular, the architecture and developments presented in this paper are now being implemented in computers forensics laboratory of Portuguese criminal police.

The distributed architecture proposed in this paper is based on open source components, available with a GPL license. This feature brings also some benefits when comparing with existing commercial solutions, as it is free of charge and easily adjustable to different business environments.

Our future work will include tests in a wider network, in which client and server belong to different IP networks. We also aim to evaluate the impact of using USB/IP protocol to extend a cloud infrastructure. That is, the USB peripherals connected to an USB/IP server can be remotely accessed in a cloud environment and benefits for being in such environment. We are also evaluating this distributed USB/IP infrastructure on Raspberry Pi computers, in order to identify the advantages and challenges on having such an implementation on this kind of computers.

Acknowledgements. This work is financed by the ERDF – European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project FCOMP-01-0124-FEDER-037281.

References

[1] Afonso, A.: USBport Through TCP/IP - USB ports replication. MSc thesis in Informatics and Multimedia, by Polytechnic Institute of Leiria (November 2012)

[2] Casey, E.: Handbook of Digital Forensics and Investigation. Academic Press (2010)

[3] AccessData FTK (December 2013), `https://ad-pdf.s3.amazonaws.com/FTK_DataSheet_web.pdf`

[4] Guidance Encase (November 2013), `http://www.guidancesoftware.com/encase-forensic`

[5] Salvaneschi, U., Balducci, P., Jacomuzzi, F., Moroncelli, P., Piazzalunga, C.: Security strength measurement for dongle-protected software. IEEE Security and Privacy 5(6), 32–40 (2007)

[6] Hirofuchi, T., Kawai, E., Fujikawa, K., Sunahara, H.: USB/IP a Peripheral Bus Extension for Device Sharing over IP Network. In: FREENIX Track: 2005 USENIX Annual Technical Conference, Anaheim, California, USA, pp. 47–60 (2005)

[7] Kwon, W., Cho, H.W., Song, Y.H.: Design and Implementation of Peripheral Sharing Mechanism on Pervasive Computing with Heterogeneous Environment. In: Obermaisser, R., Nah, Y., Puschner, P., Rammig, F.J. (eds.) SEUS 2007. LNCS, vol. 4761, pp. 537–546. Springer, Heidelberg (2007)

[8] USB/IP Project (accessed November 2013), `http://usbip.sourceforge.net/`

[9] Kyuchang, K., Dongoh, K., Kiryong, H., Jeunwoo, L.: Android phone as wireless USB storage device through USB/IP connection. In: 2011 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, pp. 289–290 (2011)

[10] Lee, C.Y., Ing, T.L.: USB Device Sharing Server for Office Environment. In: Asia-Pacific Services Computing Conference, APSCC 2008, pp. 953–956. IEEE, Yilan (2008)

[11] Incentives Pro (November 2013), `http://www.incentivespro.com/usb-redirector.html`

[12] FabuaTech (November 2013),
http://www.usb-over-network.com/usb-over-network.html
[13] Axelson, J.: USB Complete: The Developer's Guide, 4th edn. Lakeview Research LLC, United States of America (2009)
[14] Troelsen, A.: COM and.NET Interoperability. Apress Media, LLC (2002)
[15] Wibu-Systems (November 2013),
http://www.wibu.com/hardware-copy-protection.html
[16] SafeNet (November 2013), http://www.safenet-inc.com

Bandwidth Optimization for Real-Time Online and Mobile Applications

Ahmed Mokhtar A. Mansour[1] and Mona A.M. Fouad[2]

[1] Free Lancing, Nile Innovations,
Cairo, Egypt
ahmedmokhtar_adu@yahoo.com
[2] National Telecommunication Institute,
Cairo, Egypt
mfouad@nti.sci.eg

Abstract. This work proposes a novel algorithm to maximize the utilization of bandwidth for any text based application. The algorithm is used for real-time online text chatting, as well as offline SMS in mobile phones. Text is handled transparently without any modification in the core network. The proposed algorithm is an application of 'A-M' compression framework. Applying such algorithm will save at least 25% and up to 90% of the bandwidth depending on the context of the transferred data. Other applications such as email, web browsing, etc… will also gain from applying such algorithm. This paper demonstrated two types of applications; the real-time text chatting, and the mobile SMS application. The algorithm depends on a pre-defined dictionary of maximum size of 16 Kbyte installed on the client side or on the cloud near each client.

Keywords: Bandwidth optimization, mobile application, real-time application, network QoS, transparent compression.

1 Introduction

The long distance cable capacity around the world is much less in capacity than amount of data transferred. Bandwidth limitation is a source of bottleneck and delay over the network. As a consequence, QoS degradation due to delay and latency is a serious issue. Really thanks to content delivery networks (CDN) and peer to peer technologies (P2P), the end-to-end or the client server model that has optimized to avoid passing the whole network. Network and application level solutions are introduced to solve the problem of the delay of un-cached data.

In the early internet, the client server model was used. In such a model, the bandwidth was consumed very fast. The local caching at the client side was introduced as a partial solution [1]. Afterwards, proxy caches were used in the proxy level and showed a success [2]. Later, cooperative caches have been introduced such as Internet caching protocol and Cache Array routing [3]. The mutual synergy was introduced as a solution as well [4]. CDN technology is evolved to be one of the major border side

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_27, © Springer International Publishing Switzerland 2014

solutions to the problem of bandwidth consumption. All the previous solutions are application level solutions. The idea behind all this solutions is to transfer data once and retrieve it several times. Such a technique requires that the data is static and pre-retrieved. For real-time data such as video streaming, compression and scalable video coding, application level solutions are used. Some other network based solution such as integrated services (intserv), differentiated services (diffserv) and resource reservation protocol (rsvp) used either bandwidth reservation or priority classes' creation for the data [5] [6].

Real-time applications require transferring huge amount of data, instantaneously, for both images and videos. Hence, lossy coding techniques are used. This means that part of the data might be lost during data transfer. The data of these images and videos won't be affected much or even the lost data be noticed at all. For text data, the case is different; losing one single character might affect the meaning for the phrase or even the whole paragraph [7].

Currently, chatting application is the most used real-time text based application. Another text application that is not a real-time however faces a limitation in the SMS. The SMS has a limited number of characters (160 characters). Any extra character should be sent in another email. Even for the applications that can cut and regroup SMSs, the client may send or receive long (multiple) SMSs as one SMS; However, the network will deal with them as multiple SMSs without any bandwidth optimization [8].

Many other applications such as web browsing use real-time dynamic application for data exchange. In this case, the network based solution is used. The bandwidth optimization happens due to best utilization of resources rather than saving resources as for the mentioned application level based solutions (CDN, caching, P2P, etc…).[9] [10].

Other application level bandwidth optimization is data compression. An example is the email attachment compression done occasionally by mail servers. As most of the application level solution, the solution is controlled either by the proxy (service provider) as in caching or via the server (content provider) as for email attachment compression [11].

Saving bandwidth is a goal for both network operators and applications generator. Maintaining the same QoS (or better) with less bandwidth consumption is always an objective. When the application is a real-time one, the need to maintain QoS is a real challenge [12] [13].

In this paper we are offering a new transparent solution for bandwidth optimization for real-time online application. The solution is not limited to real-time applications. The solution is novel and unique for text based real-time application such as chatting.

The solution is based on a new introduced framework for text compression based on predefined dictionary. The proposed solution will guarantee at least 25% of compression. The solution is an end-to-end bandwidth optimization technique based. It guarantees both transparency and loss free transmission without any needs for modifications for the underlying transfer protocols or the network infrastructure; neither for service provider nor mobile operators.

2 'A-M' Framework

The 'A-M' code is a lossless adaptive compression technique, composes of two levels [14]. The first level ensures 25% reduction of data size of the original text, free of any overhead. The second level uses a variable length coding algorithm such as Huffman or Shannon, for further data reduction. In the first level, a fixed dictionary is generated for a given language. Two 'A-M' dictionary versions are proposed, storing the ASCII of individual letters and all possible combinations of two alphabetical letters [15].

The 'A-M' dictionary – vesion-1 generates adopted codeword of variable word length, with a maximum size of 12-bits. The Most Significant Bit (MSB) is used as indicator of the codeword length. For a given language, the dictionary is composed of all the normal ASCII letters from 0 – 127 plus all the combinations of two alphabetical letters. Three codewords of different lengths are used; 8-bits, 10-bits, and 12-bits for coding 'special element', 'simple elements', and 'compounds'; respectively. Some important letters such as space, full stop, comma, etc. are added to the dictionary. 57 compound elements resulted in 3249 records plus 15 records for special elements (numbers and special characters) and 128 ASCII character records for the simple elements. The parser can identify the length of the record by reading the first three LSBs for the 'simple elements' and the first four LSBs for 'special elements'.

The 'A-M' dictionary – version-2 is proposed for optimizing the bit-level decompression in version-1 and increasing the space of the encoded records. The 'A-M' dictionary version-2 uses fixed length codeword of 12 bits, offering a direct compression ratio of 25%. The ratio resulted from using 12 bits rather than 16 bits for two ASCII characters. The processing and the coding are different from version-1. The records in this version are divided into two parts. Each part is composed of 6 bits in the form of 'x.y', in which the two symbols 'x' and 'y' may vary from 0 to 63. The records starting from 0.0 and terminates at 63.63 (i.e., from 000000.000000 to 111111.111111). Five classes of records are used, namely; 'reserved', 'normal', 'alphabetical', 'numerical', and 'frequent' words records.

The dictionary length for each language will be around 16kbyte. It is fixed for each language and for each version. Such a dictionary might be saved for reuse for any "A-M" transformation, or might be built in real-time, using the previous algorithm.

2.1 Code Generation

In order to create a code from a given text document, the dictionary should exist either in the local storage or be retrieved from the cyber space; the right dictionary version should be retrieved. Before starting the first level 'A-M' coding, an optional flagging process should take place. The objective of the flagging process is to identify the frequent words that already stored in the list of frequent words in the dictionary, those words to be treated as whole words rather than character by character. The coding process is then goes through the text document and replaces each character, characters, or words by the corresponding code that predefined in the dictionary.

The 'A-M' compression based on 'A-M' dictionary version-2 may obtain between 25% and 75% based on the types of words in the document. The output compressed code from the fist level is compressed further via a variable length coding such as

Huffman or Shannon-Fano. If a second level coding is used, the 12 bit records output of the first level will be used as the input to the second level coding system such as Huffman or Shannon-Fano. The output of the second level will be a coded file with a mapping table that maps the input symbols to the corresponding code.

The decompression process is much easier than the compression one. It is a direct reverse mapping process from the 12 bit codes in the compressed document into the original character, two characters, or word. In all cases, it will be a direct mapping between 12 bit record and its registered correspondence in the predefined dictionary.

3 Bandwidth Optimization

Before going further explaining bandwidth optimization for real-time and mobile text-based application, we illustrate first an example to show the advantages of the 'A-M' code over the well known Lempel–Ziv–Welch (LZW) coding algorithm [16].

LZW is one of the most used codes for text compression, providing a high compression ratio. LZW builds a mapping table based on the document contents. Figure 1 shows an example of transferring text to LZW code. The left hand side is the dynamic mapping table. The right hand table shows the details of the LZW step by step application for the input string.

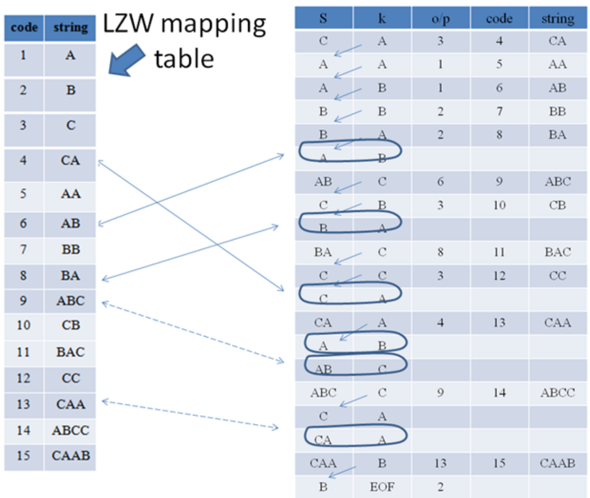


Fig. 1. LZW example

In the shown figure, the input text to the LZW coder is the 20 characters string CAABBABCBACCAABCCAAB. LZW creates, dynamically, a mapping table between character(s) and LZW code. The LZW code is 12 bit code created dynamically from the input string. The table maximum size is 4096 records. The first part of LZW is to define the basic mapping table. The basic mapping table is a table that contains the ASCII characters included in the document with their corresponding

LZW code. Each new entry will perform two actions. The first action is to encode the input to LZW code; the second step is to create new entry to the mapping table. Such event might happen for single or multiple characters depending on the existence on the LZW table. The output of this example will be 3 1 1 2 2 6 3 8 3 4 9 13 2. In this special case, there will be almost no compression (20*8/13*12=1.025). This is one of the disadvantages of LZW.

Now, we will see how the 'A-M' code reacts with the same example. The intermediate code for the text 'CAABBABCBACCAABCCAAB' accomplishes data reduction of about 30% of the original size (compression ratio = 20*8/10*12=1.3333).

It is important to say that both LZW and the 'A-M' code perform further compression by applying a variable length code. From the previous example, it is seen that in small text files the compression ratio of LZW is very low compared to the 'A-M' code. Another disadvantage of LZW is that it cannot be used for real-time application because of the dynamic dictionary creation process at the sender and the receiver sides.

3.1 Bandwidth Optimization for Real-Time Online Applications

In case of real-time applications, such as chatting, our solution is to use the 'A-M' code in both peer sides. The dictionary will be installed locally and be opened in the cache memory (16 Kbyte) of both sides. The dictionary will be used with each new line feed to compress the data from one side and decompress the data on the other side. The version-2 dictionary will be used with the same algorithm used for text documents. The utilization is transparent from both peers and from the network. When the data is transferred from one side to another, the receiver side should use the same version of dictionary used in the sender side to retrieve the correct data. The same technique might be used for both static and dynamic web pages.

To better understand the operation, figure 2 shows a chatting example between two users, assuming that user one is sending data to user two. User one is sending the message: "this is user one message". In a normal communication case, the number of transferred characters including the spaces is 24 characters. The size of 24 characters will be 24*8bit = 192 bits. Using the simplified version of version-2 'A-M' code, ignoring the frequent words, the immediate gain of the compression will be 25%. This will result from sending 12 messages, each message of 12bits. The result will be 144bits; which means an immediate gain of 25% with a very light processing operation.

Now we will take one step more to use the standard 'A-M' version-2 dictionary, where the frequent words list is used. The second user message might be: "and this is the second user message". Without any coding, the message is 35 characters that need 280 bits to be transferred. Using the light version as the first example 12bits * 18 records = 216 bits is needed. This means 23% bandwidth saving. In case of using the standard 'A-M' framework, we will need more processing. However, the number of transferred records will be only 14 records (assuming the words 'and', 'this', 'the', 'message' are in the frequent words' list). The total number of bits needed will be 14 record * 12 bits = 168 bits. The bandwidth saving in this case is 40%. In most of the chatting application, statically, 50% of used words, at least, will be from the frequent list. Concerning the frequent words' list, the most used 100 words in the English language is contained in the dictionary. According to Oxford dictionaries statistics, these top used 100 words account for 50% of the whole words listed in the Oxford English Corpus (of over billion words).

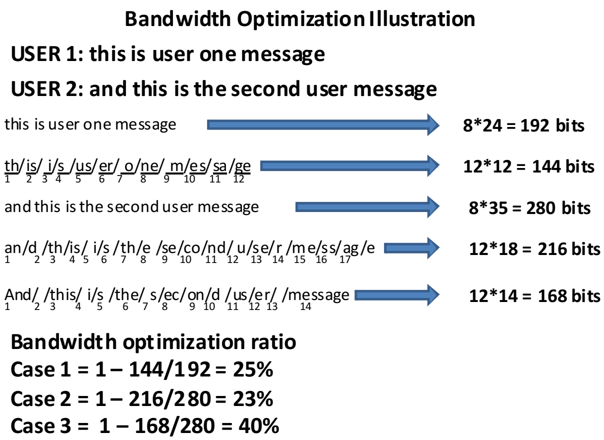


Fig. 2. Chatting application example

The minimum optimization that we will get from using the 'A-M' code will not be less than 25% in most of the cases. The only case that the compression is not useful is when only one character is transmitted. Applications should take care to count the characters before starting the operation. If the only one single character to be transmitted, the compression process should be suspended. If the number of transferred characters is even, the ratio of bandwidth saving won't be less than 25%.

Based on Internet usage statistics, the plain text consumes 2% of the internet overall bandwidth. The HTTP applications (almost web browsing) consumes around 16% of the internet bandwidth [17]. If the 'A-M' code is used for the mentioned applications, between 4% -10% of the overall bandwidth used over internet will be saved. Taking into consideration that the local resources (memory, processors) are normally free during users' online activities, using a small portion of these resources for sack of bandwidth optimization is accepted as a choice. The short size of the dictionary allows the use of cache memory for the search/conversion process. The measure of user's convenience should take place after large provisioning of the protocol.

3.2 Bandwidth Optimization for Mobile Applications

The SMS is one of the most used services of mobile system and it uses huge bandwidth. The use of 'A-M' framework as mentioned in the chatting or in the text documents will guarantee 25% reduction of the messages at least. This affords a huge save in power, bandwidth, and overall system utilization.

According to a new study from the International Telecommunications Union (ITU), more than 6.1 trillion text messages are transferred in 2010. In other words, close to 200,000 text messages are sent every second [18]. The business was around 115 billion USD in 2010. The maximum size of messages is 140 bytes. If the average is 100 bytes per message, the total size is transferred over the network in 2010 is around 610 Terra Byte. Using ‘A-M’ code would save 150 terra bytes of data transferred over the signaling channel of mobile networks. Signaling channels resources are always limited as resource. The gradual increase of using of SMS consumes that resources and causes delay of delivery. ‘A-M’ code can be seen as a possible future standard for the SMS systems. The easy processing in both sides and the guarantee of backward compatibility allows the ‘A-M’ code to be used. For online services, some modifications can be done to dictionary version-2 in for fast processing.

The advantage of using ‘A-M’ code is that the coding can be done in real-time in parallel with writing the message. Besides, there is a minimum rate of compression guaranteed. A reduced version of dictionary version-2 can be used. The reduced version is under construction and it provides 11 bits record rather than 12 bits. This results in direct gain of 32% of SMS size. This leads to a minimum of 202 words as limit of SMS size. An application such as twitter can use the solution in both broadband and mobile application.

The characteristics of the dictionary used in SMS application should take into consideration that not all special characters are needed in SMS. On the other hand, not all ASCII codes are allowed to be transferred over the signaling channel. Control codes (from 0 to 31) are not allowed in some of the networks for security reasons and reserved over other networks for control use.

4 Conclusion

This paper proposes two major contributions in two types of networks internet and mobile 2.5 G networks and beyond. The customization of the ‘A-M’ framework to be used for both real-time and online applications will result in a huge optimization on bandwidth over the whole network in the order of 25% and above for all text contents. The efficiency and the simplicity of the algorithm encourages using it for bandwidth optimization for real-time and online applications. All text based services of any mobile system may also benefit from the framework to reduce the size of transferred data by at least 25%. The use of the application guarantees between 25% and 80% reduction of the consumed bandwidth.

The chatting application is not the only online application that can benefit from the ‘A-M’ code. Web browsing, emails, etc. may use the ‘A-M’ framework to reduce used bandwidth consumption. Other applications such as map-reduce or Hadoop can also benefit from the ‘A-M’ code version-2 and the reduced version.

Future work is to customize the ‘A-M’ code for more services to reduce the consumption of network bandwidth. Implementation should take place for different types of devices and mobile operating systems.

References

1. Abrams, M., Standridge, C., Abdulla, G., Williams, S., Fox, E.: Caching proxies: Limitations and potentials. In: Proc. 4th International World Wide Web Conference (1995)
2. Shim, J., Scheuermann, P., Vingralek, R.: Proxy Cache Algorithms: Design, Implementation, and Performance. IEEE Transactions on Knowledge and Data Engineering 11(4), 549–562 (1999)
3. Valloppillil, V., Ross, K.W.: Cache Array Routing. Protocol v1.0. Internet Draft (1998)
4. Mokhtar, A., Rubino, G.J.: Mutual synergy between web caches. In: Hui, L.C.K., Lee, D.-L. (eds.) ICSC 1999. LNCS, vol. 1749, pp. 439–446. Springer, Heidelberg (1999)
5. Shioda, S., Mase, K.: Performance comparison between IntServ-based and DiffServ-based networks. In: Global Telecommunications Conference, GLOBECOM 2005. IEEE (2005)
6. Lixia, Z., Xerox, P., Deering, S., Estrin, D., Shenker, S.: RSVP: a new resource ReSerVation Protocol. IEEE Communications Society, Network 7(5) (1993)
7. Jacob, N., Somvanshi, P., Tornekar, R.: Comparative Analysis of Lossless Text Compression Techniques. International Journal of Computer Applications 56(3), 0975–8887 (2012)
8. Chandran, L., Agarwal, N., Apte, V.: Capacity Analysis of the GSM Short Message Service. In: Proceedings of NCC 2003 (2003)
9. Mansour, A.M.A., Cataline, M.: dynamic content allocation/delivery mechanism. patent WO/2003/065659 (2003)
10. Buyya, R., et al. (eds.): Content Delivery Networks. LNEE, vol. 9. Springer, Heidelberg (2008)
11. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 2nd edn., pp. 385–392. MIT Press and McGraw-Hill (2001); Section 16.3, ISBN 0-262-03293-7
12. El-Gendy, M.A., Bose, A., Shin, K.G.: Evolution of the Internet QoS and Support for Soft Real-Time Applications. Proceedings of the IEEE 91(7) (2003)
13. Qiang, N., et al.: A Survey of QoS Enhancements for IEEE 802.11 Wireless LAN. Journal of Wireless Communications and Mobile Computing 4(5), 547–566 (2004)
14. Mansour, A.M.A., Fouad, M.A.M.: A New Compression Framework: 'A-M' Code. In: Internathional Conference on Modeling, Simulation, and Applied Optimization, ICMSAO 2013, Tunisia. IEExplore (2013) ISBN: 978-1-4673-5812-5
15. Mansour, A.M., Fouad, M.: Dictionary Based Optimization for Adaptive Compression Techniques. In: MIPRO 2013 - 36th International Convention CTI - Telecommunications & Information (2013)
16. Li, Z.-N., Drew, M.S.: Fundamentals of multimedia. Pearson Education (2004) ISBN 0-13-127256-X
17. Envisional: Internet bandwidth usage estimation Technical Report: "An Estimate of Infringing Use of the Internet," Envisional Ltd. (2011), `http://www.envisional.com`
18. `http://www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf`

An Evaluation on the Usability of E-Commerce Website Using Think Aloud Method

Rogayah Abdul Majid, Mardziah Hashim, and Nurul A'syida Abdul Jaabar

Faculty of Computing and Mathematical Sciences, Universiti Teknologi MARA,
40450 Shah Aalam, Selangor, Malaysia
rogayah@tmsk.uit.edu.my

Abstract. The introduction of e-commerce websites have enabled users to shop online. To ensure positive shopping experiences by the users in a highly competitive environment, it is important for e-commerce website designers to design a good and usable site. This study attempts to evaluate the usability of the thepoplook.com website based on the model proposed by Venkatesh & Argawal. However, there is a missing element in the model - the security element. Therefore, this study has identified the problems encountered by users during the process of online shopping offered by thepoplook.com website. In addition to that, this study has also identified elements for evaluation of website usability and determined the importance of security element in evaluating e-commerce website. Think-aloud method was used to evaluate the thepoplook.com website where the respondents are required to answer questions regarding the usability of the website. The findings of the study showed that five attributes have been identified as usability elements used to evaluate the e-commerce website, namely: Ease of Use, Made-to-Medium, Emotion, Content and Security. The findings confirmed that security is the one of the important elements that the users consider in using e-commerce website. Web site designer must consider security elements as a crucial part in developing an e-commerce website, since it will give users confidence to buy from the website.

Keywords: Human Computer Interaction, Usability, Usability Evaluation, Think Aloud, B2C E-Commerce.

1 Introduction

The introduction of e-commerce website has encouraged the users to shop online. Online shopping allows the users to buy their favorite items all over the world through the internet. This shows that the Internet is an effective medium for commercial drives [1]. Nielsen's PayPal's research study stated that the Malaysian online shopping market has reached the tremendous number of RM1.8 billion in 2010 and has estimated that this will continue to grow up to RM5 billion in 2014 [2]. PayPal's Managing Director and General Manager for South East Asia and India, Elias Ghanem said that this shows that Malaysian are finding great value in buying product and services via the website as part of their daily life.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_28, © Springer International Publishing Switzerland 2014

The increase in the number of e-Commerce websites has made the evaluation of usability website know-hows relevant. The important part of online shopping experience is when users are able to buy the products and services easily. Research by Ana and Antonio [3] recommends that interface features, helps to bring positive impact on the company's success in the Internet marketing through user's experience. One thing closely associated with such experience is to make the "user interface" simplified and easy to use. On the other hand, lack of security will discourage the users to perform online payment while they are shopping online.

Izabela (2011) in her research study mentioned that Nielsen cited that one of the most important key elements for the success of an e-commerce website is usability. It is proven that the failure of many e-commerce sites in attracting the customers is caused by lack of usability element. Therefore, in order to ensure positive shopping experiences by the customers, it is important to develop an e-Commerce website that has good interface design. The Thepoplook.com website has been chosen in this study because this website is one of the very popular shopping websites which it was at the rank of 533 in Malaysia [22].

The aim of this study is to examine the problem regarding the usability encountered by customers during the process of online shopping offered by thepoplook.com website. Besides, this study attempts to evaluate the usability of the thepoplook.com website based on model propose by Venkatesh & Argawal [14]. Since the security element was not considered in the model, therefore, this study will identify elements for usability website evaluation and determine the importance of the security element in evaluating e-commerce website.

2 Literature Review

2.1 Thepoplook.com Web Site

Thepoplook.com is a B2C e-Commerce website which the seller is a business organization while the buyer is a consumer [7]. This type imitates physical retailing and therefore it is commonly called electronic retailing. B2C is the second largest and the earliest form of e-commerce [8]. Since 1995, there has been a rapid expansion of B2C e-Commerce activities and purchasing via the Internet is one of the most rapidly growing forms of shopping [9]. Study done by [21] on Malaysia's Online Shopping Scene in 2012 [18], majority of Malaysian online shoppers buy clothing and accessories.

Thepoplook.com website originates from Malaysia and the warehouse and sales office is located in Malacca. This website sells apparel and accessories through online. This online store offers free shipping via express courier to its customers within the country. From the interview with the CEO and Founder of thepoplook.com, Maryanne, 90% of the customers are Malaysian and 10% are international. According to Alexa [22], the thepoplook.com is currently ranks 533 in Malaysia.

2.2 Usability Elements

This study uses the website usability model proposed by Venkatesh and Agarwal [14] to measure website usability. Venkatesh & Argawal researched demarcated usability by applying the ISO definition on usability which is "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". The model has proposed four usability elements namely: ease of use, made-to-medium, emotion and content as shown in Figure 1 below.

Usability Elements	Sub-catogories
Ease of Use	View, Select, Add to Cart, Remove to Cart
Made-to-Medium	Product Categorization
	Accessibility and Customer service
Emotion	User's Feeling
Content	Information Layout

Fig. 1. Usability Elements proposed by Venkatesh & Argawal (2006)

Ease of use describes the cognitive effort in using a website. Made-for-the-medium relates to designing a website to match with the user's need, emotion is about users' feeling in using the website and content is describing about information layout.

2.3 Usability Evaluation

According to Abran et al. [11] usability for an application (software or system) is determined by the productivity and the acceptance of the application by its users. Nielsen [12] in his "Usability Engineering" book mentioned that usability is a single dimensional things of the user interface. In fact, usability has several components and is conventionally measured in contradiction of these five attributes, which are: learnability, efficiency, memorability, errors and satisfaction [12]. Usability is part of the significant attributes of any user interface and it quantifies user's perception of how easy the website or system interface is [13]. Recently, usability become the center of discussion [14] and defined by various definitions among the academia and industry [15]. According to Nielsen, the key elements for the success of an e-commerce website is usability. Izabela [9] in her study has proven that the failure of many e-commerce sites in attracting the customers is caused by lack of usability element. Tzanidon [16] supported that when the customers are enabled to find the item which they are looking for in e-Commerce website, they will not shop online.

As a part of the Usability Engineering process, usability evaluation is a performance measurement to ensure if usability objectives are attained. Usability evaluation is the methodical exploration of the workability of a system or software. According to

Schweibenz and Thissen [17] usability evaluation is used as a decision guideline or as an assessment of adoptions as it tests the object, provide an evaluation about it and deliver possibilities for enhancement. They also mentioned that, the objectives of the usability evaluation are to improve the existing products, the interface design and to overcome the usability issues.

A usability evaluation method consists of two which are export oriented and user oriented. Expert oriented method allows the expert to use their knowledge and experience to simulate on how users would use and handle a certain system. However, user oriented methods allow the users to evaluate the system themselves. Think-aloud, field studies, focus groups and interview are the example of this method. User oriented testing is good for identifying specific major problems related to website navigation, interface design, viewing and purchasing process, customer service and many more. This was supported by Nielsen [12] by stating that user testing provides direct information about how real users interact with an interface and the exact problems they face during their interaction. Even though user oriented method is time consuming and resource intensive, but it helps to deliver tangible data that reflect real opinions of the user. It also allows the user to express the problem that they are facing throughout the evaluation process and followed by recommending a solution or ideas for improvement.

3 Methodology

3.1 Method

Think aloud method has been chosen to be used in this study. Think aloud protocol or method is a tool that comprises a specific number of users who interact with the system individually. Think-aloud encourage the participants to provide verbal descriptions of what they are intending to do and what is happening on the screen is the main aspect of this method [10]. According to Nielsen's study, think-aloud usability testing is one of the most fundamental evaluation methods. With think-aloud, users are being observed where they have to verbalize their thoughts and voice out their comments while perform the given tasks. Furthermore, according to Lazar [23] and Nielsen [12], think-aloud is a user testing method with a condition of asking users to think-aloud throughout their interaction with an interface. Nielsen indicated that having users verbalizing their thoughts using this method allows an understanding of how users view or interpret an interface. It also helps to facilitate the identification process of the major misconceptions from users.

In this study, 15 users have been selected to evaluate the thepoplook.com website. The chosen users must have a strong internet experience background, with at least one year of experience in using websites. 15 users have participated in the website evaluation based on users' profiles. For this study, the selection of users' profile is important because the information about the target users of the websites would be helpful to identify typical users for user testing [16]. The selection of the respondents is based on the following users' profiles: (a) Gender - Male and Female participants were recruited for the evaluation activity. (b) Internet experience - To identify the frequency of internet use. (c) Online shopping experience - Participants who use online shopping are usually familiar with the use of the Internet.

The researcher has developed the task scenario for the studied website which is thepoplook.com. This included the typical tasks for the selected B2C e-Commerce website which represented the actual use of the website. In preparing the task scenario, the recommendation of Nielsen [19] and Preece et al. [18] were taken into consideration concerning the beginning and ending of the tasks. Easy tasks were elected for the first tasks with the intention to allow participants feel comfortable on the commencement stage. This is also to let the participants feel good when they reached something at the end of the tasks. The task scenario involved some typical tasks suggested for e-Commerce website from previous studies, such as finding information and finding products in the e-Commerce website [16] and using the search field to find for product [16] and changing the content of the shopping cart.

The researcher has developed the task scenario for the studied website which is thepoplook.com. This included the typical tasks for the selected B2C e-Commerce website which represented the actual use of the website. In preparing the task scenario, the recommendation of Nielsen [19] and Preece et al. [18] were taken into consideration concerning the beginning and ending of the tasks. Easy tasks were elected for the first tasks with the intention to allow participants feel comfortable on the commencement stage. According to Nielsen [19] and Preece et.al [18], this is also to let the participants feel good when they reached something at the end of the tasks. The task scenario involved some typical tasks suggested for e-Commerce website from previous studies, such as finding information and finding products in the e-Commerce website [16] and using the search field to find for product [16] and changing the content of the shopping cart.

During the usability evaluation, each participant performed the tasks given to them on the target system thepoplook.com. The participants were told to be verbalized, throw out their thought and help each other while performing the usability testing. Such clues and steps are logged by the observer on a piece of paper, recorded by Camtasia 8.0 software and camera. Through these steps, information was collected. The participants are allowed to talk freely about any error they face or the positive point about thepoplook.com while performing the given tasks. The usability test of web site interface focuses more on the collection of qualitative information such as users' thoughts, emotions and reactions. Therefore the adoption of thinking-aloud method allows directly and efficiently gather the problem arising from interface design.

3.2 Research Framework

This study adapted website usability model proposed by Venkatesh & Argawal [14] to measure the usability of the thepoplook.com website. They have proposed four elements of usability evaluation namely ease of use describes the cognitive effort in using a website, made-for-the-medium relates to designing a website to match with the user's need, emotion is user's affective reactions provided by a website, and content is described about content or information layout. However, Agarwal and

Venkatesh model do not emphasize on the security perspectives. Therefore, this study will highlight the security element in website usability evaluation. The Figure 2 below shows the proposed new website usability model.

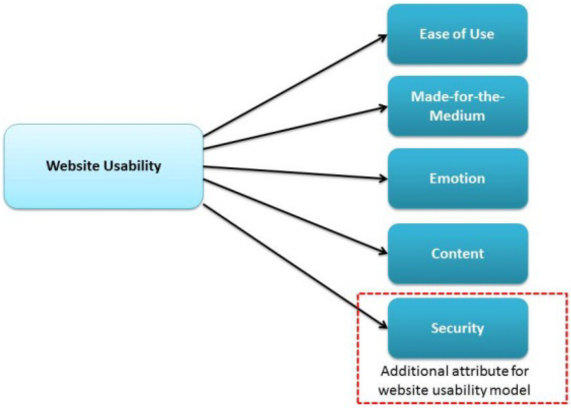


Fig. 2. Enhanced Venkatesh & Argawal (2006) website usability model

These five elements which have been sub-categorized into seven types, namely (1) View, Select, Add to Cart, Remove to Cart (2) Product Categorization (3) Accessibility and Customer Service (4) User's Feeling (5) Information layout and (6) Security and Privacy as shown in Table 1.

4 Result and Analysis

The results of the study are shown in the Table 1. From the table it shows that both Content and Made-to-Medium shows higher rate of problems which are 31% each, followed by Ease of Use 22%. Even though emotion in using the website is only 10%, and security of the website is six percent (6%), it shows that users are still concerned about these two elements.

i. Made-to-Medium

Made-to-Medium is defined as tailoring a website to apt a particular user's need. From the usability testing, it shows that Made-to-Medium is the most critical usability problem with the percentage of 31%.

ii. Content

The evaluation attributes of Content covers Information layout which contributed to 31% of the identified usability problem. It shows the arrangement of the information can delay the assessment of the data and will give negative influence to users.

Table 1. Users' Perspectives On The Usability Of Thepoplook.Com Website

Usability Elements	Problem Type	No.	(%)
Made-to-Medium	Product Categorization Accessibility and Customer service	15	31%
Content	Information Layout	15	31%
Ease of Use	View, Select, Add to Cart, Remove to Cart	11	22%
Emotion	User's Feeling	5	10%
Security	Security and Privacy	3	6%

iii. Ease of Use

Ease of Use is defined as the cognitive effort required in using a website. The evaluation attributes of Ease of Use includes View, Select, Add to Cart and Remove to Cart. Ease of Use contributed to 22% with 11 problems identified throughout the usability testing. Ease of Use of e-Commerce website emphasized the overall online shopping process which starts from product research, selection or product, placing orders and completion of online transaction which would run as smoothly and conveniently by the customer.

iv. Emotion

Emotion is defined as an affective reaction invoked by the website or user's feeling. This attribute contributes 10% with five identified problems. Users may complain, get tired and loss of patience while using unfriendly system.

v. Security

Even though security contributes the least usability problem which is only 6% of the total problems, security or safety attributes are important for an e-Commerce website. This is due to avoid unauthorized access, whether accidental or deliberate to the users' personal data and payment transaction.

5 Conclusions

The growth of e-Commerce websites in Malaysia market is currently at a fast pace, apparently creating competitive pressure from another e-Commerce websites. Therefore five usability elements have been identified for the development of usable e-commerce website. In order to evaluate the Thepoplook.com website, this study use website usability model which consists of five attributes including Ease of Use, Made-for-the-Medium, Emotion, Content and Security. From the study, it shows that security element has been identified as an important element in evaluating usability of the website. The findings of the study will become a guideline for other e-commerce website developer to develop an e-commerce website that is useful and would satisfy the users. The good user interface and good usability as well as good security of an e-commerce website will encourage users to shop online.

References

1. Candace, D.P.: E-Commerce and M-Commerce Technologies. IRM Press, Publisher of Innovative Scholarly and professional information technology titles in cyberage (2005)
2. Liveatpc. PayPayl study breaks down Malaysians' online spending habits (2011), http://www.liveatpc.com/paypal-study-breaks-down-malaysians%E2%80%99-online-spending-habits
3. Ana, R.A., Antonio, P.M.: Exploring Trust Building Mechanisms in Global B2B Electronic Markets (2006)
4. Zainudin, N., Ahmad, W.F.W., Nee, G.K.: Evaluating C2C e-commerce Website usability in Malaysia from users' perspective: A case study (2010)
5. Turban, E., King, D., Lang, J.: Introduction to Electronic Commerce, 2nd edn. Prentice Hall, New Jersey (2009)
6. Hu, H.J., Yen, H.Y.: Study on the Interface Usability of B2C Hypermarket E-commerce Website (2008)
7. Layla, H.: Usability Evaluation Framework for E-commerce Websites in Developing Countries (2009)
8. Andam, Z.R.: E-commerce and e-business. UNDP (2003), http://www.apdip.net/publications/iespprimers/eprimer-ecom.pdf
9. Izabela, M.M.: Dual-Method Usability Evaluation Of E-Commerce Websites. In: Quest of Better User Experience. University of South Africa (2011)
10. Rubin, J.: Log analysis pays off. Network Computing 15(18), 76–78 (2004)
11. Abran, A., Khelifi, A., Suryn, W., Seffah, A.: Usability Meanings and nterpretations in ISO Standards. Software Quality Journal 11(4), 325–338 (2003)
12. Nielsen, J.: Heuristic evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods, pp. 25–64. John Wiley & Sons, New York (1994)
13. Nielsen. J.: Usability 101: Introduction to usability. Useit.com (2003)
14. Venkatesh, V., Agarwal, R.: Turning Visitors Into Consumers: A Usability-Centric Perspective Purchase Behavior in Electronic Channels. Management Science 52(3), 367–382 (2006)
15. Cousaris, C.K., Kim, D.J.: A Meta-Analytical Review of Empirical Mobile Usability Studies. Journal of Usability Studies 6(3), 117–171 (2011)
16. Ekaterini, T.: Evaluating usability of e-commerce sites by tracking eye movements. PhD thesis, The Open University (2006)
17. Schweibenz, W., Thissen, F.: Quality of Website: Usability Evaluation. Springer, Heidelberg (2003)
18. Preece, J.: Human-computer interaction. Addison Wesley (1994)
19. Nielsen, J.: Usability Engineering. Academic Press, Inc., Boston (1993)
20. Chan, H., Lee, R., Dillon, T., Chang, E.: E-Commerce: fundamentals and applications. John Wiley & Sons, Ltd., England (2001)
21. Open Minds Resources, Malaysia's Online Shopping Scene (2013), http://openmindsresources.com/malaysias-online-shopping-scene-2012
22. Alexa, http://www.alexa.com/siteinfo/thepoplook.com
23. Lazar, J.: Web usability: A user-centered design approach. Pearson Education, Boston (2006)

Towards a Structured Electronic Patient Record for Supporting Clinical Decision-Making

Bente Christensen and Gunnar Ellingsen

Telemedicine and e-Health Research Group, University of Tromsø, Norway
bente.christensen@telemed.no, gunnar.ellingsen@uit.no

Abstract. At present, Electronic Patient Record systems (EPRs) in Western hospitals are mainly stores of free-text patient information and lack utility for purposes other than accessing such information. The need to improve support for clinical work processes and decision-making has put pressure on vendors and decision makers to put forward a strategy for enhancing the structure of EPR content. This paper reports on the first attempt to implement a structured EPR using a two-level modeling approach in Norwegian hospitals. Taking a work practice perspective, we show that implementing this new EPR in patient care planning implicates several socio-technical challenges that need to be solved in the process.

Keywords: EPRs, work practice, clinical decision support.

1 Introduction

Over the last few decades, Western hospitals have increasingly implemented Electronic Patient Records systems (EPRs) systems to document the treatment and care of patients. As a result, the EPRs have grown to considerable sizes, due to the accumulation of large amounts of patient information. A key problem is that much of the information in EPRs is free text (of which physicians´ narrative notes are the most prominent example). This makes it hard to use EPRs for purposes other than registering and looking up patient information [1].

Increasingly, it has been pointed out that EPRs also should be capable of providing tailored Clinical Decision Support (CDS) [2] at crucial points in patients' treatment trajectories. This has put pressure on vendors and decision-makers to put forward a strategy that facilitates such a process by: a) structuring the content in EPRs, b) identifying crucial tasks in particular need for decision support and c) automatically triggering the next step in patients' treatment pathways. Simultaneously we recognize that achieving CDS [3–5], as well as fostering the inclusion of more structured information in EPRs [6, 7], has proven hard to achieve in practice. In this paper, we look into a promising new strategy from the international openEHR organization [8]. openEHR is a two-level modeling approach aimed at supporting a high degree of interoperability between different healthcare systems and the reuse of data [8]. The openEHR approach seems promising, as the use of archetypes allows user

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_29, © Springer International Publishing Switzerland 2014

communities to define structured data in a very dynamic way, thus potentially avoiding out-of-hand user rejection of the structuration efforts due to unsuitable degrees of standardization. However, despite being a technical innovation, the openEHR approach may still face many socio-technical challenges when the new technology is put into real use. We pursue this through the following research question: *What challenges are associated with the introduction of decision-supportive EPRs based on the openEHR approach, and how can these be dealt with?*

Empirically we report from a large-scale EPR project initiated in 2011 in the North Norwegian Regional Health Authority. A key aim of this project was to replace an existing, largely free-text-based EPR with a new archetype-based (i.e., highly structured) EPR offering extensive decision support. We focus on the first attempt to implement decision support in the surgery planning process.

This study is positioned within a constructive paradigm and hence makes use of interpretive methods [9, 10] where we particularly focus on the users' work practice. A variety of data sources are used to bring out different perspectives. The first author has participated in testing sessions and workshops with developers at software vendor dubbed BigVendor and health care personnel, conducting semi-structured interviews and document studies. In addition, informal talks with people involved in the development process clarified themes that are identified in this paper. Our analysis of the data has been guided by the philosophical perspective of hermeneutics, according to which a complex whole is understood "from preconceptions about the meanings of its parts and their interrelationships"[9].

2 Theory

The openEHR framework separates the technical designs of systems from clinical concerns. A standardized reference information model represents the first level, while the openEHR archetypes, based on the reference model, represent the second level [11]. The archetypes specify constraints on the data structures in the reference model. Hence, they serve as a constraint mechanism, ensuring that stored information is valid in terms of clinical knowledge. [1].

The outcomes of the openEHR approach are systems and tools for computing with health information at a semantic level, thus enabling analytical functions like CDS. Clinical decision support depends on good quality clinical data repository and hence reinforces the need for standardized data representation and storage. Lack of good clinical data warehouse will have significant impact on the quality of advices emanating from CDS systems. Data mining algorithms require good quality clinical data repositories to be able to extract knowledge to support clinical decision-making [5]. CDS systems also depend profoundly on large volumes of readily-accessible, existing clinical datasets usually extracted from the repository content of EPRs. Lack of standardized data in the repository may lead to datasets not representative of the patient population (ibid). It is therefore essential that standardized data representation are used for leveraging the knowledge base repositories to facilitate the generation of patient-specific care recommendations for physicians (ibid). Archetypes allow flexible terminology mapping. However, to enable interoperability, they require consensus on maximum definition. Being structured data elements, archetypes can also be displayed in different presentations. Together with the ability to match individual patients to a computerized knowledge base, in order to generate patient-specific assessments or recommendations, the reuse and easy access to data is considered decision-supportive abilities.

While acknowledging the great promise with the openEHR approach, new technical solutions to be used in complex health care settings may still face many socio-technical challenges [12]. For example, the temporal aspect of health care work makes it a process, not an event [13]. In our case, this means that making a plan for surgery, the planning process is not a result of a sequence of decisions, it is the emergent effect of the interlocking of entities performing subtasks [14]. In this respect collaborative activities including informal communication and coordination are of uttermost importance [15]. Munkvold et al. [16] found that pressing too hard on a practice to achieve more formalized routines, resulted in unexpected changes of the correlated informal practice. They demonstrated how informal routines, which initially were considered a problem, became part of the solution in order to make the new formalized routines work. They even found that the formalizing of routines nurtured the need for even more informality (ibid). Viewing clinical work as collaborative productions dependent upon and constitutive of teamwork has implications for the computing that supports the primary work [13, 17].

3 The BigInvestment Project

In 2011, the North Norwegian Regional Health Authority issued a call for tender, asking for new clinical ICT systems for all 11 hospitals in North Norway. The overall goal of the NNHRAs new investment was development of an EPR supportive of physician decision-making, with the objective of facilitating execution of patient treatment pathways and clinical workflows within the region. The process of developing and implementing the new EPR system took the form of a project known as BigInvestment. The cost of BigInvestment amounted to 82 million EURO for the period 2012-2016, making it one of the most ambitious healthcare-related ICT projects in Norway at the time.

The leading Vendor of EPR systems in Norway, BigVendor, won the bid for tender. BigVendor s systems are used by 70% of hospitals in Norway, encompassing 65,000 users, including hospitals in the North Norwegian Regional Health Authority. In winning the bid for tender, BigVendor promised to develop what it denoted as the "next generation EPR", NewArena. The NewArena will replace the old EPR, which, to a large degree, contains unstructured text that offers little CDS. Together with BigVendor, the BigInvestment project has planned the progressive development and implementation of NewArena. While NewArena is being developed, a substantial amount of clinical work will have to be done on the old platform, but this is not supposed to affect users´ daily work. All information stored on the old platform must also be retrievable from NewArena after migration to this system and must be displayable as part of patients´ histories.

NewArena is built on a two-level modeling approach, using the openEHR architecture [11].

4 Assessing Patients for Surgery

Critically, NewArena provides clinicians with a totally new CDS tool that offers easy access and overview over patient treatment history and facilitates the documentation of patient treatment and care. The first functionality afforded by NewArena‘s product family benefited outpatient clinic physicians:

The ambitions for Arena are high; they must be when we are talking about an Electronic Patient Record for the future (…) The first version ready for customers in spring 2012 will allow physicians to work more efficiently in outpatient clinics. New functionalities will follow rapidly until the end of 2016 [18]

Particularly in surgery planning, the role of physicians at outpatient clinics assessing surgery patients is vastly important. Patient assessment is the first step in the costly and complicated surgery planning process, a process characterized by the prevalent waste of resources, presumably due to missing information. The collection and documentation of clinical data earlier in the planning process is anticipated to improve the situation. Today, clinical data are registered as patients are admitted for surgeries; this might be the day before or the day of surgery. Pushing information gathering forward to outpatient consultations, where surgery is decided, is expected to improve the quality of the whole planning process, including planning for bed-capacity, anesthetics, use of surgery theatres and tools.

NewArena shall also provide the ability to communicate plans and patient needs in a way that allows the system to catch up if anything is forgotten. Thus, for patients planned for surgery, for instance, a reminder will show up for the physician if treatment is not initiated according to the plan made by the physician in the outpatient clinic. Accessible and updated templates of patient pathways will also inform personnel involved in patient treatment about patient status and treatment plans and progress, thus making it easier to carry out treatment according to plans. Also in NewArena, clinical information, notes and other types of documents related to, e.g., medication, are expected be retrieved from patient records and displayed in surgery planning module, connected to “surgery cases”. Thus, the planning process will be facilitated by the reuse of structured, clinical data already registered in the EPR.

5 The First Pilot

After a one-year delay of the project plan, a version of NewArena was presented to clinicians for piloting and testing in October 2013. According to BigVendor, delivery was delayed because the surgery planning module in the new platform had to «wait» for basic functionality to be achieved on the Arena platform, like structuring of data, in order to work optimally.

The interaction between EPR and NewArena is complex, and BigVendor reports challenges related to interaction between the platforms and to the performance and implementation of new functionality in NewArena. (Dossiers to the steering committee meeting on BigInvestment, 10 October 2013)

The starting point for surgery planning is defined to be outpatient consultation. It is important that physicians be provided enough information in outpatient consultations to plan and document decisions regarding further treatment. patients enter consultations.

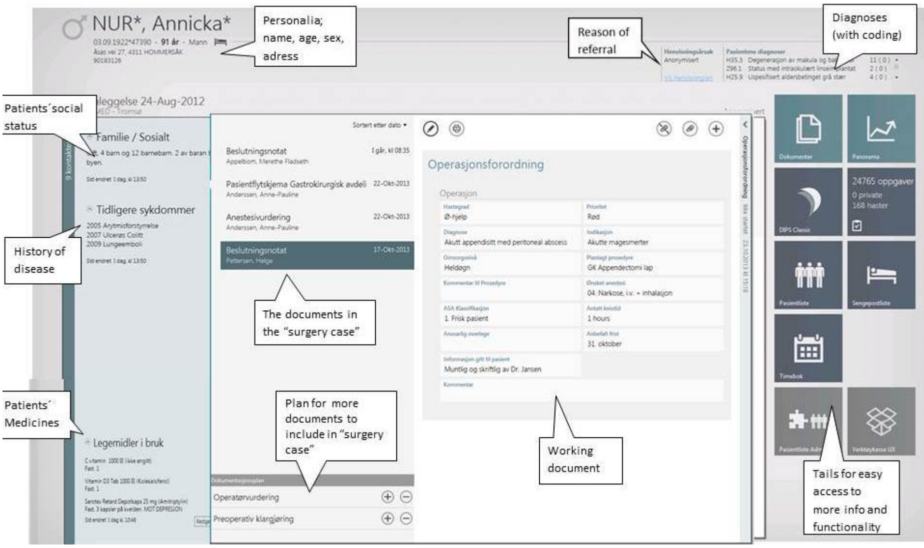


Fig. 1. The information displayed for “physician in outpatient clinic” in “NewArena”

The use of new technologies that rely on archetypes to structure data is expected to facilitate the generation of overviews of patient health information in ways that allow physicians to quickly get overviews of problems, the anamnesis and possibly test results before The information displayed in the NewArena system is adapted to each physician’s work tasks, so that if a physician enters the system from an outpatient clinic, the information needed in this setting is displayed.

A surgeon who decides that a patient is going for surgery has to create a decision-note within the NewArena system. This note is regarded as the “go” for the surgery planning process and serves as the head-document for what BigVendor has named a “case”. Other documents in a case typically include an admission note; a surgeon’s assessment note; an anesthesiologist’s note; nurses’ pre-, per- and post-operative notes and a discharge letter. Case collections of such documents make it easy to retrieve and view relevant documents in relation to one another. As the decision-note is the head-document, other documents are attached to it, and it cannot be removed from its associated case. Only when a head-document is created may other documents be added to a case.

This makes an inscription of the workflow – meaning that others cannot start their jobs before the surgeon has produced a decision-note that establishes assignments. This approach may seem straightforward but is not, as today's working routines are that secretaries prepare for the surgeons work by creating decision-notes using information retrieved from EPRs, like diagnoses and codes and indications for surgery. This way, surgeons need only fill in decision-notes with data derived from consultations and can focus on patients, instead of their computer screens, while patients are present. The inscribed workflow in NewArena does not allow for this; due to access-control rules, decision-notes, as part of surgery order forms, can only be established by physicians.

The decision-note, the anesthetic pre-operative assessment and the surgeon's assessment note together are to replace the mandatory fields in the old "surgery order form" for two reasons: First, in the old surgery order form, the fields for surgeons- and anesthesiologists´ information could be filled by anyone with access to the system. Hence, from a juridical point of view, it was not possible to tie the data in these fields to the physicians responsible for assessments. Replacing these fields with documents whose validity must be verified by their authors will create accountability. Second, the structuring of text will make the importation of data into surgery order forms much easier and, hence, hopefully make available more complete information needed to prepare for surgery. However, in many cases, outpatient consultations reveal the need for further information collection via such complementary assessments as radiology or other tests, before a final decision on surgery can be made. Thus, under NewArena, it may remain as difficult to author decision-notes or patient health information may be just as insufficient as it is today.

"In the outpatient consultation, you must consider if you have sufficient information to decide if the patient is in need of surgery or if additional tests like radiology examinations are required. I might approach this in two ways: Either I agree with the patient that he is going for surgery, but I need additional assessments that will be carried out as he is admitted for surgeries. Then, there will be a risk that the test might show surgery is not accurate treatment and hence result in cancellation. The other approach is to order the necessary tests and schedule another outpatient consultation when the results are ready, hence postponing the decision of surgery and spending more resources at the outpatient clinic." (Surgeon)

The new documents in NewArena – the decision-note, anesthetic pre-operative assessment and surgeon's assessment notes – are thus a part of the "surgery case" and are incorporated into the old standard surgery order form on the old platform, since the planning tool with current functionality is still there.

At present, secretaries processing the planning for surgery have to work on the old platform, as the functionality available in NewArena is currently only supporting physicians. This means that a secretary has to pick up the decision note in the "surgery case", and send it to the head surgeon, who must set deadlines for treatment, in accordance with national guidelines regarding prioritization of treatment. When the head surgeon has returned this note, the secretary starts booking resources and sends appointment letters to the patient. Put simply, for the patient to turn up, it is necessary that the surgery decision-note go into an electronic workflow managed by secretaries.

So far, there is no support for this electronic workflow in the new planning module. From what users have been presented at this stage, it may seem that surgery plans are made as decisions to operate are made and as surgery order forms are created. However, surgeons decision-notes are, in fact, only the "triggers" of surgery planning, and much work by secretaries and coordinating nurses goes into the planning process. This multidisciplinary perspective on the planning process is absent for now from the NewArena format.

Though NewArena focuses on physicians' work settings, data have not been structured, and hence data reuse is not yet possible. BigVendor has defined only a small number of archetypes to exemplify how these will work in NewArena. These are archetypes of clinical observations, like blood pressure, weight and height, data that surgeons register in the surgery order form. Archetypes of instructions and actions necessary to set up workflows have not been defined, and hence workflow support is missing. The further work on defining archetypes will have to be done by the hospitals themselves. The side-by side, alternating use of the old EPR and the NewArena platforms creates additional challenges: Documents produced by physicians in NewArena have to go into the old EPR to be processed further in the surgery-planning module by secretaries, and since the old EPR does not support structured text, the data structured by archetypes in NewArena cannot be employed in the old EPR. Thus, clinical data registered by archetypes in NewArena act like free-text documents during further processing in the old EPR. This represents no improvement in the existing planning tool:

"If I could decide, the surgery order form would not look like it does today at all. It is too much work to set it up; you cannot import data from the electronic record into the form. As an example, the patient may use 20 different medicines, and you have to write it all into the surgery order form. That is quite time consuming, besides being an error trap. The surgery planning tool should be connected to the electronic record, so that there would be only one click – set up the surgery order form – and then all the structured data in the record would be loaded directly and all I would have to do was to check it. No writing." (Surgeon)

At present, surgeons are not quite content with the functionality of the NewArena surgery-planning tool, as they have not yet got what they want the most: structured data that can be reused independent of documents, as described in the above quote. Still, they are happy with the new interface and the display of data. They say that the screen displayed in Figure 3 is a major improvement, as they have to open six different windows to display the same amount of information in the old system. They see the improved display as a kind of decision support in itself.

The secretaries and nurses who still have to work on the old platform cannot see their contribution to the planning process and are rather confused about how to ensure that patients are enrolled for surgery. Their major concern is the establishment of routines for being aware of new decision notes. Without an electronic workflow that notifies them of new notes for patients, they must find workarounds to replace such functionality.

6 Concluding Discussion

The intent of NewArena Clinician is to support health care workers´ decision-making at crucial steps in their clinical work. One such step occurs at outpatient clinics, where physicians decide whether a patient is going to have surgery. The key strategy supporting the design of NewArena focuses solely on this situation and the physician's role in it. Presumably, this does not only delimit the scope for the developers, but also promises to greatly enhance physicians' decision-making capabilities. However, our empirical evidence challenges this assumption, pointing to the socio-technical implications of the new software, which extend beyond the software's support for physician decision-making process at outpatient clinics. In particular, we discuss the relationship between the formal and informal as these relate to information, work responsibility and professional roles.

Formalization of Responsibility: In order to realize a working decision-support system, some degree of *formalization* appears necessary. The obvious initial step is, of course, the structuring of EPR content, which lays the basis for automatic/suggestive decision-making. However, the provision of appropriate decision-making support to the right professional (the surgeon) also requires the redistribution of work tasks between personnel involved in the work flow. This second, largely hidden step is an implicit formalization of the involved work processes. Just consider how the surgeon must set up the decision-note for the secretaries to be assigned their tasks, while earlier the secretaries set up and prepared the decision-note for the surgeon. Moreover, information that surgeons used to dictate, they must now write themselves and approve immediately in the form of decision-notes, before the rest of their teams may be assigned their tasks. The responsibility for decision-making hence becomes undisputed.

The formalization of physicians' work was expected to improve resource utilization later in the planning process. However, the information necessary to make determinations regarding surgery – e.g., the results of x-rays or heart- and lung-function tests – may not have been processed at the time of outpatient consultations, or such consultations may reveal the need for supplementary information, meaning that patient health information may not be complete even if information gathering is pushed to the front of the planning process. This jeopardizes the decision-note's role as the trigger of the planning process and the assignment of tasks to physicians' teams.

While it is relatively clear that formalization of decisions is necessary to surgery planning, a more open question is at what point in the planning process formalization should happen. Even if pushing formalization up front, to outpatient consultations, is anticipated to result in better utilization of resources, this may not be a good thing for physicians experiencing added workload (particularly when the additional workload comes as a surprise), as they will have to ensure the completeness of decision-notes before the next step in the planning process can be initiated. The effect on surgeons' work was worsened by the fact that data was not structured yet and hence could not be presented automatically to physicians, make suggestions or be reused.

The Importance of Informal Collaboration: In an implicit way, the formalization of roles and responsibilities appears to be a consequence of NewArena. The physician is accountable for the decision at the outpatient consultation stage. However, NewArena fails to capture the *informal* collaboration between physicians and secretaries in the planning process. Care processes do not rely solely on clinical information to be "seamless", and secretaries and nurses make important contributions to the logistical work involved in surgical planning, although these contributions are largely hidden to the untrained eye. Solely focusing on a single user-role in the development process ignores the interdisciplinary perspective, which is perhaps one of the main characteristics of health care that makes health care work. Secretaries facilitate physicians' work by preparing documents, transcribing notes, looking up and checking the status of activities, booking, scheduling and coordinating. They are aware of different physicians' preferences and make preparations accordingly, hence smoothing "the system" and contributing to the effective use of one of the most costly resources, surgeons` time.

Compatibility with Installed Base: This case illustrates the importance of ensuring that new ICT systems are compatible with installed bases. Stepwise implementation of the new functionality afforded by NewArena was anticipated to facilitate the surgical planning process but turned out to complicate it as well. In so far as they continue to serve as repositories for retrievable information, old EPRs must remain operable, so that users can alternate between new and old platforms. Such alternation is technically demanding, as well as challenging for users, as attention to what platform they are working on is required to accomplish their tasks. For now, secretaries are "mediators", and their mediating between platforms is actually process-supporting the system. Their concern for making routines that ensure patients are enrolled for treatment is compensating for the process-support missing in the EPR system.

References

1. Bird, L., Goodchild, A., Tun, Z.Z.: Experiences with a two-level modelling approach to electronic health records. J. Res. Pr. Inf. Technol. 35, 121–138 (2003)
2. Kawamoto, K., Houlihan, C.A., Balas, E.A., Lobach, D.F.: Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 330, 765 (2005)

3. Aarts, J., Ash, J., Berg, M.: Extending the understanding of computerized physician order entry: Implications for professional collaboration, workflow and quality of care. Int. J. Med. Inf. 76(suppl. 1), S4–S13 (2007)
4. Ash, J.S., Gorman, P.N., Lavelle, M., Payne, T.H., Massaro, T.A., Frantz, G.L., Lyman, J.A.: A Cross-site Qualitative Study of Physician Order Entry. J. Am. Med. Inform. Assoc. 10, 188–200 (2003)
5. Bonney, W.: Impacts and Risks of Adopting Clinical Decision Support Systems. In: Jao, C. (ed.) Efficient Decision Support Systems - Practice and Challenges in Biomedical Related Domain. InTeh (2011)
6. Rector, A.L.: Clinical terminology: why is it so hard? Methods Inf. Med. 38, 239–252 (1999)
7. Chase, M.D.: Engineering a Learning Healthcare System: A Look at the Future: Workshop Summary. National Academies Press (2011)
8. Garde, S., Knaup, P., Hovenga, E., Heard, S.: Towards semantic interoperability for electronic health records. Methods Inf. Med. 46, 332–343 (2007)
9. Klein, H.K., Myers, M.D.: A set of principles for conducting and evaluating interpretive field studies in information systems. MIS Q., 67–93 (1999)
10. Walsham, G.: Interpretive case studies in IS research: nature and method. Eur. J. Inf. Syst. 4, 74–81 (1995)
11. Beale, T., Heard, S.: Archetype Definitions and Principles in The openEHR foundation release 1.0. ...Architecture Overview in The openEHR foundation release 1.0.1. Presented at the (2007)
12. Berg, M.: Health information management: integrating information technology in health care work. Routledge, New York (2004)
13. Gasser, L.: The integration of computing and routine work. ACM Trans. Inf. Syst. TOIS 4, 205–225 (1986)
14. Berg, M.: Rationalizing Medical Work: Decision-support Techniques and Medical Practices. MIT Press (1997)
15. Kraut, R.E., Fish, R.S., Root, R.W., Chalfonte, B.L.: Informal communication in organizations: Form, function, and technology. In: Human Reactions to Technology: Claremont Symposium on Applied Social Psychology, pp. 145–199 (1990)
16. Munkvold, G., Ellingsen, G., Koksvik, H.: Formalizing work: reallocating redundancy. In: Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, pp. 59–68 (2006)
17. Hardstone, G., Hartswood, M., Procter, R., Slack, R., Voss, A., Rees, G.: Supporting informality: team working and integrated care records. In: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, pp. 142–151 (2004)
18. D MAG (2011), `http://viewer.webproof.com/pageflip/324/69431/index.html`. (acc. September 2013)

Architecture for Serious Games in Health Rehabilitation

Paula Alexandra Rego[1,3], Pedro Miguel Moreira[1,3], and Luís Paulo Reis[2,3]

[1] ESTG - Escola Superior de Tecnologia e Gestão, IPVC - Instituto Politécnico de Viana do Castelo, Av. do Atlântico s/n, 4900-348 Viana do Castelo, Portugal
[2] DSI - Departamento de Sistemas de Informação, Escola de Engenharia da Universidade do Minho, Campus de Azurém, 4800-058 Guimarães, Portugal
[3] LIACC, Laboratório de Inteligência Artificial e Ciência de Computadores, Rua Dr. Roberto Frias s/n, 4200 465 Porto, Portugal
{paularego,pmoreira}@estg.ipvc.pt, lpreis@dsi.uminho.pt

Abstract. Serious Games is a field of research that has been growing substantially with valuable contributions to many application areas. Traditional rehabilitation approaches are often considered repetitive and boring by the patients, resulting in difficulties to maintain their interest and to assure the completion of the treatment program. This paper describes a framework for the development of Serious Games that integrates a rich set of features including natural and multimodal interaction, social skills (collaboration and competitiveness) and progress monitoring which can be used to improve the designed games with direct benefits to the rehabilitation process. This improvement in the games' rehabilitation efficacy mainly arises from an increase in the patient's motivation when exercising the rehabilitation tasks due to the rich set of features provide by the framework.

Keywords: Serious Games, Rehabilitation, Architecture, Natural User Interfaces, Multimodal Interfaces, Health Informatics.

1 Introduction

Nowadays, the attention given to the development of tools for the rehabilitation of patients suffering from some kind of disability is growing. Additionally, the use of serious games as part of these rehabilitation tools is also increasing. Serious Games is a field of research that has been growing substantially and with valuable contributions in many areas.

As a multidisciplinary field of research, Serious Games can be applied into different problems in various and diverse areas such as: education, military, health rehabilitation. In such areas, computer games design and technology can be used to provide the user with an entertainment or enjoyment component, while attaining other main purpose that can be, respectively: educate, train (military strategies), and rehabilitate. The main purpose while playing the game is not the entertainment, or enjoyment, or fun, but another (more serious) purpose. In this research we focus the health rehabilitation domain.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_30,

The costs in rehabilitation of a variety of disabilities resulting from diseases or traumatic incidents are high and contribute to the social costs growth. It has been shown that most of these patients sometimes get into depression as soon as the rehabilitation program begins, contributing to increase even more the associated healthcare costs. In previous studies [1-4] we discuss issues and identified features that could benefit the rehabilitation process using Serious Games. Based on that the main objective of this work is to describe our work regarding the design and implementation of an hardware/software architecture for the development of Serious Games in Health Rehabilitation, to augment patients' motivation during their recuperation plan and that integrates the set of features we find as relevant to this process.

The rest of the paper is structured as follows. Section 2 presents the motivation for using serious games in health rehabilitation. Following, section 3 presents the state of art in natural and multimodal user interfaces in rehabilitation, section 4 presents design considerations and system requirements for the developed architecture and section 5 describes the architecture. Section 6 presents results of user testing using the described architecture, current work and some directions for future work. Section 7 presents the major conclusions.

2 Serious Games for Health Rehabilitation

In the rehabilitation area, games can be used to increase the motivation of patients in the rehabilitation sessions, which is the main problem evidenced in traditional therapy sessions [5], caused by the repetitive nature of exercises.

Most studies on rehabilitation of patients with some disability or impairment show that an effective rehabilitation must be early, intensive and repetitive [5, 6]. As such, these rehabilitation approaches are often considered as repetitive and boring by the patients, resulting in difficulties in maintaining their interest and in assuring that they complete the treatment program [5]. On the other hand, due to its nature, games can motivate and engage the patients' attention and distract him from his rehabilitation condition, as they require some motor and cognitive activity, they have a story, and can offer feedback and levels of challenge and difficulty that can be adapted to the patient skills. However, designing a game with all the features that could benefit the outcome of patient's rehabilitation is a complex task. Computer games design can thus offer valuable contributions in the development of more effective games for rehabilitation programs. In that sense, the identification, classification and assessment of game features relevant in the health rehabilitation area is very important for designing the games.

In a previous study by Rego et al.[1], relevant criteria for classifying Serious Games in Health Rehabilitation were defined. The study shows that the adoption of Serious Games in rehabilitation area is very recent in general and the majority of reported researches include only prototypes and games in early stages of test and development. Additionally, their effectiveness is hard to validate due to the fact that few prototypes were tested with a significant number of real patients. Additionally, in the same study, interaction technology was pointed as an important factor that can influence game effectiveness in a rehabilitation program. Human Computer

Interaction (HCI) is a very important part of a systems design. The way the user communicates with the application/computer system determines its interactivity, accessibility and affects the user satisfaction with the system. In the rehabilitation area in particular, the interface can be used to improve interaction with the patient in a way that he can attain his goals more easily, more rapidly and with an higher level of satisfaction.

3 Natural and Multimodal User Interfaces in Rehabilitation

Recent directions in HCI design are toward the use of new methods of interaction: intelligent/adaptive, multimodal and natural. Natural interaction techniques aim to mimic real-world interaction by using user body movements, gestures, voice or sound input recognition and actions similar to those used in the physical world to accomplish the same task and to diminish the artificial communication devices for the interaction with the computer system. A multimodal interface integrates information from different input modalities (speech, gesture, writing, and others), providing a flexible use of input modes and an increase in the user degrees of freedom. In adaptive interfaces, the system tries to adapt itself to the way the user experiences the interaction. User experience can also be measured using several modalities of input, including biofeedback, information from voice, face and head. Brain signals contain information about user experience, along with other information measured from the body: bio signals as heart rate, body temperature, respiration, and muscle tension.

In the last years several prototypes using these new forms of interaction have been developed and published in rehabilitation area. Rego et al. [2] summarize relevant works of serious games for rehabilitation using natural forms of interaction.

A significant number of recent works involve motion tracking devices used to track patient's movements to detect and correct incorrect motion patterns during the rehabilitation exercises and/or to recover from several motor capabilities problems. The success of motion controlled game hardware and peripherals enables the application of games for rehabilitation, not only in hospitals and clinics, but also at home [7]. Several studies about the feasibility and results of using Wii and Wii sports game software showed similar improvements in functional outcomes, compared to conventional therapy, but the level of satisfaction and motivation reported by the users was higher using the Wii [8]. Some authors, besides developing custom applications to tailor the Wii game system with game tasks specific of rehabilitation, also use physical custom setups using the Wiimote as the base of the game system to obtain more affordable systems [9, 10], making it possible to be used at a home rehabilitation. Studies of balance and stability improvement on users with impairments typically involve the use of the Wii Balance Board integrated in custom applications [11]. On the other hand, actual mobile phone devices are available with integrated sensors enabling the development of applications that can sense motion and react to the mobile phone orientation. An example is given by Spina et al. [12] presenting an application that is based only in smartphone-integrated sensors to monitor and access the performance of the rehabilitation exercises in patients suffering from chronic pulmonary obstructive disease.

Kinect can be used in health and rehabilitation applications in several scenarios. Chang et al. [13] describe the use of Kinect in physical rehabilitation. The pilot study showed that the participants improve exercise performance during the intervention phases and that their motivation for physical rehabilitation has increased.

Data Gloves can be used as input devices to capture physical data such as bending of fingers, providing accurate data for motion capture that is interpreted by the software that accompanies the glove. In [14], the 5DT glove is connected to a PlayStation3 game console installed at their home and remotely monitored to study hand function in adolescents with hemiplegic cerebral palsy.

In a tangible user interface (TUI), the user interacts with digital information using physical objects of the real world as user interfaces, which is important in rehabilitation to diminish the cognitive load needed in the interaction. A relevant example is given in [15], describing the development of applications for motor rehabilitation based on a multi touch tabletop screen, providing tasks that closely mimic existing rehabilitation activities .

Touch/Multi-touch is an interaction modality that enables do detect contact of body parts (usually fingers) with a surface. The interface mechanism typically tracks the position of the contacts to recognize gestures. Additionally, the use of pressure information is present in some systems. In [16], a low-cost multi-touch surface platform was developed with several games for elderly cognitive training, and using a robust visual finger tracking algorithm.

Another emergent form of interaction is the one provided by BCI (brain-computer interfaces) systems enabling human-computer communication only by analysis of human brain signals. This technology can improve the daily life experience of people with neurological and motor control disorders after stroke or other traumatic brain disorders, by giving information of the current state of brain activity. On the other hand, it may increase the efficacy of a rehabilitation program and improve muscle control for the patient, since the use of brain signals can supplement an impaired muscle control. Its uses include helping disabled users to communicate with a machine [17], in neurofeedback games [18] and in video games as game controllers [19].

Facial expressions recognition systems and eye gaze tracking may also play a key role in HCI and natural user interfaces and in Serious Games in particular. The user facial expression gives a measure of his satisfaction or experience in the game and thus the level of interaction or experiences in the game can be adapted to that measure [20]. Eye gaze tracking is the process of detecting and tracking the point of where the user is looking at – the point of gaze. There are some examples of the use of eye gaze tracking in rehabilitation domain [21, 22].

Some of more recent projects are based on multimodal or adaptive interfaces, enabling various forms of inputs. Examples of these solutions can be seen in [23] and [24-32]. Faria et al. [24-32] uses the concept of an intelligent wheelchair that is controlled using high level commands in a multimodal interface that uses voice, facial expressions, head movements and joystick as the main input forms and users have the option to choose the input type that best fits their needs, enabling to increase their safety and comfort. Users can convey information also in the form of gestures. Trigueiros et al. [33-36] uses a gestures concept for a vision-based hand gesture recognition system that can be integrated in a HCI system.

4 Design Considerations and Requirements

In the current section we present our developed framework for serious games. We focus the presented discussion on the distinguishable features that we have considered to include on its design. In particular the developed architecture encompasses modules that allow for a set of natural and multimodal interaction, social skills (collaboration and competitiveness) and progress monitoring which can be used to improve the rehabilitation process.

4.1 Design of the Rehabilitation Tasks

The design of the rehabilitation tasks is a complex issue that depends on the kind of disabilities of the patient in rehabilitation. These disabilities can be diverse and so it would be necessary a program adapted to the patient needs in order to achieve the best results in the rehabilitation program. Based on that, the design of the rehabilitation tasks would require a multidisciplinary team from the medical field composed with doctors, nurses, psychologists, physiotherapists, and therapists. Also it would require professionals of the game design field. This is difficult since it implies communication between different professionals (and languages).

4.2 Interface with the System

Patients in rehabilitation have several cognitive and motor disabilities which make their interaction with the system more difficult. Eliminating or diminishing the use of artificial devices specific for interaction and enabling users to interact directly with objects of the real world (for example, using a racquet in a tennis game) without using an intermediary device, natural interfaces enable a more easy, intuitive, and realistic interaction. The use of only a single device can limit accessibility, which is why the combination of multiple input and output modalities is an important objective People in rehabilitation can have disabilities that prevent them from using traditional techniques, such as mouse or keyboard, and for these, alternative means of interaction should be considered. With a multimodal or adaptive interface the patient can choose the input form that suits better to his disability condition. Additionally, more immersive and realistic user experiences are provided.

4.3 Social Interaction

Games and game concepts have a stimulating effect on engagement in tasks. In addition, social interaction in games has demonstrated positive effects in healthy subjects in terms of player experience [37]. Social factors included in many games have had effects on people's enjoyment when playing the games. This may increase the social connectedness and motivation in rehabilitation patients to keep practicing the exercises they need. Social play can be achieved by the incorporation of competitive and collaborative tasks in the rehabilitation games [38].

4.4 Home Rehabilitation and Progress Monitoring

From literature review we can see work developments are being made towards the creation of applications that use low-cost devices to be accessible and affordable for people in general. In particular in the Rehabilitation domain this is an essential aspect.

The use of low cost systems can foster a home rehabilitation and create more opportunities for the patients to train the necessary exercises in their recuperation plan. This is another important aspect to have in consideration in our framework: the use of an interaction technology that can be inexpensive to be feasible for home use and thus promoting for the patients to play more frequently the games. The home rehabilitation tool should not be used to replace the therapist, but instead serve as a supplement to the assisted rehabilitation, allowing a recuperation plan done partly in home. The patient can then exercise the therapy plan without needing the continued assistance from a health professional, which may not be always available. Additionally, it can save the patient from having to move to the rehabilitation center so often, which in some cases might be problematic or uncomfortable, depending on factors such as the distance from the rehabilitation center and on the disability condition of the patient.

Facilities to detect, log, and analyze patient's tasks (cognitive or motor) prescribed by the therapists, in order to evaluate the patient's progress in the therapy are also key feature to include with respect to serious game for health rehabilitation.

5 Architecture for Serious Games for Health Rehabilitation

In order to demonstrate that Serious Games can be used to increase motivation of patients in rehabilitation, a framework was developed for Serious Games in Health Rehabilitation that integrates the set of features identified as relevant to improve the rehabilitation process, such as: natural and multimodal interaction, social skills (collaboration and competitiveness) and progress monitoring. Figure 1 presents the architecture diagram. The architecture comprises several distinct layers of input recognition. The first layer represents input modalities in raw form. The second layer represents an abstract recognition from the inputs received in raw form. In the third layer we can see a combination/fusion of more modalities among each other: gestures recognition, emotion recognition, representing higher levels of recognition processes.

The main modules of the architecture are:

Game Engine: is the game component that is most common to all games, representing the most generic component of the game logic.

Game Database: the most specific component of the games that represents the repository of all the games that are available for use in the rehabilitation session.

Social Networking: this module is responsible for creating the mechanisms for users (patients) to group together in social networks in order to communicate with each other and providing tools to mediate and facilitate social interactions.

Competition/Collaboration: responsible for creating the interaction mechanisms of competition and collaboration among users. It includes user modelling and profiling for establishing the handicaps with which each user will play with others, even if at different levels of the rehabilitation process.

User Management and Profiling: responsible for managing the information associated to the users, including their profiles, therapies they have to follow, state of the therapy, progress indicators, handicaps, etc.

Logging & Monitoring: module responsible for registering the users' logs, session duration, last difficulty level attained in each session in order to monitor the progress of each patient during the therapy.

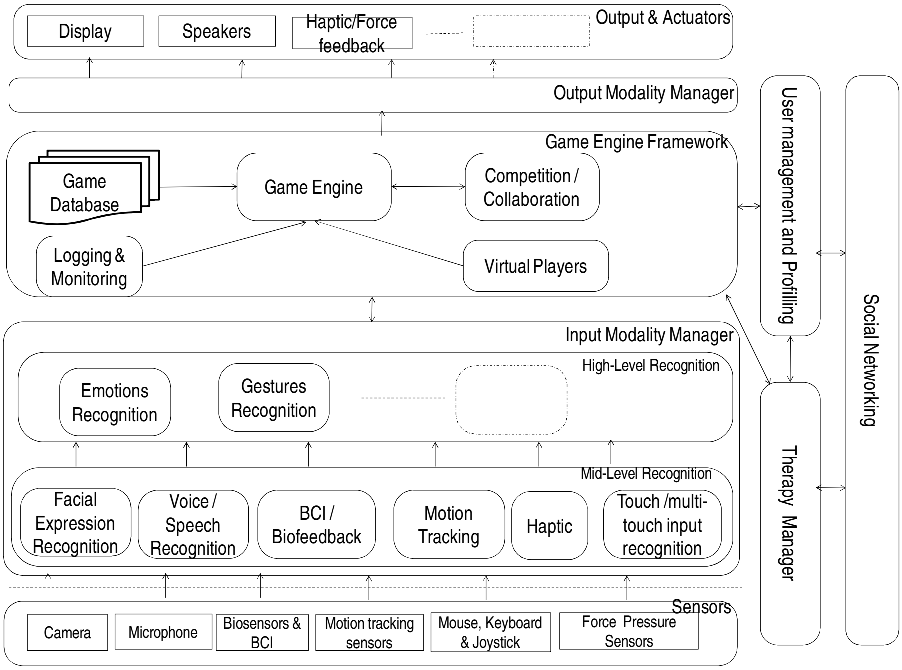


Fig. 1. Architecture Diagram

Virtual Players: module responsible for defining the agents that will assume the responsibility of users/players in the application sessions. This will enable some multi-player games even with only one real user.

Therapist Manager: provides a set of tools to configure, monitor and analyse the configuration of the prescribed therapy (the games that each patient will play, the duration of the sessions, the progress, etc.)

Input Modality Manager: manages the different input modalities and it can be adapted to the different disabilities of the users. It can be decomposed in two levels of recognition: in a mid-level we have individual forms of recognition: facial expressions, voice/speech, BCI/Biofeedback, motion tracking, Haptic and Tourch/multi-touch. In a higher level we can have a fusion of some of the modalities present in the mid-level that could compose a module of emotions recognition or gestures recognition, for example. The machine interpretation of human behaviour is very important in HCI, for achieving natural forms of interaction. In a mid-level recognition, users can communicate or interact with the machine by several forms that need to be interpreted by the machine. Users can convey messages in the form of body gestures, facial signals, speech, and emotions, among others. In a higher level of recognition, gestures recognition systems can have a multimodal nature, representing a fusion of the different input modalities present on the mid-level: gestures can be translated in: hand gestures; or body positions and movements that control a virtual environment on the screen, being captured by cameras or by the use of data sense gloves; or they can be a sequence of finger positions and movements in a multi-touch table; or they can be face gestures. Emotions have also a multimodal nature. They can be expressed through several modalities (and channels) of the mid-level recognition module: verbally by the use of emotional vocabulary or by expressing several non-verbal cues such as facial expressions, voice intonation, postures, gestures, and physiological changes. The decoding of these cues is essential to interpret the correct message. It is therefore necessary in a multimodal system to make a fusion of these different modalities to achieve a reliable (accurate) assessment.

Output Modality Manager: manages the different output modalities. It can include a display, sound device, haptic device and force-feedback equipment and it can be adapted to a specific game, environment or user.

The platform of the architecture provides virtual player and competition/ collaboration modules that integrate with the social/users community component (social network).

6 Results and Discussion

We developed a prototype of the architecture to test the effect of the introduction of new forms of interaction in rehabilitation serious games. The prototype included three different interaction modalities: mouse, sound and motion. We tested the usability of the prototype with 20 healthy users and results showed that the sound and motion options were the modalities that provided the most involvement of the users and were the options they showed most interest to play again.

We are analyzing the multimodal capacity of the platform such that we can choose the input modality that best suits to the user capabilities or disability condition. One of the main contributions of our architecture is to allow that the user interface with the system is adaptive in the sense that it can be configured to the best input modality taking into consideration the user capabilities and the proposed tasks, reducing the physical and mental fatigue in the interaction with the system.

We also intend to study the efficacy of more natural modalities of interaction in the modules of competition/collaboration and virtual players. We point the originality of our architecture that derives from the fact that it provides a competitive/collaborative component, a multimodal integrated management component and virtual players that can surrogate real players when they do not exist. In that sense, in a multi-player game, the virtual player can substitute the patient when there are not patients available.

7 Conclusions

Several motion tracking devices have been introduced by the gaming industry, enabling the use of game consoles for home-based rehabilitation with some type of gesture based game play. Noticeable examples of the later are the Nintendo Wii with the Wiimote, PlayStation 2 EyeToy, PlayStation 3 Eye and Microsoft Kinect and PlayStation 3 Move systems. Modalities as speech recognition, facial expression recognition and touch input are also used in rehabilitation to become accessible by patients with several motor and cognitive impairments. Although many prototypes are being developed for rehabilitation serious games, the solutions available don´t fulfill yet all the requirements/features that must be defined for each type of therapy/patient. In order to demonstrate that Serious Games can be used to improve the rehabilitation process, a framework is described for the development of Serious Games that integrates a set of features identified as relevant to improve the rehabilitation process, such as: natural and multimodal interaction, social skills (collaboration and competitiveness) and progress monitoring.

References

1. Rego, P., Moreira, P.M., Reis, L.P.: Serious Games for Rehabilitation: A Survey and a Classification towards a Taxonomy. In: 5th Iberian Conference on Information Systems and Technologies, pp. 349–354 (2010)
2. Rego, P.A., Moreira, P.M., Reis, L.P.: Natural User Interfaces in Serious Games for Rehabilitation: a Prototype and Playability Study. In: First Iberian Workshop on Serious Games and Meaningful Play (SGaMePlay 2011) - Proceedings of the 6th Iberian Conference on Information Systems and Technologies, pp. 229–232 (2011)
3. Rego, P.A., Moreira, P.M., Reis, L.P.: New Forms of Interaction in Serious Games for Rehabilitation. In: Cruz-Cunha, M.M. (ed.) Handbook of Research on Serious Games as Educational, Business, and Research Tools: Development and Design. IGI Global (2012)
4. Mendes, L., Dores, A.R., Rego, P.A., Moreira, P.M., Barbosa, F., Reis, L.P., Viana, J., Coelho, A., Sousa, A.: Virtual centre for the rehabilitation of road accident victims (VICERAVI). In: 7th Iberian Conference on Information Systems and Technologies (CISTI 2012), pp. 817–822. AISTI (2012)
5. Burke, J.W., McNeill, M.D.J., Charles, D.K., Morrow, P.J., Crosbie, J.H., McDonough, S.M.: Optimising engagement for stroke rehabilitation using serious games. The Visual Computer 25, 1085–1099 (2009)
6. Burke, J.W., McNeill, M.D.J., Charles, D.K., Morrow, P.J., Crosbie, J.H., McDonough, S.M.: Augmented Reality Games for Upper-Limb Stroke Rehabilitation. In: 2010 Second International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES), pp. 75–78 (2010)
7. Veltkamp, R.: Wii games for Physical therapy at home. Overview of the GATE Project (Game Research for Training and Entertainment) 2007-2010. GATE | Design & Production, Utrecht (2010)
8. Joo, L.Y., Yin, T.S., Xu, D., Thia, E., Fen, C.P., Kuah, C.W.K., Kong, K.-H.: A feasibility study using interactive commercial off-the-shelf computer gaming in upper limb rehabilitation in patients after stroke. Journal of Rehabilitation Medicine 42, 437–441 (2010)

9. Decker, J., Li, H., Losowyj, D., Prakash, V.: Wiihabilitation: Rehabilitation of Wrist Flexion and Extension Using a Wiimote-Based Game System. Rutgers University (2009)
10. Brown, D., Standen, P., Barker, M., Steven, B., Lewis, J., Walker, M.: Designing Serious Games Using Nintendo's Wii Mote Controller for Upper Limb Stroke Rehabilitation. In: Proceedings of the 3rd European Conference on Games Based Learning, pp. 68–77. Academic Publishing Limited, FH JOANNEUM University of Applied Science (2009)
11. Saposnik, G., Teasell, R., Mamdani, M., Hall, J., McIlroy, W., Cheung, D., Thorpe, K., Cohen, L., Bayley, M.: Effectiveness of virtual reality using Wii gaming technology in stroke rehabilitation: a pilot randomized clinical trial and proof of principle. Stroke 41, 1477–1484 (2010)
12. Spina, G., Huang, G., Vaes, A., Spruit, M., Amft, O.: COPDTrainer: a smartphone-based motion rehabilitation training system with real-time acoustic feedback. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 597–606. ACM, Zurich (2013)
13. Chang, Y.-J., Chen, S.-F., Huang, J.-D.: A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. Research in Developmental Disabilities 32, 2566–2570 (2011)
14. Golomb, M.R., McDonald, B.C., Warden, S.J., Yonkman, J., Saykin, A.J., Shirley, B., Huber, M., Rabin, B., AbdelBaky, M., Nwosu, M.E., Barkat-Masih, M., Burdea, G.C.: In-Home Virtual Reality Videogame Telerehabilitation in Adolescents With Hemiplegic Cerebral Palsy. Archives of Physical Medicine and Rehabilitation 91, 1.e1–8.e1 (2010)
15. Annett, M., Anderson, F., Goertzen, D., Halton, J., Ranson, Q., Bischof, W.F., Boulanger, P.: Using a multi-touch tabletop for upper extremity motor rehabilitation. In: Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group, OZCHI 2009, pp. 261–264. ACM, New York (2009)
16. Theodoreli, V., Petsatodis, T., Soldatos, J., Talantzis, F., Pnevmatikakis, A.: A Low-Cost Multi-Touch Surface Device supporting Effective Ergonomic Cognitive Training for the Elderly (2010)
17. Rebsamen, B., Guan, C., Zhang, H., Wang, C., Teo, C., Ang, M.H., Burdet, E.: A Brain Controlled Wheelchair to Navigate in Familiar Environments. IEEE Transactions on Neural Systems and Rehabilitation Engineering 18, 590–598 (2010)
18. Wang, Q., Sourina, O., Nguyen, M.: Fractal dimension based neurofeedback in serious games. The Visual Computer, 1–11 (2011)
19. Lécuyer, A., Lotte, F., Reilly, R.B., Leeb, R., Hirose, M., Slater, M.: Brain-computer interfaces, virtual reality, and videogames. Computer 41, 66–72 (2008)
20. Hudlicka, E.: Affective Game Engines: Motivation and Requirements. In: Proceedings of the 4th International Conference on Foundations of Digital Games (2009)
21. Trojano, L., Moretta, P., Estraneo, A.: Communicating using the eyes without remembering it: cognitive rehabilitation in a severely brain-injured patient with amnesia, tetraplegia and anarthria. Journal of Rehabilitation Medicine 41, 393–396 (2009)
22. Trojano, L., Moretta, P., Estraneo, A., Santoro, L.: Neuropsychologic Assessment and Cognitive Rehabilitation in a Patient With Locked-In Syndrome and Left Neglect. Archives of Physical Medicine and Rehabilitation 91, 498–502 (2010)
23. Omelina, L., Jansen, B., Bonnechère, B., Van Sint Jan, S., Cornelis, J.: Serious games for physical rehabilitation: designing highly configurable and adaptable games. In: Proc. 9th Conference on Disability, Virtual Reality and Assoc. Technologies (ICDVRAT) (2012)
24. Faria, B.M., Vasconcelos, S., Reis, L.P., Lau, N.: A Methodology for Creating Intelligent Wheelchair Users' Profiles. In: ICAART 2012, 4th International Conference on Agents and Artificial Intelligence, pp. 171–179 (2012)

25. Faria, B.M., Reis, L.P., Lau, N.: Cerebral Palsy EEG Signals Classification: Facial Expressions and Thoughts for Driving an Intelligent Wheelchair. In: IEEE 12th International Conference on Data Mining Workshops (ICDMW) 2012, Biological Data Mining and its Applications in Healthcare Workshop, pp. 33–40 (2012)
26. Faria, B., Ferreira, L., Reis, L., Lau, N., Petry, M.: Intelligent Wheelchair Manual Control Methods. In: Correia, L., Reis, L.P., Cascalho, J. (eds.) EPIA 2013. LNCS, vol. 8154, pp. 271–282. Springer, Heidelberg (2013)
27. Faria, B., Reis, L., Lau, N., Soares, J., Vasconcelos, S.: Patient Classification and Automatic Configuration of an Intelligent Wheelchair. In: Filipe, J., Fred, A. (eds.) Agents and Artificial Intelligence, vol. 358, pp. 268–282. Springer, Heidelberg (2013)
28. Faria, B.M., Vasconcelos, S., Reis, L.P.: Evaluation of Distinct Input Methods of an Intelligent Wheelchair in Simulated and Real Environments: A Performance and Usability Study. Assistive Technology: The Official Journal of RESNA 25, 88–98 (2013)
29. Petry, M., Moreira, A.P., Faria, B.M., Reis, L.P.: IntellWheels: Intelligent Wheelchair With User-centered Design. In: IEEE 15th International Conference on e-Health Networking, Applications and Services - Healthcom, pp. 392–396 (2013)
30. Faria, B.M., Reis, L.P., Lau, N.: Adapted Control Methods for Cerebral Palsy Users of an Intelligent Wheelchair. Special Issue on Autonomous Robot Systems, Journal of Intelligent and Robotic Systems (to appear)
31. Faria, B.M., Reis, L.P., Lau, N.: A Survey on Intelligent Wheelchair Prototypes and Simulators. In: WorldCIST 2014 - World Conference on Information Systems and Technologies, Madeira, Portugal (to appear, 2014)
32. Faria, B.M., Silva, A., Faias, J., Reis, L.P., Lau, N.: Intelligent Wheelchair Driving: A Comparative Study of Cerebral Palsy Adults with Distinct Boccia Experience. In: WorldCIST 2014 - World Conference on Information Systems and Technologies, Madeira, Portugal (to appear, 2014)
33. Trigueiros, P., Ribeiro, F., Reis, L.P.: A comparison of machine learning algorithms applied to hand gesture recognition. In: CISTI 2012 - 7th Iberian Conference on Information Systems and Technologies, pp. 41–46 (2012)
34. Trigueiros, P., Ribeiro, F., Reis, L.P.: A comparative study of different image features for hand gesture machine learning. In: ICAART 2013 - 5th International Conference on Agents and Artificial Intelligence, pp. 51–61 (2013)
35. Trigueiros, P., Ribeiro, F., Reis, L.P.: Vision Based Referee Sign Language Recognition System for the RoboCup MSL League. In: 17th RoboCup International Symposium, Robocup 2013. Springer (2013) (to appear)
36. Trigueiros, P., Ribeiro, F., Reis, L.P.: Vision-based Portuguese Sign Language Recognition System. In: World Conference on Information Systems and Technologies, WorldCIST 2014, Madeira, Portugal (to appear, 2014)
37. Vanacken, L., Notelaers, S., Raymaekers, C., Coninx, K., van den Hoogen, W., Jsselsteijn, W.I., Feys, P.: Game-Based Collaborative Training for Arm Rehabilitation of MS Patients: A Proof-of-concept Game. In: Proceedings of the GameDays, pp. 65–75 (2010)
38. Alankus, G., Lazar, A., May, M., Kelleher, C.: Towards customizable games for stroke rehabilitation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2113–2122. ACM, Atlanta (2010)

A Simple Movement Classification System for Smartphones with Accelerometer

Micael Costa[1] and Paulo Fazendeiro[2]

[1] Departamento de Informática, Universidade da Beira Interior, Covilhã, Portugal
[2] Instituto de Telecomunicações, Universidade da Beira Interior, Covilhã, Portugal
acmicael@gmail.com, fazendeiro@ubi.pt

Abstract. The penetration rate of smartphones has been growing during the past years. Today's smartphones provide access to the Internet, GPS navigation and are equipped with cameras and various sensors: accelerometer, gyroscope, proximity sensor and light sensor among others. The main objective of this paper is to propose a movement classification system whose main characteristic is to obtain the numeric acceleration values along the three axes of the accelerometer and the subsequent conversion to one limited set of linguistic terms. The resulting simple classification of movements is sufficient to classify correctly (from a person's point of view) the smartphone movements as well as their intensity. The validation tests and the proof of concept presented in this article open the path to the development of applications for physiotherapy and mobile health, especially those aimed at improving health and welfare through motivation for physical activity.

Keywords: Mobile Computing, Smartphone, Accelerometer, mHealth.

1 Introduction

The mobile phone industry is paradigmatic of a high-tech sector which R&D component is heavily dependent on the developments in the computer industry. Nowadays, there are smartphones, "intelligent" mobile phones that emerged in the 90s, with characteristics similar to the ones of a personal computer. Those are mobile phones with a built-in operating system that have diverse advanced features, such as image and video capabilities, as well as email, sensors, and GPS navigation. Moreover they allow the installation of additional programs according to the preference of each user, thus facilitating a wide range of activities and knowledge acquisition tasks. A research made by Strategy Analytics [1] revealed that there are over one billion smartphones operating worldwide and the number is expected to continue to grow exponentially in the upcoming years.

The mobile phones are focused on the user and the smartphone further enhances this aspect, allowing a greater degree of customization by installing applications that are no longer just a means of communication, but also a means of entertainment, multimedia and welfare center [2]. Among the applications that have been finding an increased number of developments stands out the ones related with mobile health (mHealth).

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_31,

One of the key aspects of mHealth is focused on the use of the mobile phone as a means to help users to be more aware of their healthcare needs and, consequently, their quality of life. These applications, when incorporated in smartphones, help to reduce health costs and bring significant improvements in health in the years to come [3]. Actually the connection of mobile devices to health is already a reality; mHealth is growing mainly because the mobile technologies provide increased resources (as sensors, internet access and better quality of picture and sound, among others) at a low cost. The use of sensors like the accelerometer in smartphones allows the increase of interactivity and dynamism with users.

In the last few years the first applications that connect the mHealth with sensors such as the accelerometer began to appear. Some examples include: smartphone as pedometer [4], detection of falls [5], detection of seizure-like movements [6], diagnosis and monitoring of patients with Parkinson [7].

This paper proposes a movement identification system that aims to convert the numerical acceleration values in linguistic terms, so they are easily understandable and also illustrate the description of the movements and the intensity with which they are performed by the smartphone. The paper is organized as follows: Section 2 describes the validation tests for the smartphone used in this study. Section 3 presents the proposed movement classification method. In Sec. 4 an example of application of this method is demonstrated. Section 5 draws some conclusions and directions for future developments.

2 Testing and Validation

This section describes the validation tests aimed at comparing simultaneously the behavior of an accelerometer embedded in a smartphone against an accelerometer used for scientific applications. The smartphone Sapo a5 was used with operating system Android version 2.2 Froyo. The Sapo a5 is based on ZTE Blade a mobile phone manufactured by ZTE Corporation. The dimensions of the device are 57mm long, 115mm high and 12.5mm wide. The weight is 110g and it uses a 600 Mhz processor. It has an accelerometer, a digital compass, a proximity and light sensors.

The accelerometer used in Sapo a5 is LIS302DL developed by Intersil Corporation. The LIS302DL is a three-axis accelerometer with a sampling output rate of 100Hz or 400Hz. It uses a power rate 2.16V to 3.6V and a power consumption of less than 1mW [8].

The TSD109F is an accelerometer from BIOPAC Systems Inc., a company specialized in instruments, components and accessories designed for scientific and applied research. More specifically the TSD109F ±50g allows to simultaneously measure acceleration in three-axis with a sampling rate of 2000 Hz. It uses a power rate of +5V, by cable, weighs 17g and its dimensions are: 33mm long, 19mm high and 28mm wide [9].

The accelerometers were placed simultaneously on the weights of a latissimus machine, used for dorsal strengthening (see Fig. 1). The repetitions of the movement were measured in one-minute time lapse.

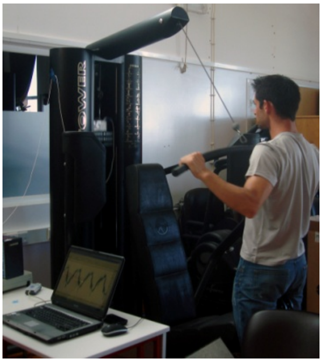


Fig. 1. Motion capture using accelerometers LIS302DL and TSD109F, coupled in a latissimus machine

Based on these two accelerometers LIS302DL (smartphone Sapo a5) and TSD109F (Biopac) three sets of tests were performed, corresponding to the axis X, Y and Z to show that the smartphone's accelerometer can achieve results comparable to an accelerometer used in biomechanics research even with a smaller sample rate. The maximum supported sampling rate was used in both cases. In the Biopac accelerometer, the sampling rate was fixed at 2000 values per second. In the smartphone accelerometer the sampling rate was variable, for example it may present 8, 10, 12, 16, 17 values per second once the process of capturing the sensor values runs in parallel with the others that are being executed on the smartphone.

In the first example, we tested the X-axis in the accelerometer smartphone. Fig. 2 (left) shows a graph with data from the TSD109F accelerometer. The Fig. 2 (right) refers to the values of the smartphone's accelerometer LIS302DL.

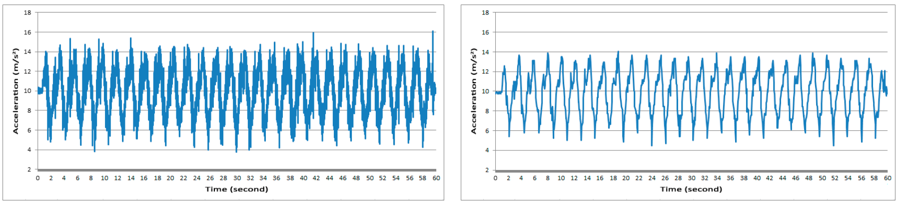
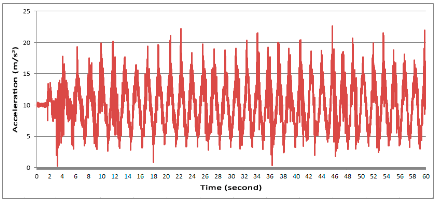


Fig. 2. Acceleration values in the X-axis for accelerometers coupled with a latissimus machine. Left: TSD109F accelerometer, right: LIS302DL accelerometer.

A small program was created to read the data captured by the two accelerometers, TSD109F e LIS302DL. The same program checks when the acceleration values exceed a high threshold. After that it checks whether the values pass by the origin. Only one movement counts, if a lower threshold is reached next. The values for the upper and lower thresholds of both accelerometers were established empirically as 12 and 8. When the program was executed it generated the result of 27 for both accelerometers.

Fig. 3 (left) shows the values obtained for the axis Y of the TSD109F. Fig. 3 (right) represents the values obtained for the axis Y of the smartphone's accelerometer after properly rotated and again united to latissimus machine. The same method was executed and in this case the output generated for the two accelerometers was 34.



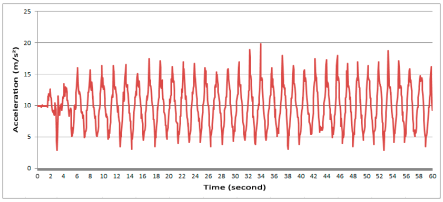
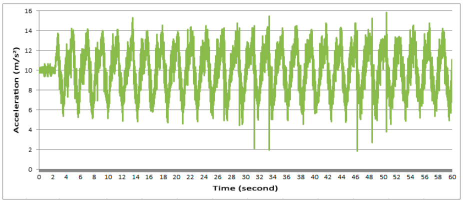


Fig. 3. Acceleration values in the Y-axis for accelerometers coupled with a latissimus machine. Left: TSD109F accelerometer, right: LIS302DL accelerometer.

The final axis to be tested was Z. Fig. 4 (left) illustrates the values acquired by the TSD109F accelerometer in which the total number of movements was 27. The graph of the Z-axis for the smartphone is represented in Fig. 4 (right) and the LIS302DL accelerometer output is also 27.



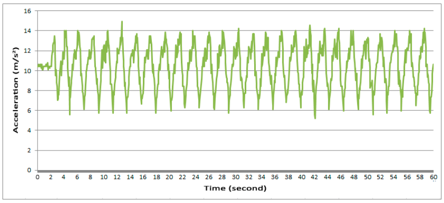


Fig. 4. Acceleration values in the Z-axis for accelerometers coupled with a latissimus machine. Left: TSD109F accelerometer, right: LIS302DL accelerometer.

It should be noted that the acceleration varies from user to user. In order to verify this fact, different users performed the tests in different axes. There is a remarkable inter-variability between subjects in terms of the maximum and minimum values of acceleration. However, the number of recorded movements was always coincident in the two accelerometers and it was consistent with visual observation.

Based on these tests we can say that the used smartphone accelerometer, even though it presents values of low precision in fine grain movements due to the very low and irregular sampling rate, under controlled conditions can have results comparable to the Biopac's accelerometer. In other words, for applications in which absolute accuracy is not required and the aim is to count the perceptible movements, that are visible to the naked eye, the accelerometer of a low-cost smartphone presents a performance as good as a top of the range accelerometer.

3 Movement Classification

We developed a movement classification system, which converts the numerical values of the smartphone's accelerometer in linguistic terms by characterizing the direction of movement (up, down, left, right, front, back) and describing the intensity that is applied (strong, normal, weak).

3.1 Determining the Direction of Movement

The first step was to read the accelerometer values through a smartphone application that stores values in an external file. Many sets of testing were conducted with upward, downward, right, left, front and back movements.

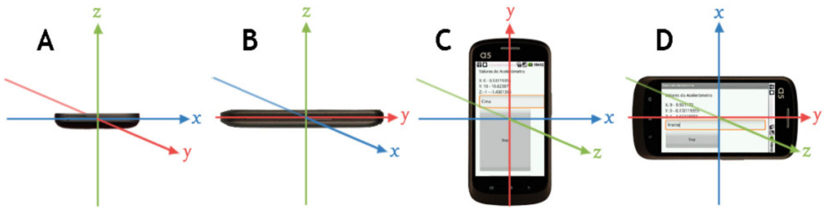


Fig. 5. A: smartphone lying down in a portrait position; B: smartphone lying down in a landscape position; C: smartphone upright in a portrait position; D: smartphone upright/sideways in a landscape position

Table 1 identifies the movements with the smartphone in a lying down portrait position (Fig. 5A). The Z-axis represents the 'up or down' movements and its value is approximately 10 because, in this case, it's where the force of Earth's gravity is applied. The X-axis represents the 'left or right' movements and the Y-axis represents the 'front or back' movements.

The intensity of the movement is better explained in session 3.2. The maximum and minimum absolute values depend on the applications and can be calibrated in order to change the intensity and define the level of sensitivity.

Table 1. Movements with the smartphone lying down in a portrait position (Fig. 5A)

Axis	Axis value	Movement	1st limit to attain	2nd limit to attain
Z	10	Up	Maximum	Minimum
Z	10	Down	Minimum	Maximum
X	0	Left	Maximum	Minimum
X	0	Right	Minimum	Maximum
Y	0	Front	Maximum	Minimum
Y	0	Back	Minimum	Maximum

It is concluded that there is an upward movement when in the Z axis there is an absolute maximum only after an absolute minimum; symmetrically there is a downward movement, when on the Z axis there is an absolute minimum and then an absolute maximum. Fig. 6 illustrates a real case of upward and downward movements.

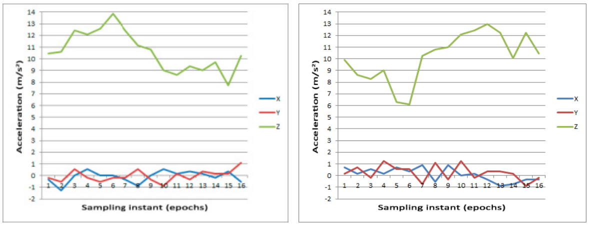


Fig. 6. Left: example of an upward movement. Right: example of a downward movement.

The distinction of right and left movements is verified on the X-axis. If the motion is done to the right, there is also an absolute maximum and an absolute minimum. Inversely, there is an absolute minimum followed by an absolute maximum on the left movement (Fig. 7).

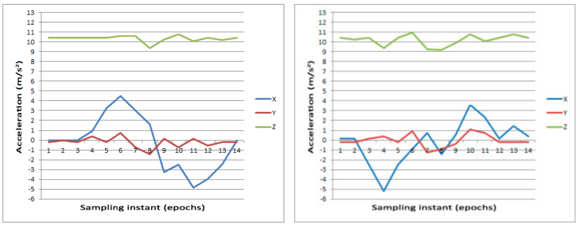


Fig. 7. Left: example of a movement to right. Right: example of a movement to left.

The Y-axis is the axis that varies the most on front and back movements (Fig. 8). During frontward movements, the absolute maximum occurs before the absolute minimum. During the backward movements, it is the exact opposite.

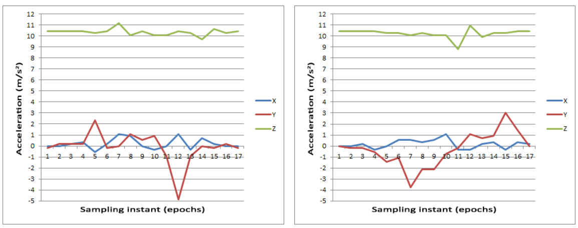


Fig. 8. Left: example of a frontward movement. Right: example of a backward movement.

After a first analysis of the upward, downward, right, left, frontward and backward movements with the smartphone in a lying down position (Fig. 5A), the same was done for the remaining positions depicted in figures 5B, 5C and 5D.

Among the positions of the smartphone: upright or lying down (Fig. 5), it is possible to have the smartphone in an inverted position, that is the smartphone might be in a position of 0º or 90º and 180º or 270º. To identify 6 distinct movements (up, down, right, left, front and back) it is necessary to detect the approximate initial position (this will be one among 8 possible positions that result from the combination of 4 departure angles and 2 positions of the smartphone - upright or lying down) obtained by the detection of the axis where the force of Earth's gravity is being exerted with higher intensity. The smartphone orientation sensor was used returning one of 4 integer values according to the positions 0º, 90º, 180º or 270º.

3.2 Determination of the Intensity of the Movement

The intensity of the movement is dependent on the user and the context of the completion of the movement. For example, the strength of a movement differs according to the age of the user, usually the intensity of a movement performed by a younger person is different from that applied by an older one and it depends on the purpose of the movement.

Therefore, we studied what might be a strong, normal and weak movement. The smartphone was placed in an upright position and on portrait mode (Fig. 5C) and an upward movement was made.

Fig. 9A represents a strong movement that reached a maximum value of 23.04 and a minimum value of -22.86 on Y axis. These findings emphasize the difference between the maximum and minimum values reached during the data capture, which only accumulated 12 values.

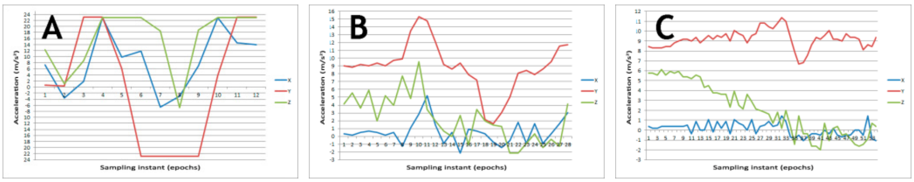


Fig. 9. A: Upward movement with the smartphone in an upright position and with strong intensity. Maximum value on the Y axis of 23.04 and minimum value on the Y axis of -22.86. B: Upward movement with the smartphone in an upright position, with normal intensity. Maximum value on the Y axis is 15.30 and minimum value on the Y axis is 1.62. C: Upward movement with the smartphone in an upright position with weak intensity. Maximum value on the Y axis is 11.69 and minimum value is 6.66 on the Y axis.

In the representation of a normal movement (Fig. 9B) by doing the same movement, the time of execution of the movement is superior than the strong movement, that is, if the movement is not as strong, the number of captured values on this movement will be higher. In this case, this process went from 12 values captured from a strong movement to 28 values captured from a normal movement. In relation to the values of maximum and minimum acceleration achieved on Y-axis, 15.30 and 1.62, there is a sharp decrease in the maximum value, but mostly a big difference in the minimum value reached. In this case it doesn't reach negative values, as in the stronger movements (Fig. 9A).

In the embodiment of the weak movements (Fig. 9C) the time is increased comparatively to the normal and strong movement and we obtained 53 values. If a movement is weaker, the time of its execution takes longer and there's less difference between the maximum and minimum values. In this particular case, the maximum value is 11.69 and the minimum is 6.66 in the Y-axis.

In conclusion, for different intensities it is apparent that if the movement is strong, the time of execution is shorter and there is a greater difference between the maximum and minimum acceleration (Fig. 9A). In a movement with less intensity (Fig. 9B), that the amplitude of the interval between the maximum and minimum values is reduced and the execution time (values captured) increases. On the other hand, if the movement is even weaker, the time to achieve the same values increase and the difference between the maximum and minimum values is shortened as is shown in the graph of Fig. 9C.

We also performed tests with the norm of acceleration, $||\vec{a}|| = \sqrt{x^2 + y^2 + z^2}$ where x, y and z represent the values of the acceleration on axes X, Y and Z respectively. Analyzing the same movements by the standard acceleration, the ratio of direct proportionality between the standard and intensity of movements is clear. Due to the simplicity of calculation and the good results obtained with the exclusive use of the acceleration components that influenced the movement the most, we decided to use this method in determining the intensity of the movements.

4 Example of Application

The application developed encompasses the whole study and it was conceived for the operating system Android version 2.2 Froyo. It consists in counting the movements performed by the user, either by direct contact with the smartphone or indirectly, for example, with the smartphone coupled to a gym machine. The application allows the counting of repetitions and movements in different directions as: upward, downward, right, left, frontward and backward.

Taking into account that the intensity of movement may vary from user to user and depending on the context of their use, the user can choose the level of sensibility or calibrate the movement manually. In the latter case, the movement is recorded and will be counted as the only movement in this direction and with this intensity. At the end of usage, it is possible to save and share the total of repetitions and the time of the physical activity.

With the intention of evaluating the implementation in a practical context, a prototype was developed and submitted to testing on 30 subjects, users in the gymnasium Nova Academia Fitness e Spa in Covilhã, Portugal. The application was tested on a smartphone Sapo a5. The smartphone was placed together with the weights of a latissimus machine (gym machine for dorsal strengthening), as is illustrated in Fig. 10.

Fig. 10. Application, positioning smartphone and tested machine

The total number of repetitions performed by the sample (30 users) was 601, whereas the application recorded a total of 609 repetitions, obtaining an error rate of less than 3%, i.e. with an absolute error rate of 16 repetitions. Notice that the application did not undergo any adjustments of calibration for the different users.

This preliminary test assessed the performance of the application in a real context and will help improve the application, according to the functional requirements of future users. We obtained positive feedback both in terms of the results achieved with the application, as well as to its usefulness. It was possible to notice that applications of this kind can help increase the levels of motivation and concentration during exercise and can contribute to a better performance.

5 Conclusion

The smartphones are increasingly user-friendly and easy to use. Their utilization rate shows solid and gradual growth and its distribution covers different age groups. The rich linguistic description of the movements of a smartphone as well as the characterization of their intensity enhances the human-centric nature of the applications that can be developed in areas such as ambient assisted living, elderly care and mHealth. It was verified that the proposed movement identification system, developed for smartphones with accelerometer, has potential to be used in a wide range of applicational scenarios such as: games, to move objects like cars and characters; verifying sudden movements such as falls and counting movements in different directions, to name a few.

The application developed confirms that it can be used independently by individual users of different age groups, but also has good indicators in relation to the business world, especially in gymnasiums and physiotherapy centers. The result of this work is also promising as a means to calculate physical activity. In this sense, it can have an important role in fighting obesity. Another application area that we intend to explore is mobile health (mHealth) by developing applications that stimulate interaction and physical activity, by functioning as regulators, as well as promoting and encouraging the need for the practice of physical activity.

Acknowledgments. This work was partly supported by Fundação para a Ciência e Tecnologia (FCT) under the project PEst-OE/EEI/LA0008/2013. The authors wish to thank Dr. Aurélio Faria from the UBI Sports Science Department for his availability providing all the necessary equipment to the validation tests. A special acknowledgment is due to Dr. Albino Ferreira and to Dr. Filipe Ferreira for their valuable suggestions and ideas.

References

1. Blog Strategy Analytics. Worldwide Smartphone Population Tops 1 Billion in Q3 (october 2012), http://blogs.strategyanalytics.com/WDS/post/2012/10/17/Worldwide-Smartphone-Population-Tops-1-Billion-in-Q3-2012.aspx (retrieved from)
2. GSMA Intelligence. Smartphone users spending more 'face time' on apps than voice calls or web browsing (2011), https://gsmaintelligence.com/analysis/2011/03/smartphone-users-spending-more-face-time-on-apps-than-voice-calls-or-web-browsing/271 (retrieved from)
3. GigaOM. 4 ways mobile health could save $400B in health costs (February 2013), http://gigaom.com/2013/02/25/4-ways-mobile-health-could-save-400b-in-health-costs/ (retrieved from)
4. Stankevich, E., Paramonov, I., Timofeev, I.: Mobile Phone Sensors in Health Applications. In: Proceeding of the 12th Conference of FRUCT Association, pp. 136–141 (April 2012)
5. Sposaro, F., Tyson, G.: iFall: an Android application for fall monitoring and response. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6119–6122 (2009)
6. Lockman, J., Fisher, R., Olson, D.: Detection of seizure-like movements using a wrist accelerometer. Epilepsy & Behavior 20(4), 638–641 (2011)
7. Parise, B., Delano, R.: iTrem, http://eosl.gtri.gatech.edu/Capabilities/LandmarcResearchCenter/LandmarcProjects/iTrem/tabid/798/Default.aspx (retrieved from)
8. STMicroelectronics. LIS302DL (October 2008), http://www.st.com/st-web-ui/static/active/en/resource/technical/document/datasheet/CD00135460.pdf (retrieved from)
9. BIOPAC Systems, Inc. Tri-Axial Accelerometers (November 2013), http://www.biopac.com/Product_Spec_PDF/Accelerometers.pdf (retrieved from)

Intelligent Wheelchair Driving: A Comparative Study of Cerebral Palsy Adults with Distinct Boccia Experience

Brígida Mónica Faria[1,2,3], Ana Silva[1], Joaquim Faias[1], Luís Paulo Reis[3,4], and Nuno Lau[2,5]

[1] ESTSP/IPP – Esc. Sup. de Tecnologia da Saúde do Porto / Inst. Polit. do Porto, Portugal
[2] IEETA – Instituto de Engenharia Electrónica e Telemática de Aveiro, Aveiro, Portugal
[3] LIACC – Lab. Inteligencia Artificial e Ciência de Computadores, Porto, Portugal
[4] EEUM/DSI - Escola de Engenharia da Universidade do Minho, Guimarães, Portugal
[5] DETI/UA – Dep. Eletrónica, Telecomunicações e Informática, Univ. Aveiro, Portugal
{btf,jfaias}@estsp.ipp.pt, silva.anato@gmail.com, lpreis@dsi.uminho.pt, nunolau@ua.pt

Abstract. An electronic wheelchair facilitates the autonomy and independence of a person, however specific cognitive, sensorial and perceptual skills are needed to conduct the assistive technology. These skills are also inherent to the sport boccia. Thus, the aim of this study is to understand the relationship between the experience of the participant in driving a wheelchair in relation to their autonomy and independence and also examine the practice of boccia in relation to the cognitive skills and performance in driving an intelligent wheelchair using a simulator. It was performed an evaluation of 28 participants, 6 of whom had no experience driving an electronic wheelchair and 22 had experience, 15 practice boccia and 13 did not practice this type of adapted sports. In the collection of data was tested three interfaces command of a smart wheelchair in a simulator. It was showed a good performance of the participants with experience in using electronic wheelchair and practitioners of boccia. It was also possible to observe that the autonomous and independent participants showed good results.

Keywords: Intelligent Wheelchair, Simulator, Boccia, Cerebral Palsy.

1 Introduction

Cerebral palsy has been defined as a group of permanent developmental disorders of movement and posture, causing limitations in terms of activity due to a non-progressive disorder that occurs in the brain during fetal and childhood [1] development.

The motor disorders in cerebral palsy are often accompanied by sensory deficits as well as deficits in perception, cognition, communication and behavior, also with epilepsy and secondary musculoskeletal problems [1].

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_32, © Springer International Publishing Switzerland 2014

Studies of adults with cerebral palsy show that the main limitations in daily activities are present in the areas of self-care and mobility [2] [3] [4].

It also appeared in a study by Nieuwenhuijsen et al. [5], that the problems in mobility are reported frequently and are extremely important to overcome by young adults with cerebral palsy.

Andersson and Mattson [6] showed that nearly half of adults with cerebral palsy without learning difficulties could not move independently and autonomously in the community as desired or as needed. Being independent is having the physical ability to perform personal and social tasks. Autonomy is reflected by the freedom to determine their own actions or behaviors [7]. Therefore, independent mobility is vital for activities and participation, thus reducing dependence on caregivers and the environment adaptation [8] [9] [10].

1.1 Assistive Technologies

Assistive technologies are defined as any product, instrument, equipment or adapted technology specially designed to improve the functional levels of the disabled person. Resorting to these products can help reduce the limitations in mobility [8] [10] [11]. The electronic wheelchair wheels are an enabling technology, used by people, due to a wide range of diseases including cerebral palsy [12].

Since users of such assistive technologies are increasing, manufacturers have been developing various kinds of wheelchairs with different functional characteristics, as Cooper [13] and Edlich [14].

To use an electronic wheelchair, cognitive and physical skills, which not all people possess, are required. In a study by Fehr et al. [15] demonstrates that between 10 and 40% of people who wished to use the electronic wheelchair with the existing interfaces could not do so due to sensory deficits, poor motor function or cognitive deficits making it impossible for safe driving.

However, to accommodate users who are of the opinion that the performance of mobility equipment standard is difficult or impossible, some investigators have used technology that was originally used for the mobility of robots aiming to develop a wheelchair with some level of "intelligence". These assistive technologies are designed to help users move in different ways, also ensuring mobility without collisions, helping the performance of specific tasks, like opening the door and independently transporting the user between locations [16].

In a study by Montesano et al. [17], with a sample of people with cognitive deficits due to cerebral palsy, it was concluded that these users after a training phase were able to command the wheelchair smart wheels safely, even in the most difficult situations.

Children with cerebral palsy, with no previous experience of using a wheelchair showed a marked improvement in their performance after the training phase with the simulator or with serious games [18][19][20][21][22] and is therefore the key to prepare children for powered mobility simulator [23][24]. Nevertheless, the performance of children who already had experience in using the wheelchair was substantially better [23].

1.2 Boccia Sport

One of the most mentioned activities in which young people with Cerebral Palsy are involved are adapted sports activities. In accordance with the information of Cerebral Palsy - International Sport and Recreation Association (CP - ISRA), one of the most popular activities, both nationally and internationally, in adapted sport is Boccia [25]. In Portugal, there are more than 300 practitioners of this sport [26], ranked by Gross Motor Function Measure (GMFM), as having severe disability [27]. This sport is directed to people with disabilities, specifically for cerebral palsy and muscular dystrophy, but is also currently practiced by the elderly [28] [29].

The purpose of this sport is to put the colored balls together to the target ball and prevent the opponent does. The game consists of thirteen balls, six red, six blue and white, this being the target ball. To perform the release, the athlete can use the more functional end of the body [30] [31].

This sport is adapted to people with physical or functional disability, with serious problems of coordination and motor control. It is also considered a game with technical accuracy and strategy, which emphasizes the skills and technical capabilities. The Boccia involves a lot of precision, coordination, attention, concentration, perceptual-motor, especially at the level of visual-motor coordination [31] capabilities. These skills are also mobilized in the action of driving a wheelchair [32].

The aim of this study was to understand the factors inherent to a user, which influence their performance in driving an intelligent wheelchair using a simulator. There were other specific objectives, such as, to verify how the performance in driving the intelligent wheelchair in the simulator depends on the user experience of driving the usual wheelchair, especially in terms of their independence and autonomy.

The second specific aim of this study is to examine whether there are differences in performance outcomes in a course in driving simulator chair “smart” (the level of cognitive skills), among practitioners Boccia wheels for over a year and people with cerebral palsy never practiced this kind of adapted sports.

2 Methodology

This section refers to the methods used for the development of the comparison of the performance in driving the intelligent wheelchair between the cerebral palsy adults with and without experience in playing Boccia.

2.1 Study Design

To materialize this study a quasi-experimental quantitative study was developed. It were included two control groups, the control group composed with adults without experience in playing boccia and a group of adults without experience in driving a wheelchair. This study was also cross-sectional the data was collected at a define time [33].

2.2 Participants

The sample was composed by clients of a Cerebral Palsy Institution. It was prepared a questionnaire with the inclusion criteria and possible interfaces for driving the intelligent wheelchair. Since the elements were selected based on their reasonable ability for at least one interface, according to therapist's assessment the sample is designed as a convenience sample.

For this study two groups were constituted. In Group 1 were defined as inclusion criteria: diagnosis of cerebral palsy, aged 16 years or higher, being classified as a level IV or V according to the GMFM and boccia practitioners for at least a year. In Group 2 all criteria were applied, except the last, so participants never practiced boccia.

2.3 Instruments and Procedures

In this study, a user profile application was used. Several small tasks should be accomplished through the available interfaces [34-37] in order to verify which interfaces could be tested by the cerebral palsy users.

The first task was button sequences were several buttons should be pressed. In the second task the participant was asked to repeat the phrases “go forward”, “go back”, “turn right”, “turn left”, “right spin”, “left spin” and “stop”, using voice sequences. In the third task it was asked to control the placement of a red circle on a green circle, by manipulating the joystick. When the red circle overlaps the green circle for two seconds, automatically passes to a next level, since there were three different positions. The last interface tested requires the use of head movements and the goal is the same as the previous interface, i.e., put the red circle on the green circle for two seconds.

Fig. 1. Tasks examples using the user profiling

The simulator is displayed in a LCD (42 inches) and represents the interior spaces of the Cerebral Palsy Institution. Over these spaces in the simulator, 14 balls were placed well visible, to define the points where the participants had to pass. When passing the ball with the virtual intelligent wheelchair, the simulator produces an audible feedback and the object disappears. At the end of the route a different object (a star) should be collected and it is also represented by a distinctive sound. The participant can get 15 points at the end of the route, one point for each ball and a point by the star.

After the completion of the circuit the participants were asked to fill out a questionnaire organized with the demographic data as well as information about the motor, cognitive and sensory characteristics. The questionnaire also had other objectives: to gather information about the users' experience using information and communications technologies; the usability about the intelligent wheelchair and safety of intelligent wheelchair in simulation mode; the level of satisfaction with the type of control of intelligent wheelchair; fatigue and/or frustration to use the interface; the level of difficulty of use of the interface and the degree of preference for each control interface.

The data about the performed route and time spent was accessed and the Statistical Package for Social Sciences [38] version 19.0 was the tool used to create the database and to perform the statistical analysis. For the sample characterization the measures of central tendency (mean, mode, median) and dispersion (standard deviation) were calculated. It was applied the Kolmogorov Smirnov test for analyze the normality of the data and the Mann-Whitney was applied to verify the differences between the independent groups in terms of performance (time and score achieved). The Chi-square test was applied to test the association between variables. The significance level was established with the value 0.05.

3 Results

The results are organized with the information and characterization of the sample and the outcomes for the purposes of this study.

The sample consisted of a total of 28 participants 8 were female and the remaining 20 were male. The age range of the sample varies between 18 and 47 years old, with an average of 28 years old and with a standard deviation of 7.69. All participants have been diagnosed with cerebral palsy, 6 classified with the level IV and 22 with the level V of the GMFM. Regarding to the education level, 2 of the participants are illiterate, 5 completed the elementary school, 7 the middle school, 9 the high school, 1 have a BSc and 4 did not answered.

Table 1. Sample characterization about constraints, autonomy and independency

Variables		Frequency	Variables		Frequency
Use manual wheelchair			Cognitive constraints		
	no	17		no	10
	yes	11		yes	15
Use electric wheelchair			Motor constraints		
	no	6		no	0
	yes	22		yes	28
Autonomy using wheelchair			Visual constraints		
	no	7		no	14
	yes	21		yes	13
Independence using wheelchair			Auditive constraints		
	no	6		no	28
	yes	22		yes	0

Regarding dominance, 8 participants are righty and 15 lefty, 5 participants did not answer this question since the control was done with the head.

Table 1 presents the characteristics of the sample according deficits of the participants how to use the wheelchair and the autonomy and independence.

Besides this distribution, 6 participants are not independent or autonomous, 1 participant is independent but not autonomous and 21 participants are independent and autonomous.

The experience in driving an electronic wheelchair and the autonomy and independence of the participant are significant related. In fact, and applying the Fisher's exact test it was obtained a p value of 0.01. The Table 2 shows the distribution of data related with the variable autonomy and independence and the variable experience in using an electronic wheelchair.

Table 2. Experience in driving an electronic wheelchair and autonomy and independency

Experience in driving an Electronic Wheelchair	Yes	No	Total
Autonomous and Independent	4	2	6
Neither Autonomous or Independent	2	19	21
Independent but not Autonomous	0	1	1
Total	6	22	28

The association between the cognitive constraints and the boccia's practitioners (users that responded) is not significant (p value of the Fisher's exact test is 0.226). The Table 3 presents the distribution of these variables.

Table 3. Boccia Practitioners and cognitive constraints

Boccia Practitioners	Yes	No	Total
Presents cognitive constraints	6	9	15
Do not presents cognitive constraints	7	3	10
Total	13	12	25

After the profiling session with was possible to access which interface could be used for driving the intelligent wheelchair using the simulator. Table 4 shows the results.

Table 4. Interfaces distribution by users

	Yes	No
Joystick	18	10
Head Movements with wiimote	27	1
Voice Commands with microphone	4	24

Table 5 presents the analysis of the performance of the participants in the simulator with and without experience in driving electronic wheelchairs.

Table 5. Users with and without experience with wheelchairs - results using joystick

Driving the Intelligent Wheelchair with Joystick										
		Time (min)			Collisions			Score		
Experience	**n**	aver	md	std	aver	md	std	aver	md	std
With	**13**	11.8	9.5	5.4	6.9	6	5.5	12.9	14	3.7
Without	**5**	15.3	8.8	15.3	17.2	6	22.9	9.4	9	5.0
P value		0.775			0.428			0.209		

Applying the Mann-Whitney test it was possible to verify that there are not statistical evidences to affirm that there are differences between the two groups (with and without experience in driving an electronic wheelchair) related with the time, collisions and score in the serious game by driving an intelligent wheelchair with joystick. Nevertheless, it is possible to observe better results in the group with experience using an electronic wheelchair.

The same overall results were obtained when using the head movements. Table 6 shows the results about the performance in terms of time, number of collisions and score.

Table 6. Users with and without experience with wheelchairs - results using head movements

Driving the Intelligent Wheelchair with Head Movements										
		Time (min)			Collisions			Score		
Experience	**n**	aver	md	std	aver	md	std	aver	md	std
With	**22**	16.3	15.0	7.4	8.5	6	5.7	13.8	15	2.6
Without	**5**	18.2	18.7	8.0	7.6	7	4.0	12.8	15	4.4
P value		0.605			0.940			0.902		

The next results are related with the performance of users divided in two groups: the practitioners of Boccia and the users without experience in playing Boccia. Table 7 and 8 present the results of these groups using the joystick and the head movements to drive the intelligent wheelchair.

Table 7. Users with and without experience in paying Boccia – results using joystick

Driving the Intelligent Wheelchair with Joystick										
		Time (min)			Collisions			Score		
Boccia Player	**n**	aver	md	std	aver	md	std	aver	md	std
yes	**9**	11.1	7.6	6.4	6.3	6	5.2	13.2	13.2	3.6
no	**9**	14.3	9.7	11.0	13.2	7	17.3	10.6	10.6	4.6
P value		0.222			0.373			0.142		

Table 8. Users with and without experience in paying Boccia – results using head movements

Driving the Intelligent Wheelchair with Head Movements										
		Time (min)			Collisions			Score		
Boccia Player	**n**	aver	md	std	aver	md	std	aver	md	std
yes	**15**	16.7	15.7	8.4	7.4	6	5.2	13.9	15	2.6
no	**12**	16.6	17.6	5.3	9.3	7.5	5.7	13.3	14	3.3
P value		0.952			0.384			0.263		

There are not statistical evidences to affirm that there are significant differences between the group of players and no players in terms of time, collisions and score.

4 Discussion

After the experiments it was found that participants with experience in driving a wheelchair achieved better results with all interfaces when analyzing the median value. It was interesting to observe that the group with experience in driving an electric wheelchair could perform the experiment without many errors and if the participants had a period of training in driving the intelligent wheelchair in the simulator, would be expected almost perfect results in the group of participants with experience. It also will be expected an improvement of the results in the group of participants with no experience in driving a wheelchair as Hasdai et al. [23] also have concluded. It was possible to verify that the experience in driving an electronic wheelchair and the autonomy and independence of the participant are significant related.

For practitioners of boccia it was found a better performance in the simulator in comparison to the group that do not have experience in playing this sport although not very different. Participants with cognitive deficits have not completed the route because of the deficits in attention, concentration and the considerable extension of the circuit. Even so, from the total of 11 participants with cognitive deficits that began driving the intelligent wheelchair with the joystick, 7 completed the route and from 14 that used head movements, 12 managed to finish the route. The voice command interface had a sample of 2 participants who completed the route and both are practitioners of boccia. The time difference between the two participants is due to the fact that the participant females have better diction of the verbalization of the words and the system quickly recognized the commands. The three interfaces available for use in the simulator, which had more participants to complete the course was with the head movements possible with the wiimote. Although at the beginning of the session the participants show up apprehensive regarding this interface thinking that were not capable to control the wheelchair with only the head movements, at the end of the experience most participants considered the wiimote easier to use than originally thought.

In short and analyzing the few results in the literature [31] about the skills of young adults with cerebral palsy, in playing boccia and the influence in using an electric wheelchair it is possible to relate this with the results here presented. Nevertheless it will be required a greater sample in the group of participants without experience, to see if this trend is maintained and has the same results as Hasdai et al. [23]. Moreover other possibilities for driving the wheelchair should be implemented and analyzed [39][40][41].

5 Conclusions and Future Work

The simulator used for collecting data for this study allowed us to assess the performance of participants with different interfaces: joystick, head movements and voice commands. The simulator allowed several cerebral palsy adults to have actions in an autonomous way, making them very motivated by their performance. The head movements would be a good interface to be adapted for driving an intelligent wheelchair and for future could be used to access other kind of devices, such as a computer.

This study encountered some limitations, including the fact that there are few references related to these topics. Also the number of participants should be increased in order to have more reliability in the statistical evidence. For a further study is suggested to verify the performance of the participants, after a training period for driving the intelligent wheelchair in the simulator. Also it is suggested to verify the performance of participants with neuromuscular diseases, to determine which is the best interface for this type of population, this would require taking into account the level of functionality that the participant had.

Therefore, the simulator used in this study, along with similar to those used in electronic wheelchairs interfaces, evaluated in a protected environment the skills of clients. It also makes it possible to assess what assistances still needed to develop for driving an electronic wheelchair in real context. This equipment, in simulation mode, it would be a good training tool, since it enabled experiences in a safe manner. As an intervention tool, the simulator also helps to develop cognitive and sensorimotor abilities necessary for conducting electronic wheelchair. Thus, after a suitable period of intervention, the client may conduct an electronic wheelchair independently and autonomously, improving their quality of life by allowing access to other occupational forms.

References

1. Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M., Bax, M., Damiano, D., et al.: A report: the definition and classification of cerebral palsy April 2006. Dev. Med. Child Neurol. Suppl. 109, 8–14 (2007)
2. Andren, E., Grimby, G.: Dependence and perceived difficulty in activities of daily living in adults with cerebral palsy and spina bifida. Disability & Rehabilitation 22(7), 299–307 (2000)

3. Andren, E., Grimby, G.: Dependence in daily activities and life satisfaction in adult subjects with cerebral palsy or spina bifida: A follow-up study. Disability & Rehabilitation 26(9), 528–536 (2004)
4. Bottos, M., Feliciangeli, A., Sciuto, L., Gericke, C., Vianello, A.: Functional status of adults with cerebral palsy and implications for treatment of children. Developmental Medicine & Child Neurology 43(8), 516–528 (2001)
5. Nieuwenhuijsen, C., Donkervoort, M., Nieuwstraten, W., Stam, H.J., Roebroeck, M.E.: Transition Research Group South West N. Experienced problems of young adults with cerebral palsy: targets for rehabilitation care. Arch. Phys. Med. Rehab. 90(11), 1891–1897 (2009)
6. Andersson, C., Mattsson, E.: Adults with cerebral palsy: a survey describing problems, needs, and resources, with special emphasis on locomotion. Dev. Med. Child. Neurol. 43(2), 76–82 (2001)
7. Haak, M., Fänge, A., Iwarsson, S., Ivanoff, S.D.: Home as a signification of independence and autonomy: experiences among very old Swedish people. Scandinavian Journal of Occupational Therapy 14(1), 16–24 (2007)
8. Tefft, D., Guerette, P., Furumasu, J.: Cognitive predictors of young children's readiness for powered mobility. Developmental Medicine And Child Neurology 41(10), 665–670 (1999)
9. Fernandes, T.: Independent mobility for children with disabilities. International Journal of Therapy & Rehabilitation 13(7), 329 (2006)
10. Palisano, R.J., Tieman, B.L., Walter, S.D., Bartlett, D.J., Rosenbaum, P.L., Russell, D., et al.: Effect of environmental setting on mobility methods of children with cerebral palsy. Developmental Medicine & Child Neurology 45(2), 113–120 (2003)
11. Wiart, L., Darrah, J.: Changing philosophical perspectives on the management of children with physical disabilities—their effect on the use of powered mobility. Disability & Rehabilitation 24(9), 492–498 (2002)
12. Laffont, I., Guillon, B., Fermanian, C., Pouillot, S., Even-Schneider, A., Boyer, F., et al.: Evaluation of a stair-climbing power wheelchair in 25 people with tetraplegia. Arch. Phys. Med. Rehabil. 89(10), 1958–1964 (2008)
13. Cooper, R.A.: Engineering manual and electric powered wheelchairs. Critical Reviews in Biomedical Engineering 27(1-2), 27–73 (1999)
14. Edlich, R.F., Nelson, K.P., Foley, M.L., Buschbacher, R.M., Long, W.B., Ma, E.K.: Technological advances in powered wheelchairs. Journal of Long-Term Effects of Medical Implants 14(2), 107–130 (2004)
15. Fehr, L., Langbein, W.E., Skaar, S.B.: Adequacy of power wheelchair control interfaces for persons with severe disabilities: a clinical survey. Journal of Rehabilitation Research & Development 37(3), 353–360 (2000)
16. Simpson, R.C.: Smart wheelchairs: A literature review. J. Rehabil. Res. Dev. 42(4), 423–436 (2005)
17. Montesano, L., Diaz, M., Bhaskar, S., Minguez, J.: Towards an intelligent wheelchair system for users with cerebral palsy. IEEE Trans. Neural Syst. Rehabil. Eng. 18(2), 193–202 (2010)
18. Rego, P., Moreira, P.M., Reis, L.P.: Architecture for Serious Games in Health Rehabilitation. In: World Conference on Information Systems and Technologies, WorldCIST 2014, Madeira, Portugal, April 15-18 (to appear, 2014)
19. Mendes, L., Dores, A.R., Rego, P., Moreira, P.M., Barbosa, F., Reis, L.P., et al.: Virtual Centre for the Rehabilitation of Road Accident Victims (VICERAVI). In: CISTI 2012, Iberian Conf. Information Systems and Technologies, vol. 1, pp. 817–822 (June 2012)

20. Rego, P., Moreira, P.M., Reis, L.P.: New Forms of Interaction in Serious Games for Rehabilitation. In: Cruz-Cunha, M.M. (ed.) Handbook of Research on Serious Games as Educational, Business, and Research Tools: Development and Design, pp. 1188–1211. IGI Global, Hershey (2012)
21. Rego, P., Moreira, P.M., Reis, L.P.: Natural user interfaces in serious games for rehabilitation. In: Proc. 6th Iberian Conference on Information, Systems and Technologies, CISTI 2011, vol. II, pp. 229–232 (June 2011)
22. Rego, P., Moreira, P.M., Reis, L.P.: Serious Games for Rehabilitation: A Survey and a Classification Towards a Taxonomy. In: Rocha, A., Sexto, C.F., Reis, L.P., Cota, M.P. (eds.) Proc. 5th Iberian Conf. Information Systems and Technologies, CISTI 2010, S. Compostela, Spain, June 16-19, pp. 349–354 (2010)
23. Hasdai, A., Jessel, A.S., Weiss, P.L.: Use of a computer simulator for training children with disabilities in the operation of a powered wheelchair. Am. J. Occup. Ther. 52(3), 215–220 (1998)
24. Faria, B.M., Teixeira, S.C., Faias, J., Reis, L.P., Lau, N.: Intelligent Wheelchair Simulator for Users' Training: Cerebral Palsy Children's Case Study. In: 8th Iberian Conf. on Information Systems and Technologies, Lisbon, Portugal, June 19-22, vol. I, pp. 510–515 (2013)
25. About Boccia Cerebral Palsy Internacional Sport and Recreation Association, `http://www.cpisra.org.za/index2815.html?id=80` (cited March 6, 2012)
26. Ranking Nacional de Boccia 2012 PC – Ass. Nacional de Desporto 2012 (2012), `http://www.pcand.pt/sites/default/files/documentos/Ranking%20Nacional%20de%20Boccia%202012.pdf.` (cited june 29, 2012)
27. Rosenbaum, P.L., Walter, S.D., Hanna, S.E., Palisano, R.J., Russell, D.J., Raina, P., et al.: Prognosis for Gross Motor Function in Cerebral Palsy Creation of Motor Development Curves. JAMA: J. American Medical Ass. 288(11), 1357–1363 (2002)
28. Oterelo, H.: A Importância Diferenciada dos Lançamentos na Discriminação da Vitória e da Derrota nos Parciais do Jogo de Boccia. Universidade do Porto, Porto (2006)
29. Soares, A.: A Utilização das Competências Psicológicas pelos Atletas de Boccia. Universidade do Porto, Porto (2007)
30. Boccia, M.L.: Estudo sobre o Estado de Conhecimento na Modalidade. Universidade do Porto, Porto (1998)
31. Leite, V.: As Competências Psicológicas no Desporto: Estudo com Atletas de Boccia. Universidade do Porto, Porto (2007)
32. Simpson, R.C., LoPresti, E.F., Cooper, R.A.: How many people would benefit from a smart wheelchair? J Rehabil Res Dev. J. Rehabil. Res. Dev. 45(1), 53–71 (2008)
33. Kielhofner, G.: Research in Occupational Therapy. F. A. Davis Company, Filadélfia (2006)
34. Faria, B.M., Reis, L.P., Lau, N.: Knowledge Discovery and Multimodal Inputs for Driving an Intelligent Wheelchair. Int. J. Know. Discov. Bioinformatics 2(4), 18–34 (2011)
35. Faria, B.M., Reis, L.P., Lau, N., Soares, J.C., Vasconcelos, S.: Patient Classification and Automatic Configuration of an Intelligent Wheelchair. In: Filipe, J., Fred, A. (eds.) ICAART 2012. CCIS, vol. 358, pp. 268–282. Springer, Heidelberg (2013)
36. Faria, B.M., Vasconcelos, S., Reis, L.P., Lau, N.: A Methodology for Creating Intelligent Wheelchair Users' Profiles. In: Filipe, J., Fred, A. (eds.) ICAART 2012. CCIS, vol. 358, pp. 171–179. Springer, Heidelberg (2013)

37. Faria, B.M., Vasconcelos, S., Reis, L.P., Lau, N.: Evaluation of Distinct Input Methods of an Intelligent Wheelchair in Simulated and Real Environments: A Performance and Usability Study. Assistive Technology: The Official Journal of RESNA 25(2), 88–98 (2013)
38. Maroco, J.: Análise Estatística com utilização do SPSS. Síbalo, Lisboa (2003)
39. Trigueiros, P., Ribeiro, F., Reis, L.P.: A comparative study of different image features for hand gesture ML. In: 5th Int. Conf. Agents Art. Int., ICAART 2013, pp. 51–61 (2013)
40. Trigueiros, P., Ribeiro, F., Reis, L.P.: Vision Based Referee Sign Language Recognition System for the RoboCup MSL League. In: 17th RoboCup International Symposium on Robocup 2013. Springer, Eindhoven (July 1, 2013)
41. Trigueiros, P., Ribeiro, F., Reis, L.P.: A comparison of machine learning algorithms applied to hand gesture recognition. In: Iberian Conf. Information Systems and Technologies, CISTI 2012, vol. 1, pp. 41–46 (2012)

The Process-Oriented Implications in the Development of Health Informatics Curriculum for Teaching and Learning

Suama Hamunyela and Tiko Iyamu

Namibia University of Science and Technology, Department of Business Informatics, Windhoek, Namibia
{slhamunyela,tiyamu}@polytechnic.edu.na

Abstract. The healthcare environment is complex and delicate in terms of its activities and processes. Some of the activities include storage, retrieval, use, and management of individual and group patient's record and clinical data. As such, healthcare informatician needs to be equipped with diverse knowledge and skills for better healthcare services deliver, planning and management.

This is attributed to the complexity of developing the curriculum. Also, the comprehension of the informatics curriculum is critical for accreditation purposes. Thus, this research paper focuses on the development of curriculum at institutions of higher learning, to examine and understand the implications of the processes that are involved in health informatics, for academic purposes. The curriculum is purposely to promote, practice and transfer of technological knowledge and skills on health informatics.

This is an applied research that was carried out at the Namibia University of Science and Technology. The study employed different approaches in the design, and collection of empirical data. Based on the interactive nature amongst stakeholders, the interpretive method was employed in the data analysis, to examine how the health informatics curriculum is developed for undergraduate level programmes.

Keywords: health informatics, technology, process, people, curriculum development, interpretive.

1 Introduction

Health informatics is an emerging field of study [1], designed to cater for the educational needs in health environment, for various healthcare professionals, and applied in different relevance and contexts. The health informaticians (HI) are professionals that manage healthcare data and information resources to provide relevant knowledge, to relevant healthcare sector and stakeholders. The profession encompasses planning, collecting, aggregating, analysing and disseminating of individual patient and aggregate clinical data. In some arguments, healthcare informaticians needs to be equipped with diverse knowledge and skills for better healthcare services deliver and planning [2].

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_33, © Springer International Publishing Switzerland 2014

This paper provides insight on the development of health informatics curriculum at the Polytechnic of Namibia which recently transformed into Namibia University of Science and Technology (NUST). Health informatics is an interdisciplinary of health science and information systems. The health informatics curriculum was based on premises, such as the followings:

a) As mandated by its constitution, the NUST develops instructional programmes that meet the needs of the Namibian industry, internationally recognised field of study, as well as, to promote practice and transfer technological knowledge and skills [3]. The health informatics is within the scope of the University's School of Information Technology (SIT). The development of the curriculum therefore followed a process as designed by the University.
b) Informatics-based solutions are essential to solve data redundancy, data fragmentation and information sharing: Some of the health information systems (HIS) studies that were reviewed indicated that deployment of HIS depends on the competencies of the users [4], [5]. [6] Argued that HIS are needed to improve the process of data handling in order to extract useful information for health planning, decision making, and resources allocation.
c) To contribute to addressing the shortage of skills in the health informatics environment. This contributes to achieving the Country's vision of efficient healthcare delivery and automation of health processes by year 2030. It is thus, on the agenda of the Ministry of Health and Social Service (MoHSS) to embark on skill development for healthcare sector in the country.

In lights of the above rationale and motivation, there is the need for educational program that will produce qualified health informaticians in Namibia. The remainder of the paper is divided into literature review, methodology, development process, and results' implications. Finally, a conclusion was drawn.

2 Literature Review

Academic institutions strive to provide service to the general society through its course offerings. The course offerings are also driven by competitiveness amongst the institutions. The courses are guided by curriculum and syllabus. Curriculum development is based on requirements which are often derived from the past experiences, present experiential and futurist factors. According to [7], there are many drivers of curriculum development, most importantly the needs and desires of employers for educated people who have the skills and competencies that can help their organizations survive and succeed. As different field of interest emerge, curriculum is developed for its study purposes [1].

As did the current fields of studies, the emergence of Health Informatics requires curriculum to be developed [8]. Due to the newness of the field, many institutions of higher learning have not been able to integrate it into their programmes, for different reasons such as lack of resources to develop and implement the course.

Health Informatics is a combination of both information systems (IS) and health sciences disciplines. The IS and health sciences are distinct disciplines, which have separately continually undergone evolution and transformation over the years. However, it is clear that the needs of the healthcare industry and the ambitions of IS professionals coincide [9]. There is no clear indication of where IS ends, and where Health Sciences begins [2]. As such, it becomes problematic on which academic Department or Faculty should host the discipline. This could be attributed to the challenges that some Universities encounters in an attempt to develop the Health Informatics curriculum.

2.1 Organisations and the Need for Health Informatics

The term health information system (HIS) refers to the use of information with the support of information technology (IT) in healthcare sector, to optimise work flow processes, and to ultimately improve overall healthcare services [10]. Given the definition of HIS, the IT in the health sector can be seen as an enabler, proving efficient use of information in healthcare organisations such as hospitals, clinics, pharmacies etc. A significant use of IT in the health sector is the design and implementation of Telemedic systems and Hospital Information Systems.

HIS aim at helping health service providers in decision and action making, improves health service provision processes flexibility, access and effectiveness [11]. It is further argued that, the adoption of HIS in healthcare oganisation is expected to pave the way for better healthcare service [11]. A better healthcare service could improve relationship amongst stakeholders, through effective information flows. HIS has the potential to restructure, and revive healthcare service delivery processes and activities.

Because of the benefits that HIS has to offer in the healthcare sector, there have been a thriving numbers of healthcare organisations in developing and developed countries. These organisations can further be divided into two categories: a) onsite software developing organisations. In this category, the organisations develop software and applications that are necessary to addressing its business processes. b) The off the shelf product dependant. The later depends on the off the self products to carry out business activities. Consequently, and by implications, the organisations need IT professional with the knowledge and skill, to support the usage of technologies and systems. Also, it is argued that, it is very important for health informaticians to understand, and be able to facilitate the development of organisations' vision, in what a HIS can do or is required to do for the specific organisation [4].There is seem to be a desperate need for people who are able to develop, implement and use new tools to support and manage healthcare activities.

2.2 HI Curriculum Implications and the Context

Academic fields emerge when there is a body of specialized knowledge and practice that can be provided by an academic discipline [8]. The demands for education and training in health informatics go well beyond basic computer skills. Information literacy courses are designed to teach the basic information management skills, and the quantitative courses are designed to teach how to better analyze data. According to [9], these have become increasingly helpful in providing the knowledge that is needed for patient care delivery. In a *nutshell*, the health informatics field deals with more than just patient data. It also deals with the collection and analysis of pertinent healthcare data, to create information and knowledge hub, on which communities and health policy decision makers can act upon [9].

The Namibian Government has adopted an ICT policy, which is aimed to support and simplify work and processes, service delivery and interaction, between different stakeholders [12]. The ICT policy provide guidelines for an over-arching framework that allows provision of speedy, transparent, accountable, efficient and effective processes service delivery [13]. Government organisations, ministries and offices adopt the ICT policy and operate within the policy's mandate. In addition, the Ministry of Health and Social Services (MoHSS) in Namibia is mandated to manage and provide an integrated, affordable, accessible, quality health and social welfare services, which is responsive to the needs of the Namibian population [14]. Unfortunately, currently the ministry is experiencing a shortage of skilled employees especially in the information systems and data management sector [14]. This has huge set-back for the health services which the MoHSS and its subsidiaries provide to the nation.

The adoption of the ICT policy necessitates the need for healthcare informaticians. However, there are no formal programs in existence in Namibia that deal with this area on at academic level. As stated by [12], Namibia as a developing country is experiencing numerous challenges in managing health information systems. Some of the challenges, such as ICT skill are unique. There was a strong desire for health informatics specialists in the Namibian environment. The observed health systems, activities and data management of the healthcare sector provide the basis for this curriculum.This curriculum is aimed to provide skills and knowledge that are required, to meet the shortage of healthcare informatician in Namibia, and in other developing countries. In order to achieve the objectives of the study, a methodology which includes research approaches and methods was employed.

3 Methodology

The study employed the qualitative case study approach, mainly because it facilitates exploration of a phenomenon within a context, using a variety of data sources. Yin [15] argued that the case study approach should be considered when the "when", "how" and "why" questions are used to sought answers. Based on its constructivist paradigm, the case study approach can guide researcher to explore individuals or organizations, through complex interventions, relationships, communities, or programs [15], [16]. The NUST was used as the case, in the study. This was primarily because health informatics was been developed at the time of this study, so, we could follow the follow the process, empirically.

The NUST host about eleven thousand students at both undergraduate and postgraduate levels, within its seven academic Faculties, School of IT included. The School of IT is divided into Four Departments, among them was Informatics.

The Participatory, involving experiential, and documentation approaches were used in the data collection. The researcher attended and participated in 13 meetings and 5 workshops during the development of the curriculum. Access was granted to documents, outcomes from strategies (Advisory Board, Faculty Board, and Senate Committee), need analysis of report from feasibility enquiry, and the draft of the health informatics curriculum. Also, the data collection was participatory and experiential in that, the researchers were participated in the some of the committees which developed the Health Informatics curriculum in the University.

There was comparison exercise, which was part of the data collection. The benchmarking was carried out in parallel, to the data collection, at the case study.

Benchmarking or a comparative exercise as referred to in this paper was done with the purpose of ensuring internationalization of the health curriculum. There was a need and necessity to benchmark the curriculum against health informatics curriculum of other universities, in both developing and developed countries. Some of the universities were University of Kwazulu Natal, in South Africa, and University of Michigan; University of Toronto; Georgia State University; and University of Illinois, in North America.

The data was analysed, using the interpretive technique to gain better understanding of why things happened in the way that they did in the development of health informatics curriculum at the NUST. The interpretive approach was used to attempt to gain deeper understanding of phenomena being study through the meanings that people assign to them. According to [17], Interpretive methods of research in IS are "aimed at producing an *understanding* of the context of the information system, and the process whereby the information system influences and is influenced by the context".

This paper followed document review method to gather relevant information adopted in developing the health informatics curriculum. Developing a programme or syllabi at the intuitions follows process-oriented stages as shown in Figure 1. The figure outlines the stages that guide the curriculum development and app**ro**val process. As part of the curricula development method, the health informatics curriculum development workshops were carried out as follows:

Representatives of the School of Information Technology (SIT) held four meetings with stakeholders, which included IT experts who are employed in the industries. In addition, the Health Sciences department and the SIT management team to deliberate on the curriculum development. Each of the sessions lasted for approximately two hours, during which participants' had the opportunity to contribute their inputs, of knowledge and documentations.

4 Developing the Health Informatics Curriculum

The processes that were followed, and the data that was collected were analysed. Other materials that were also included in the analysis were the results from the benchmarking exercise that was conducted, as discussed in the methodology section.

Table 1 below presents the results from the analysis. The table explains the components and elements which were involved and considered in the development of health informatics course. As shown in Table 1, the development of health informatics curriculum entails the different entities, divided into nine categories.

Table 1. Components of Health Informatics Curriculum

Definition	**Internalisation**	**Externalisation**
The health informatics is an interdisciplinary field, which combines Computing and Health Science fields of studies. It is primarily concerns with the health cognitive, information processing, technology communication, and education. In addition, the field entails socio-technical, and research activities.	The health informatics curriculum development followed internal processes, which were rigorous nature. The processes were characterised by institutional committees, of subjects' expertise and Quality Assurance professionals. This was to ensure that graduates who obtain health informatics qualification are eligible for employment, and equivalent to international standard.	The curriculum processes involved external entities. These were to ensure appropriate context, relevance of the curriculum. The external entities, such as such as National Qualification Authority (NQA) agencies were key, as they were responsible for the approval and accreditation of the programme. This was to align with by International standard.
Course delivering strategy	**Undergraduates level**	**Post graduate level**
The course delivering strategies includes face-to face-contact on the part-time and full-time mode of study, at both undergraduate and postgraduate levels. This is mode of implementation of the curriculum enables, progressions into maturity stage.	The undergraduate level is the foundation, and building block into practice (professional) and advancement for postgraduate studies. This curriculum was relevant to undergraduate students: To undertake the health informatics course, students were required to have a professional background in IT or have completed a diploma in IT at the University, or any other institution with a similar curriculum (exemptions were done according to the University acceptance procedures).	There is focuses on research in all areas of health informatics. The research output is aimed at reviewing the curriculum for undergraduate's teaching and learning. Also, the research is used to review and understand better the professional activities, within context. Part of the post graduate learning is research and publications. Students enrolling at post graduate level are expected to research and publish in the field of health informatics.

Table 1. (*continued*)

Curriculum objectives	**Academic Faculty**	**Advisory Board**
The objectives of the health informatics curriculum were driven, primarily by academic goals and professional services.	Based on the strategic goal which includes competitiveness, the objectives of the health informatics course were formulated. This focused on both undergraduate and postgraduate, including short courses trainings and certification, for already in-service healthcare providers and IT personnel.	The advisory board provided and contributed experiential, real-life experiences, to guide the development of the health informatics course.

As shown in Table 1, the development and implement of health informatics curriculum is enabled and constraint by three main components, technology, people and process. Figure 1, illustrate the iterative and dependence of the components in the development and implementation of Health Informatics curriculum in institution of higher learning. The components depend on each other, and are inseparable in the development and implementation of Health Informatics curriculum.

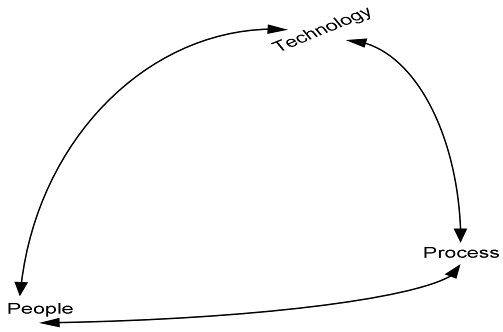


Fig. 1. Health Informatics elements

Technology **-** Technology generally impacts academic, the society and organisation processes positively or negatively. Positive impacts emerge when the technology adopted in an organisation supports its processes and when the people are effectively making use of the technology. Negativity surfaces when there is a divergence between people, process and technology. This interactive relationship between the three components instigates inclusion in curriculum development in higher learning institutions where curricula are designed for people in accordance with institutional processes and mandates.

Technology in health informatics can be grouped into two categories, the infrastructure and the informatics:

i. Healthcare Infrastructure ranges from the hardware and applications including internets, intranet and extranets that run healthcare data and services applicable to the organisation. Infrastructures enable the people to process and share healthcare data.
ii. Informatics in this case refers to the computing of health data and guidelines.

Process - The process component in the development of Health Informatics is considered to be critical. This is attributed to some of its roles which include validation of context, and quality assessment. Due to the nature of the roles, the processes as discussed below are sequential and dependent on each other, as shown in Figure 1.3:

Based on the interpretation of table 1, a process-oriented framework as shown in Figure 2, was developed. Figure 1 illustrates the process-oriented of the activities that were involved in the development of health informatics curriculum at the Namibia University of Science and Technology.

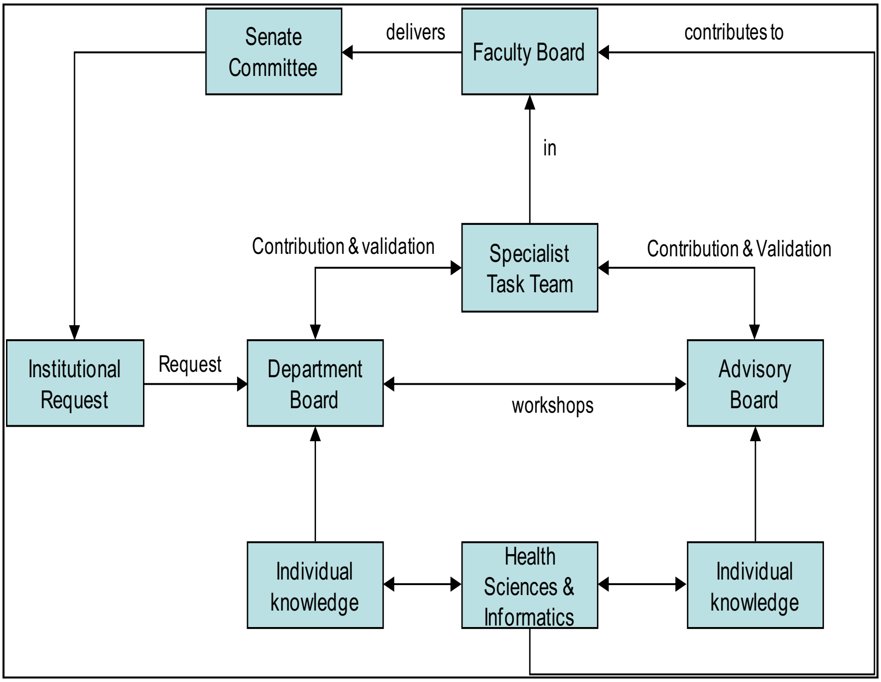


Fig. 2. Health Informatics Process-oriented

The components as shown in Figure 2 are briefly described as follows:

i. Specialist Task Force investigates the relevance of the curriculum, needs of the students and benchmark the curriculum.
ii. Departmental Committee collaborates with the specialist task force in selecting courses and providing course details.

iii. Faculty (SIT) Board is responsible of researching on the new trends in the field, presenting programme documents and consultation. It is constitutes of academics form the IT and health informatics faculties because the curriculum is cross-cutting.
iv. The advisory Board guides, contributes and validates curriculum development processes plus aligning the curriculum with national needs. The advisory board constitute both academic and industry personnel from both information systems and technologies, and healthcare sector, including government agencies. The board is constituted in accordance to the institutional mandate. The role of the board in the development of Health Informatics includes validating the course content and teaching and learning strategies.
v. Senate Committee approves the curriculum documents reviewed and recommended by the faculty board before submitting to the accreditation bodies such as the Namibia qualification Authority (NQA).

People **-** Healthcare informaticians are the people that deliver healthcare services. It is vital for healthcare informaticians to be equipped with recent and relevant knowledge and technological skills.

i. Skill: Informatics majors learn to critically analyze various approaches to processing information using information systems and develop skills to design, implement, and evaluate different generation of information technology tools.
ii. Development: The healthcare informatician should therefore be concerned of how technologies can affect healthcare delivery.

A step-by-step approach is provided to guide the development of health informatics curriculum. The guide comes from the deeper sense making of the analysis and results, of the data.

4.1 Guidelines for Developing a Curriculum

The process of curriculum developing is essential for success achieving educational goals [18] in any field. Several steps and elements are adopted in developing a curriculum. According to [19], guidelines are essential for developing comprehensive curriculum. Ljuca emphasised the need to include elements, such as educational strategy, goals and objectives, and implementation plan, in the development of a curriculum [18]. The following is a discussion on the steps that were considered vital, in the development of a curriculum in institutions of higher learning:

i. **The needs assessment:** an assessment that is aimed to determine a degree, to which the intuition of higher learning implements a stated philosophy of education [19]. Within the frame, a needs assessment will be carried out, to guide the relevance of the courses' outline and their contents. [2] Extended needs assessment by stating that, there are specialized university programs that can be integrated within other professional educational programs (such as medicine, nursing, informatics and computer science). Therefore educational components varies in depth and breadth, to suit specific student groups, for this reason a comprehensive needs assessment is required.

ii. **Relevant stakeholders.** It is critical to identify and involve relevance stakeholders in the development of curriculum. [19] Stated that, curriculum development should be carried out in an inclusive and cooperative manner. All relevant stakeholders should consult each other for the purposes of the development process. A number of committees consisting of academics and non-academic, external and internal stakeholders should be formed, to carry out the processes.

iii. **Curriculum Goals and objectives.** A curriculum should aim to achieve specific objectives, resulting from the needs analysis. This means that specific measurable knowledge, skill and processes targets of the curriculum should be stated [18]. This indicates the direction of a curriculum in terms of the content.

iv. **Educational strategy.** One of the key drivers of the curriculum development is the institution's educational strategy. This support and motivate the goals and objectives of the curriculum. , an inclusive pedagogical strategy is adopted in the development of a curriculum. This includes course level, type of qualification offered. This stage can be expanded to cover content of certain courses. Ljuca [18] stated that, it is necessary to have a plan on how to maximize the impact of the curriculum. This includes which contents should be included, how the contents should be organised, which educational methods should be used, and how the elements of curriculum should be communicated to the general stakeholders.

v. **Development.** The development of a curriculum be carried out by highly qualified expertises, which are the necessarily the tutor or lecturer of the course. As stated above, the development of curriculum should be inclusive in nature, in order to attain high level of comprehension that is required for accreditation. Also, the development of a curriculum must have timeframe, within which it must be completed. This must be in alignment with the academic calendar.

vi. **Implementation:** The implementation stage consists of different attributes and elements, such as management, and organisational:

   a. *Management.* This is the management of the process. It includes and assimilates getting approvals, and accreditation of the course, from the respective authorities. It is also at his stage that the quality checks of individual and group responsibilities are done.

   b. *Organisational.* This aspect of implementation is the organisation of course. It is specifically on how the: contents are organised; delivering methodology is structured; teaching and learning outcomes are formulated; ration between the practical and theoretical concepts are integrated; and assessment strategies are planned.

5 Health Informatics Curriculum for Teaching and Learning

This section presents the implication of developing the Health Informatics curriculum for teaching and learning in institution of higher learning. As revealed in the analysis, particularly on the internationalisation and externalisation in Table 1, the development of the HI curriculum is driven and influenced by two primary factors: rigor and relevance. According to [20], many information systems studies are challenged mainly because they are not of rigor, this is as a result of little or lack of intensity, synthesis and theoretical assumptions' appropriate guidance. On another hand, some research articles which consist of practical in-depth, and application of real-world situation, but lack rigor faces criticisms from IS practitioners.

Rigor - Various inquiring techniques were employed to achieve the rigour of developing a curriculum for institution of higher learning. Theoretical density, including the procedural methods and the elaborative approach, entailed in the development of the HI curriculum makes it rigorous.

Relevance - The course is developed based on its relevance to practitioners' concerns, which were achieved through the advisory board. Relevance was in accordance to the meaning the practitioners associated to the content of the HI curriculum. The advisory board allowed the practitioners to transcend their own knowledge, experience, and need in applying the substantive theory, based on rigour.

The primary factors that were considered by members of the advisory board include economic needs and social context, as well as viability and sustainability, which are to be enabled and supported by people, technology, and through legitimised process.

6 Conclusion

The study revealed that the development of a comprehensive curriculum is process-oriented, which inculpate different factors of positive influences. This includes inclusiveness of major stakeholders, quality of reputable standard, rigor and relevance in the context of academic front.

The process-orient approach enacts two critical factors, acceptability and quality. Due to the inclusive approach, the curriculum is generally accepted by the stakeholders, making the programme more marketable to both potential candidates and employers. The employers enrol to accept the programme, in that it represents their interest. The students find comfort in the programme as a result of employers' buy-in. Also, the extensive consultation and inclusiveness enhance the quality of the curriculum. This includes contributions from experiential learning, and benchmarking with already matured curriculum of other institutions.

The combination of the factors, rigour and relevance add and in-depth theoretical and practice values to the development of the HI curriculum for institutions of higher learning. The quality and the credibility that are instilled in the HI curriculum through rigour and relevance give the stakeholders the confidence to invest in the newly developed course. We have concluded that there is the need for educational programme that will produce qualified health informaticians.

This study will benefit academic managers who intend to development health informatics or other curriculum, for teaching and learning, in institutions of higher learning. Also, it will help those, who intend to undertake review of their curriculum – it is an opportunity to imbibe learning from this study.

References

[1] Malik, M.A., Khan, S.: Health informatics training at the postgraduate level: a curriculum development exercise in Pakistan (Electronic Version). Journal of Health Informatics in Developing Countries 1(1), 16–18 (2007), http://www.jhidc.org/index

[2] Mantas, J., Ammenwerth, E., Demiris, G., Hasman, A., Haux, R., Hersh, W., Hovenga, E., Lun, K.C., Marin, H., Martin-Sanchez, F.: Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics – First Revision. EJBI 7(2), 3–17 (2011)

[3] Polytechnic of Namiba, Strategic plan three -SP3 (2010), http://www.polytechnic.edu.na/strategicplan

[4] Covvey, H.D., Zitner, D., Rernstein, R.: Pointing the Way: Competencies and Curricula in Health Informatics (2001)

[5] Sullivan, F., Wyatt, J.C.: ABC of Health Informatics, Massachusetts, USA (2006)

[6] Kimaroa, C.H., Nhampossa, J.: The challenges of sustainability of health information systems in developing countries: comparative case studies of Mozambique and Tanzania. Journal of HID 1(1), 1–10 (2007)

[7] McCuddy, M.K., Pinar, M., Gingerich, E.F.R.: Using student feedback in designing student-focused curricula. International Journal of Educational Management 22(7), 611–637 (2008)

[8] Baskerville, R., Stage, J., DeGross, I.: Organizational and social perspective on Information Technology. Kluwer Academic Publisher, The Netherland (2000)

[9] Martz, B., Zhang, X., Ozanich, G.: Developing an Integrative Health Informatics Graduate Curriculum. Communications of the Association for Information Systems 19, 645–654 (2007)

[10] Hersh, W.: A stimulus to define informatics and health information technology. BMC Medical Informatics and Decision Making, 1–6 (2009)

[11] Garrib, A., Stoops, N., McKenzie, A., Dlamini, L., Govender, T., Rohde, J., Herbst, K.: An evaluation of the district health information system in rural South Africa. SAMJ 98(7), 549–552 (2008)

[12] Hamunyela, S., Iyamu, T.: Readiness assessment model for the deployment of health information systems in the Namibian- mohss. In: Proceedings of the 12th International Conference on Social Implications of Computers in Developing Countries, Ocho Rios, Jamaica, pp. 842–852 (2013)

[13] Namibia. OPM, The e-governance policy for the public service of Namibia (2005), http://www.grnnet.gov.na/grnnews/ (accessed September 12, 2010)

[14] MoHSS, Ministry of Health and Social Services Strategic plan (2009), http://www.africanhealthleadership.org/wp./MOHSSStrategicPlan.pdf

[15] Yin, R.K.: Case study research: Design and methods, 3rd edn. Sage, Thousand Oaks (2003)

[16] Hancock, D.R., Algozzine, B.: Doing case study research: A practical guide for beginning researchers. Teachers College Press, New York (2006)

[17] Walsham, G.: Interpreting Information Systems in Organizations. Wiley, Chichester (1993)

[18] Ljuca, F., Lozo, S., Simunovic, V., Bosse, H., Kadmon, M.: Curriculum development, ch. 11 (2009), http://www.bhmedemanual.org/book/export/html/99

[19] Wiles, J., Bondi, J.: Curriculum Development: a guide to practice, 2nd edn. Charles E. Merrill, US (1984)

[20] Iyamu, T.: Underpinning Theories: Order-of-Use in Information Systems Research. Journal of Systems and Information Technology 15(3), 1–13 (2013)

Visualization of Services Availability: Building Healthy Communities

Waqar Haque[1], Bonnie Urquhart[2], and Emery Berg[1]

[1] University of Northern British Columbia, Prince George BC, Canada
[2] Northern Health, Prince George BC, Canada
{waqar.haque,eberg0}@unbc.ca, bonnie.urquhart@northernhealth.ca

Abstract. Business Intelligence techniques have been applied to data related to healthcare infrastructure. Data visualization has eliminated tedious analysis of legacy reports and provides mechanism for optimally aligning resources with the needs. Once the data from disparate sources is cleansed and integrated into a data warehouse, the OLAP (Online Analytical Processing) cube allows slicing along multiple dimensions determined by key performance indicators representing population and patient profiles together with complex management groups. In addition, comparison, availability, service levels and community health reports are also generated on demand. All reports can be drilled down for navigation at a finer granularity. The use of mapping tools, customized shape files and embedded objects further augments the navigation. Finally, web forms provide a mechanism for remote uploading of data and transparent processing of the cube.

Keywords: Business Intelligence, Healthcare Informatics, Services Availability, Data Visualization.

1 Introduction

Healthcare extends beyond medicine in many ways. One of these is the ability to access healthcare services particularly when one is located far from the core infrastructure. Access to relevant information in an intuitive form not only benefits the patient, but also assists the administration in identifying areas where resource allocation may have the highest impact. This ultimately leads to healthy communities and optimal use of healthcare funding. Fortunately, large volumes of information have been gathered over years and serves as a base for achieving the envisioned goals. Unfortunately, this information sits in legacy formats and the sheer volume of data makes it incomprehensible for any use other than the specific purpose for which each dataset was gathered. In addition, the data is of poor quality, suffers from inconsistencies and lacks integrity. Despite having the data, healthcare providers and supporting staff are faced with the challenge of determining the type and location of resources accessible to them and their patients. In order to locate this information, an extensive search through numerous workbooks, databases and statistical web sites is quite common. Even then, it can be extremely tedious to find the needed information from these sources because they generally differ in purpose and tend to be inconsistent with each other.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_34,

In this paper, we propose how Business Intelligence (BI) techniques can be applied to such data in order to make this information more accessible and comprehensible for a broader group of people. Our envisioned goal has been achieved by first gathering the sources into a singular entity, second providing interactive access and control of the underlying data, and finally visually representing the data through reports. By applying these techniques, the resulting information can be accessed through dashboards which provide a quick overview of the key performance indicators (KPIs) and allow navigation to underlying reports of finer granularity. Thus, instead of sifting through massive spreadsheets for the desired information, one can now access a centralized system which renders reports in matter of seconds. The system also extends to other tools such as web forms for updating data by designated staff without the need for going through complex IT protocols. The underlying data represents geography and services for the entire region covered by Northern Health (NH), that is, a population of ~300,000, land mass of ~600,000 km^2 and the breadth of services from health prevention and promotion through to acute care services.

1.1 Methodology

Business Intelligence tools and techniques are an effective way to integrate and analyze large data repositories. However, the integration process becomes challenging when the data is not collected with analytics in mind. In our solution, we have used Microsoft SQL Server's BI tool stack [1] and ASP.NET [2] to make the data more accessible and reduce the time that data analysts spend searching through large collections of sources. We have also merged the disparate data sources to eliminate data conflicts and create a singular source for reporting. The resulting data warehouse becomes the central source for all analysis and reporting. An ETL (extract-transform-load) [3] process is used to populate the data warehouse. During the extract phase, connections are created to various data sources and the required information is pulled in a temporary storage. In the transform phase, the format of stored data is made consistent with metadata prior to loading into the data warehouse. The SSIS component in Microsoft's BI tool stack is used to accomplish this integration. An Online Analytical Processing (OLAP) cube is then created using the Analysis Services. This cube is an n-dimensional structure which can be used to reveal more complex details at various levels of granularity through predesigned and ad-hoc queries. The cube consists of several dimensions and fact tables [3]. Using the cube structure, reports are created and rendered through Microsoft's Reporting services [1]. SQL Server Reporting Services provides a rich set of data visualization features such as charts, tables, matrices, gauges, maps, and tooltips. A dashboard gives a high-level overview of the KPIs and acts as a central navigation hub to other reports. Mapping allows the user to see the information based on regions and provides visual representation of distances between locations.

Web forms are used to allow users to remotely update information with automatic consistency checks. Normally, updates to a database require knowledge of the underlying structure and the associated query language. By providing a web form, these queries are created automatically and the database structure is represented visually for easy understanding.

2 Related Work

Historically, the healthcare field has been slow to adapt new computer technologies; this has been largely due to hardware limitations, computer literacy, mechanical user interfaces and privacy concerns. The first two causes have been mostly overcome due to penetration of computers in daily lives and the technological advancements in computer hardware, but many applications still are mechanical in nature and are not intuitive to the user [4]. The Minnesota Health Association developed a pilot program to combine clinical information with administrative data, which faced many challenges such as the expertise of those involved and communication issues resulting from distributed data sources [5]. BI tools and techniques have been used to provide insight into ambulatory care sensitive conditions within Northern Health by analyzing data and identifying areas that need attention [6]. These techniques have also been successfully used to improve the management of large quantities of medical information [7]. Historical information and comparisons with the United States' primary care system has shown that providing improved access to primary care reduces the cost of healthcare and enhances the care provided to patients [8]. A comparison survey observed that with the increased access to healthcare in Canada, the general health of the public was superior to that in the United States [9]. Another study showed that though Canada has relatively lower cost of healthcare, the wait times negatively affect the perceived availability of care [10]. Additionally, disparities in healthcare and its access have been shown to be negatively related to lower income, education, and race both due to perception and access [11]. Unfortunately, there has been little or no significant evidence of work that incorporates the concept of Business Intelligence in analysis of data related to services availability.

3 Data Challenges and Cube Design

The underlying data was collected over several years for varying purposes including generation of community health reports. The complexity of the underlying data posed several challenges in the integration phase. The first and foremost challenge was the sheer number of workbooks which have been the primary source of information for the data analysts for several years. An initial screening eliminated irrelevant data, but even after this exercise a very large number of workbooks remained. Most of these contained several sheets which were created for a variety of (sometimes unrelated) purposes, which meant the data did not always match in content or level of granularity. Even the repeated numbers sometimes differed across the workbooks. The data was available at various levels of hierarchy making aggregations unpredictable. Similarly, different naming schemes were used for locations without specifying any clear relationship(s) among them. To deal with this, fuzzy lookups [3] were used by specifying a threshold to match names which are similar enough but not identical. For locations which failed to match, a manual mapping table was created and the names were corrected at the database level through SQL queries.

A relational model was developed to create a singular source of information. This database consisted of twenty four relations which were populated by three dump sheets via web forms (described later). An integration package was built to cleanse, combine and group the data based on its purpose and granularity. When there were conflicts due to repeated information, the selection was based on conformity with other sources and the age of data. In rare cases, informed calculations were performed to correctly reflect missing values. The next step was to create an analysis cube using this database. Normally such cubes use star schema with a single fact table and multiple dimensions [3]. While this structure gives superior performance due to a reduced need for joins, it requires all information to exist at the same level of granularity. This was impractical in our case and could potentially lose any advantage because of the need for added rows and empty cells if the data were to be re-structured. Thus, in our somewhat unusual design, eight fact tables and nine dimensions were used (Fig. 1), primarily for performance and granularity.

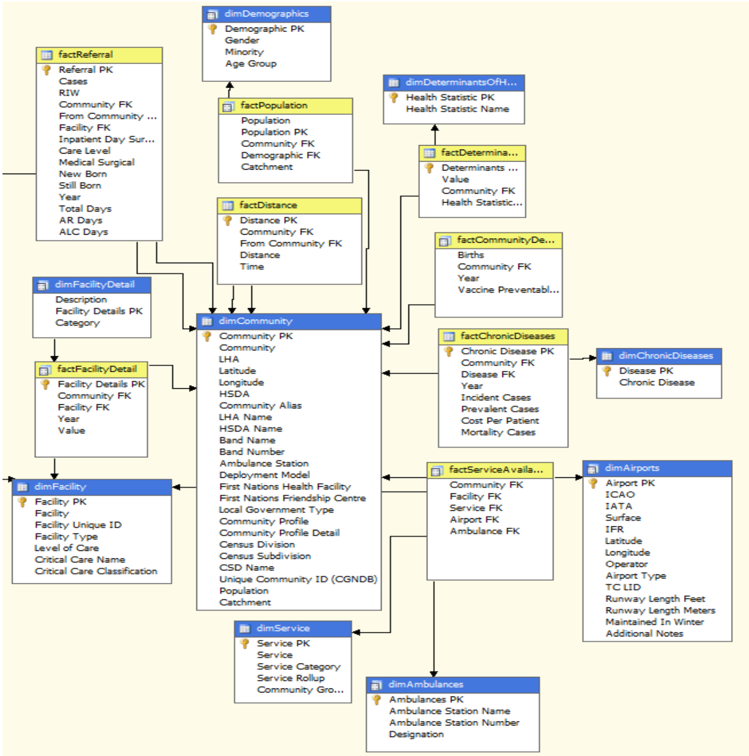


Fig. 1. The OLAP Cube

Another challenge was to allow seamless update of data by analysts and staff unfamiliar with the underlying schemas and not trained to write sophisticated queries. Besides having a capability for bulk loading of large volumes of data, there was also a need for the ability to update individual rows without affecting the integrity of the database. To provide this functionality, a web form was created in ASP.NET [2].

An intuitive combination of tabs, groupings and drop down lists allow data entry into individual cells of the selected table. The entered values are checked against metadata before updates are committed. Another mechanism to prevent inconsistencies is the use of drop down lists when names are referenced. For bulk loading, two dump sheets reflecting the database structure were created. These sheets allow data to be compiled or manually entered and uploaded to the web form. An integration process then triggers to transparently upload the data and reprocess the cube. The integration is fast, stable and reliable due to instant validation and simplified logic.

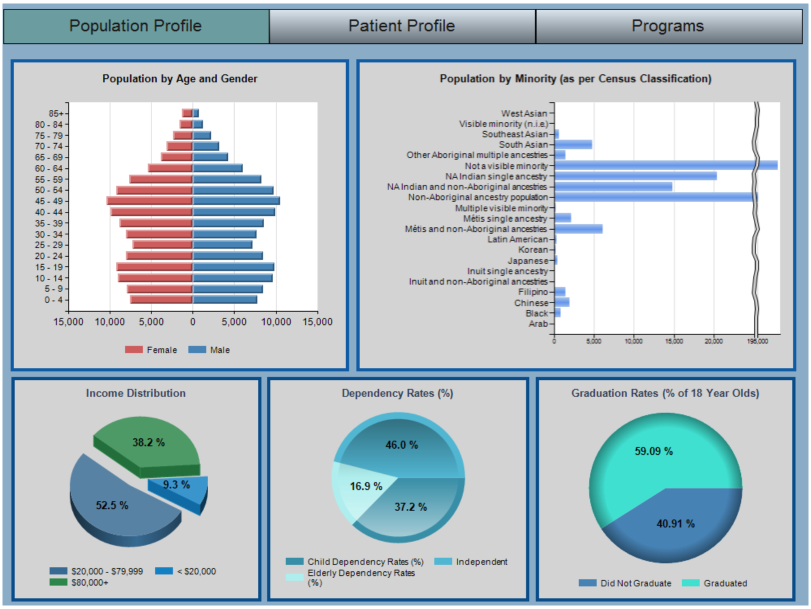


Fig. 2. Main Dashboard: Patient Profile

4 Data Visualization

For interactive access to information and data visualization, MS SQL Reporting Services [1] were used to generate dynamic reports. In addition to conventional charts and graphs, access to advanced features like mapping, navigational controls and parameterization of reports is also provided. The information contained in these reports can be updated through the web form which automatically reprocesses the cube and immediately reflects the changes.

4.1 Main Dashboard

The main dashboard provides an overview of the KPIs and includes navigation controls including a toggle control to switch between demographic information and patient profiles. The demographic information displays information relevant to the population status including factors such as wealth, education level, origin, and dependency rates (Fig. 2). These metrics assist in identifying potential areas of concern and any needed level of support and services. The patient profile gives an overview of the health related metrics within the selected region by showing information such as births, commonality of chronic conditions, and vaccine preventable diseases (Fig. 3). This information is available at all levels of hierarchy with the granularity becoming finer from Northern Health Authority to Health Service Delivery Areas (HSDA) to Local Health Authorities (LHA) to Communities. This hierarchy can be selected from the maps which in turn generates the parameters for necessary filtration of information.

Tabbed controls allow switching to other reports while maintaining the current level of hierarchy. These tabs allow access to subgroups such as availability of services, comparisons of selected regions, service levels, and direct access to community profiles. To improve navigational performance, the header is embedded with parameters to track the current tab select, type of report (drill-down), and the current level. This allows transparent passing of parameters to determine which report or level to load next based on the direction of navigation.

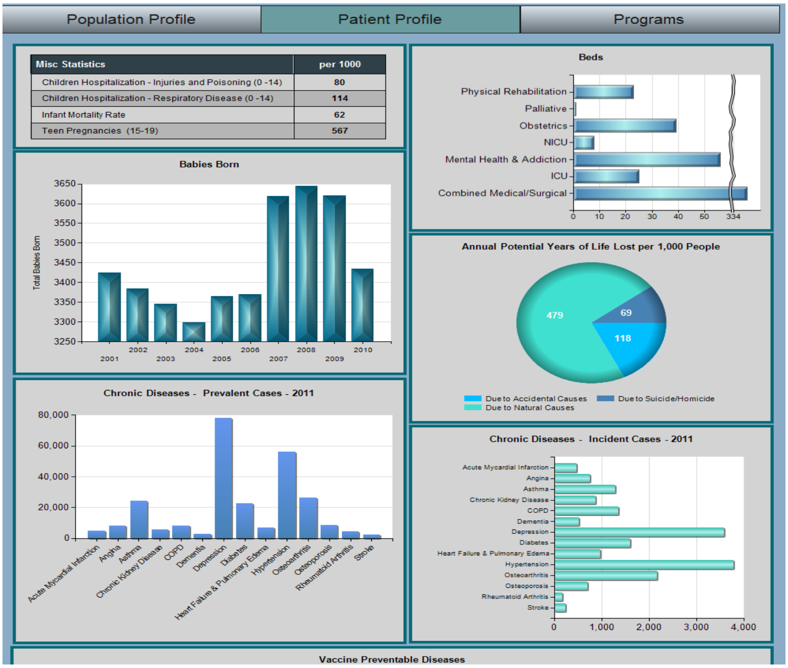


Fig. 3. Main Dashboard: Patient Profile

4.2 Mapping Functionality

The mapping features allow location based visualization and navigation. A challenge, however, was the lack of availability of full range of maps. While default maps are provided by the tools, no maps of British Columbia (BC) were included; this resulted in the need for a shape file to store geographic information such as the shape of regions, locations of communities or other geographic features. The shape files of BC were obtained from [13] and modified to better fit our needs. These modifications were done through an open source Geographic Information System (GIS) application, QuantumGIS [14]. Using this application, the shape file of BC was restricted to the area covered by Northern Health; another shape file was created for storing the locations of communities within the region. To address stability issues created by large size of single shape file, steps were taken to limit the information contained therein, primarily by removing information that was available elsewhere.

Maps were also used for controls in the comparison report, and to visualize availability of services in the community or proximity. Examples of these controls can be seen in Fig. 4 where the map has been used to select LHAs for comparisons, and Fig. 6 where the location of selected services are shown by indicators on the map.

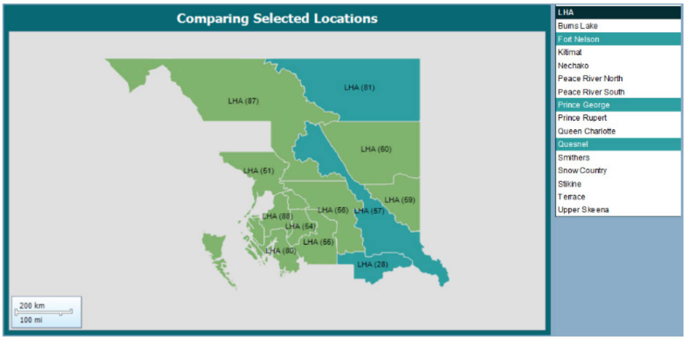


Fig. 4. Comparison Map

Community Map Comparison Tables

	Fort Nelson	Quesnel	Prince George
Total Population	4,664	9,710	74,547
⊞ Community Services Available	2	2	4
⊞ Airports	4	2	4
⊟ Medical Services Available	74	145	249
Aboriginal Health	2	3	3
General Services	4	9	14
HCC	7	12	17
Imaging	3	8	12
Mental Health & Addictions	5	6	23
Outpatient	1	8	20
Perinatal	2	2	3
Population Health	9	9	9
Preventative Public Health	14	17	18
Public Health	2	23	2
Public Health Protection	18	18	18
Specialist MDs	5	7	40
Surgery	2	23	70

Fig. 5. Community Comparison Tables

4.3 Comparison

Comparisons between differing regions provides further insight into the state of health care within a region and potential causes for disparities [9]. The regions of interest can be selected by simply clicking on the map. The comparison can be performed at all levels and allows for comparing up to three regions at a time (Fig. 4). When a new region is selected for comparison, the three most recent selections are maintained. Metrics such as population, facilities, community services (airports, ambulances, etc.) and medical services are displayed for each selected region (Fig. 5).

4.4 Service Availability

Examining the demographic information, patient details and availability of services has been shown to be effective in identifying possible needs of a region and improving the care of patients [11] [8] [12] [10]. This information has been provided using colour coded markers on the map. For readability purposes, we show up to four services in circles split into quarters (Fig. 6). The services are selected from a categorized list; additional information about the locations is displayed through tooltip display when hovering over the circles. Each time a service is selected, the map is updated to show the communities which have facilities offering the selected services. All services offered in NH, whether or not those are available locally, can also be seen in the availability report at the community level. This expandable list shows the proximity of where missing services can be found and the distance/travel time from the current community (Fig. 7).

4.5 Other Reports

The reports illustrated in this paper are a small representative sample due to space limitations. There are several other main and drill-down reports which provide various perspectives of the services availability. For instance, a community health report contains transfers/referrals information from/to the selected community in addition to charts and graphs which appear elsewhere in the application. These reports are printable and generally made available to communities. Similarly, many charts open a popup window instead of loading another report. These popups windows consist of descriptive charts or tables, definitions and contain information about the source of data together with names of data analysts responsible for the information.

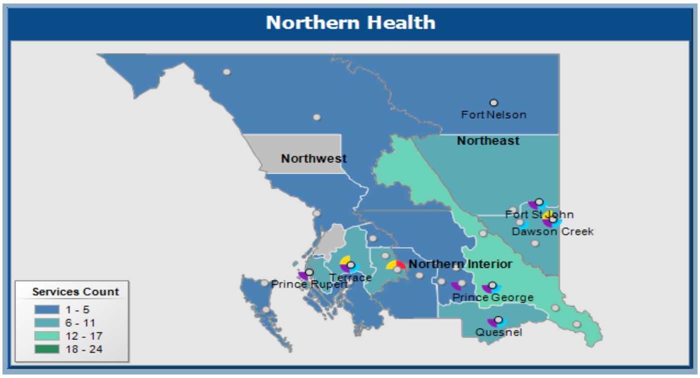


Fig. 6. Availability Map

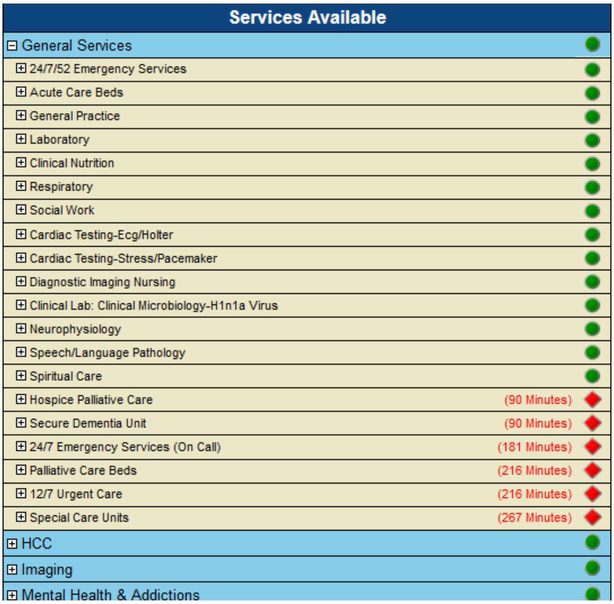


Fig. 7. Community Profile Service Availability

5 Conclusion

We have demonstrated how BI techniques and tools can be used in non-traditional areas of healthcare environment to make informed decisions with reference to resource allocation and enhancement of the quality of patient care. The multidimensional cube allows analysis of data in several dimensions and reports are generated within seconds. The data can be kept up to date all year around while preserving integrity during interim reporting. Originally, the data was updated annually due to complexity of data collection and compilation. The versatility of reports is enhanced through parameterization which allows values to be passed between sub-reports. The interaction of web forms with the underlying database and cube allows for transparent data upload and integrity checks. The interactive reports provide users with valuable information such as proximity to location of available services, facilities with specific needs, comparative analysis and tools for resource reallocation, if necessary. For privileged information, access controls have been implemented. The solution is modular and new datasets such as for smoking rates, teen pregnancies, HIV rates, immunization coverage, and vital statistical summaries can be easily integrated into the existing dashboard. The model can also be extended to other programs such as Home and Community Care, Critical Care, and Mental Health and Addictions.

Acknowledgments. This work was funded by a collaborative research grant from Northern Health, British Columbia, Canada. Among others, Kari Harder, James Haggerstone, Keely Maxwell and Matthew Amsel have been very instrumental in compilation, loading, and verification of underlying data.

References

1. Microsoft Corporation: Business Intelligence, http://www.microsoft.com/en-us/bi/ (accessed 2013)
2. Microsoft Corp.: The Offical Microsoft ASP.NET Site, http://www.asp.net/ (accessed 2013)
3. Larson, B.: Mircosoft SQL Server 2008 Delivering Business Intelligence. McGraw-Hill Companies (2009)
4. Cantrill, S.: Computers in Patient Care: The Promise and the Challenge. Communications of the ACM 53(9), 42–47 (2010)
5. Pine, M., Sonneborn, M., Schindler, J., Stanek, M., Maeda, J., Hanlon, C.: Harnessing the Power of Enhanced Data for Healthcare Quality Improvement: Lessons from a Minnesota Hospital Association Pilot Project. Journal of Healthcare Management 57(6), 406–418 (2013)
6. Haque, W., Edwards, J.: Ambulatory Care Sensitive Conditions: A Business Intelligence Perspective, York, pp. 31–39 (2012)
7. Olszak, C., Batko, K.: The use of business intelligence systems in healthcare organizations in Poland. Computer Science and Information Systems, 969–976 (2012)
8. Starfield, B., Shi, L., Macikno, J.: Contribution of Primary Care to Health Systems and Health. Milbank Quarterly 83(3), 457–502 (2005)
9. Lasser, K., Himmelstein, D., Woolhandler, S.: Access to Care, Health Status, and Health Disparities in the United States and Canada: Results of a Cross-National Population-Based Survey. American Journal of Public Health 96(7), 1300–1307 (2006)
10. Sanmartin, C., Berthelot, J.-M., Ng, E., Murphy, K., Blackwell, D., Gentleman, J., Martinez, M., Simile, C.: Comparing Health And Health Care Use In Canada And The United States. Health Affairs 25(4), 1133–1142 (2006)
11. Lurie, N., Dubowitz, T.: Health Disparities and Access to Health. The Journal of the American Medical Association 297(10), 1118–1121 (2007)
12. Mäntyselk, P., Halonen, P., Vehviläinen, A., Takala, J., Kumpusalo, E.: Access to and continuity of primary medical care of different providers as perceived by the Finnish population. Scandinavian Journal of Primary Health Care 25(1), 27–32 (2007)
13. BCStats: Translations and Data Sets, http://www.bcstats.gov.bc.ca/StatisticsBySubject/Geography/TranslationsDataSets.aspx (accessed 2013)
14. Team, Quantum GIS Development: Quantum GIS Geographic Information System, http://www.qgis.org/ (accessed 2013)

Information Persistence Architecture for Informal and Formal Care Providers

Yosvany Llerena Rodriguez[1], Carlos Cardoso[2], Miguel Grade[2], Filipe Augusto[2], Alexandra Queirós[3], João Quintas[4], and Nelson Pacheco Rocha[5]

[1] Instituto de Engenharia Electrónica e Telemática de Aveiro, University of Aveiro, Campo Universitário de Santiago, Aveiro, Portugal
[2] Maisis, Rua D. Manuel B. Vasconcelos, Aveiro, Portugal
[3] Health Sciences School, Instituto de Engenharia Electrónica e Telemática de Aveiro, University of Aveiro, Campo Universitário de Santiago, Aveiro, Portugal
[4] Instituto Pedro Nunes, Rua Pedro Nunes, Coimbra, Portugal
[5] Health Sciences Department, Instituto de Engenharia Electrónica e Telemática de Aveiro, University of Aveiro, Campo Universitário de Santiago, Aveiro, Portugal
{carlos.cardoso,miguel.grade,filipe.augusto}@maisis.pt,
{llerena,alexandra,npr}@ua.pt, jquintas@ipn.pt

Abstract. Health and social care systems currently face with a set of challenges that requires integrated approaches. Integrated care comprises formal care networks, both health care and social care, and informal care networks resulting from a diversity of community-oriented structures (e.g. relatives, friends or voluntary groups). This paper presents an information persistence architecture designed to support the development of information services specially focused on informal care providers. Informal care is inherently slightly structured and, consequently, imposes high demanding requirements related with the flexibility and interoperability of the information persistence.

Keywords: Integrated Care, Personalized Care, Personal Health Record, Social Health Record, Interoperability.

1 Introduction

Needs based approaches require interventions not solely focused on medical purposes but also on a range of essential activities for the maintenance of individuals' quality of life and that are part of the normal living of every citizen. Concerning frail people, namely elderly people, those activities require additional informal and formal care services. The integrated care concept is associated with the patient perspective, as well as the implications of organizing and managing the different health and social care services in order to maximize quality, access, efficiency, effectiveness and satisfaction of the patients.

Various interacting organizational structures may co-exist within integrated care networks. First of all, there are the formal care networks, both health care and social care networks regulated by collaboration contracts or agreements. Secondly, there are

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_35, © Springer International Publishing Switzerland 2014

informal care networks resulting from a diversity of community-oriented structures including relatives, friends, voluntary groups, churches, time banks or non-governmental organizations. The synergies between all the stakeholders, independently of the networks types, can be properly mobilized by adequate organizational and support structures, including information systems and services.

The We.Can platform intends to answer to the identified needs. Its main objective is to provide information systems and services related to health and quality of life domains to support a digital health continuum where health and social evidence and personal data can be used in the entire range from lifestyle management to chronic diseases management [1, 2]. It should be pointed the We.Can platform does not intent to substitute hospital health information systems neither the distributed clinical information networks. It is being designed to interact with the existing information sources but not to duplicate them and it should support care outside hospital wards, according an integrated care perspective. The We.Can platform is being designed to support informal and formal care providers that do not have access to structured information services related with their patients.

The platform comprises a set of common services, namely the required security functions (e.g. authentication, authorization, confidentiality and integrity, non-repudiation, security compliance, logging, or security monitoring and auditing). Specifically, this paper will emphasise the information persistence architecture of the mentioned platform, which have high demanding requirements due to the existence of open slightly structured usage scenarios, in association with the need to guarantee internal and external interoperability. Concerning the interoperability issues, the authors propose the use of the Reference Information Model (RIM) [3], developed by Health Level Seven (HL7), as the foundation of the information persistence services.

The paper comprises, besides this introductory section, four more sections: Related Work, Methods, Application and Conclusion.

2 Related Work

The socio-economic implications resulting from the increase in the proportion of elderly in the world population are seriously worrisome in a large number of countries: the decrease of supply in the workforce, the decrease in productivity and the reduction of mobility and flexibility of workers, despite their greater experience, will lead to an increase in the burden of tax deductions on employers, resulting in a reduction in household (as there are fewer people earning money) and State savings (as a result of social expenditures, namely retirement plans) [4]. On the other hand, the frequency, severity and complexity of treating chronic illnesses, which appear or worsen with the natural process of ageing, will, inevitably, increase the cost of health care and social care per capita. Furthermore, a profound alteration in the household structure implies a decreasing number of unpaid care providers.

All these trends imply the need for a close relationship between health care and social care services [4]. The integrated care concept is being considered with an increasing importance. Health systems, particularly the ones of the developed countries, must serve a population with high academic qualifications and a relation with health and disease quite different from the one that existed in the past, i.e. with expectations to live longer and with autonomy. It is foreseen high demands in terms of the quality of care, so that integrated care arises as a possible response to ensure a more efficient and effective use of resources as well as the provision of care according to an individualized paradigm.

Health domain information and associated data structures are complex and formalized with different methods by different institutions. Nowadays, the needs for retrieving, managing and delivering large amounts of health information are being met by Electronic Health Records (EHR) [5]. An EHR, in its simplest form, consists of an electronic file containing clinical information of the individuals and can help to personalize care, prevent medical errors, promote the consistency of care, support the referral of patients to the correct services or control the costs [6].

However, health conditions are influenced by the individual's medical history, as well other factors distributed across different levels of impact that interact with each other continuously and in subtle ways [7]. These include behavioural (e.g. data associated to medication adherence), social (e.g. data associated with activities and participation) and environmental factors (e.g. quality and allergens presented in the air). For instance, a diet plan is influenced by individual's medical history as well as behavioural factors (e.g. physical activity) and environmental factors that either hinder or facilitate these behaviour factors. In this particular, it is important to consider a significant range of Ambient Assisted Living (AAL) systems and services, including biomedical devices, namely mobile and wearable sensors able to monitor physiological parameters, activities and behaviours [8].

Therefore, it is needed to consider new requirements in terms of distributed management, integration and use of a whole range of information, namely Personal Health Records (PHR), Electronic Social Records (ESR) and AAL generated information.

If the citizens are able to contribute with documentation, namely observations of theirs daily living [9], they can have an active role in the management of their health and care pathway. Therefore, PHR have, nowadays, an increasing importance. They include data and information related with the individuals' lifetime and individuals' care maintain by each individual, namely patient-reported outcome data. Furthermore, PHR can represent more than a repository for the individual data, because they are able to combine data, information, knowledge and tools to help any individual to be proactive in their own care [10]. This stands in contrast with EHR, which are operated by organizations and contains data entered by professionals.

Since the service models employed by health care providers and social care providers are different, there are important differences between EHR and ESR [11]. Healthcare records are focused on a single patient, often with considerable details and depth, and have a sensitive nature, which is why several laws and regulations mandate to protect the privacy of the patients. On the other end, social care records place the individuals in their daily living context of family and other informal carers, including the attitudes and effects on each, so as to ensure mental support and understanding [11].

Finally, the interconnection and interoperability of medical devices and sensors have brought new possibilities in terms of automatic collection of information about individuals and environment [12].

3 Methods

Due to the complex reality associated with the integrated care ecosystem, there is the need of innovative solutions for the management of all the available information that is required for the mediation between users, careers and institutions. These solutions should present efficient information persistence services for a broad range of applications (including PHR, ESR and AAL generated information), communication services to interoperate with external data sources through normalized protocols (e.g. HL7) and compliant devices interfaces to integrate different sensors (e.g. sensors to monitor the general conditions of the individuals or specific devices to monitor health related parameters).

The goals of the integrated care paradigm were considered for a preliminary definition of the information persistence requirements. These preliminary requirements together with a deep analysis of the related work, namely scientific reports, specifications, technologies and standards allowed the definition of the concrete information persistence architecture. Since the required reorganization processes can be highly unpredictable [13], it was considered a fundamental requirements the flexibility of the information models and structures. One of the obstacles that should be surpassed is the creation of flexible models to ensure the information persistence and respective semantics for all the potential applications related with different specialization areas. For that can be used an open management information paradigm [14] with robust and stable domain models separated from the concrete implementations. Therefore it is convenient the separation of the so-called knowledge and information models. This approach has been considerably developed by the openEHR promoters [10].

3.1 Knowledge Model

The knowledge model considers the domain and application specific concepts and should be able to support specific adaptations. For that archetypes with constraint rules define the underlying information model. According to this approach, archetypes are instances of an object oriented implementation, which means they can be created and manipulated by adequate tools, without changing the underlying information technical specifications. The definition of the archetypes structure was based on the following reasoning: there is contextual information that should always be recorded and there is information related that will vary according to patients, events and service providers. Therefore, the care processes require different types of archetypes.

A generalization of the basic archetype was performed in order to systematize the overall information considered essential to characterize an encounter between a patient and one or more care providers, whether informal or formal care providers. Therefore, an archetype consists of a common part and a flexible part that varies according to the specialization of the general archetype. The common part is the header and is inherited by all the base archetypes. The flexible part is a specialization and may take three information structures: i) observation - keeps track of all that is seen, measured or obtained by patients' responses to certain questions; ii) evaluation - related to the analysis made by the involved care providers and that can be related with problems, justifications or reasoning; iii) instruction - indicates what should be performed, including requests and responses related with the assessments that have been made, the care that should be provided or recommendations to third parties.

Fig. 1 presents a mind map representing the basic structure of any archetype consisting of a header, a body and data. The header contains information such as the type, name, author and header details, which, in turn, comprise several attributes.

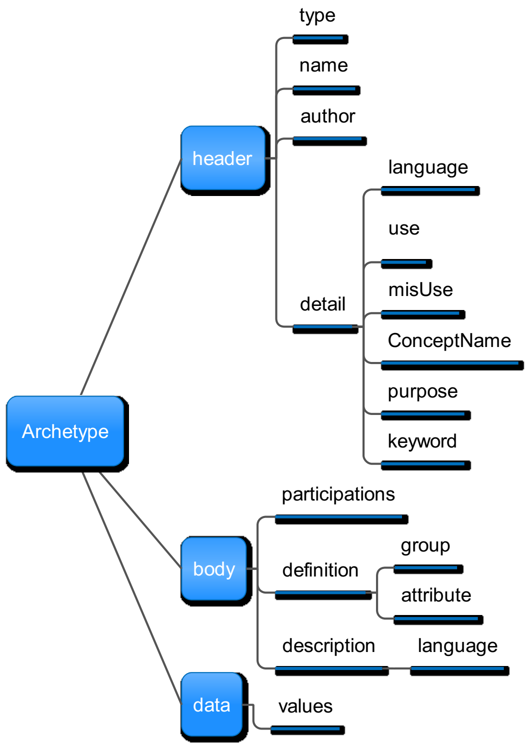


Fig. 1. Archetype Structure

The body of the archetype is a dynamically modeled field that contains three main elements: participations, description and definition. Participations are contextual information and state the responsible for the introduction of the information related with each patient. The body description indicates what are the body elements, being possible a description in multiple languages. Finally, definition is used to specify one or more groups of information contained on the data field. Each group has name, id, the minimum and maximum number of occurrences, and if the field is required or not. The Fig. 2 presents the formalization of the body structure.

Collections contain elements, while an element consists of a code, a description, and the definition of the respective type, range, units, mandatory and the default value. Regarding the data types of the date field, these refer to data types properly normalized. The groups and elements data structures are validated during the generation of the eXtensible Markup Language (XML) files that formalize the archetypes.

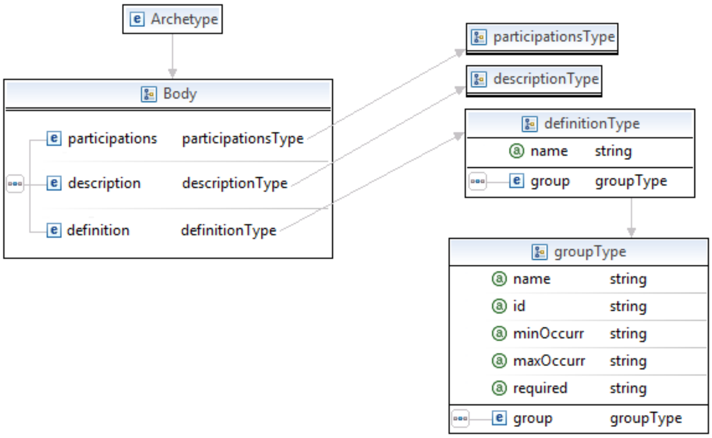


Fig. 2. Formalization of the body structure

3.2 Information Model

The existence of generic models designed to support a comprehensive range of clinical information, such as the HL7 RIM, ensures some stability to the information systems. The concept was validated by defining a data repository, accordingly to RIM and a set of functions for the management of the archetypes, in order to shape the RIM to specific application domains.

The RIM v3 model is based on a simple structure with three main classes: Act, Entity and Role. Each event is an Act that involves the participation of one or more Entities. An Entity is anything that exists or has existed and that may or may not have life. For example, an Entity can represent people, objects, whether concrete materials or chemicals or abstract concepts such as an organization. Finally, a Role may be defined as a competence of an Entity or a function that an Entity may perform.

The three main classes can be joined by three association classes: ActRelationship, Participation and RoleLink. ActRelationship is a relationship between two actions, from a source action to a destination action. Participation defines the involvement of a Role with an Act. It shows how an Entity that plays a certain Role, works to accomplish a particular Act. RoleLink is a relationship between two Roles and it allows, in a simple way, to relate various Roles as, for example, the relationship between different family members or between different elements of a team of care providers.

Subjacent to the implementation of the RIM based information container, there was the concern to provide comprehensive operations to access the underlying relational databases. It should be pointed that the delete operation was not considered because the RIM implementation should not support the deletion of previously persisted records. However, the other required create, read and update operations were designed to provide a simplified access to the relational databases by the final information consumers (i.e. applications).

Furthermore, considering the adopted approach (i.e. archetypes to model the domain specific concepts) it is critical to consider adequate information retrieving mechanisms for the final information consumers. This flexibility was achieved by the implementation of a domain specific language with a syntax specially designed to allow the formalization of flexible queries.

The implemented information services follow the architectural style Representational State Transfer (REST), a set of principles that define the use of Hypertext Transfer Protocol (HTTP) and Uniform Resource Identifiers (URI) for the client-server communications. REST style allows technology independence, improves the overall efficiency and promotes greater scalability.

4 Application

Accordingly to the followed approach it is necessary to assume that the RIM and archetypes have different goals and formalization principles. RIM is generic, complex and difficult to manage while archetypes should be facilitators of the information access. For that it is necessary to ensure that the information persistence obeys to predefined structures by archetypes and that the access to the information in RIM obeys to constraints that are also dependent on the archetypes and that any data structure from the archetypes (e.g. person, devices or organizations and related information) should be mapped and persisted in RIM without any limitation.

Ensuring data consistency is not natively given by the RIM data model. Thus, it would be extremely difficult and painful to ensure consistency with RIM if there were no archetypes and no applications and services to support the development of the final applications. The effectiveness of the proposed model requires a specification of a metamodel that should express the knowledge of this domain. This formalization must consider the basic archetypes structure previously referred. This allows, in certain situations, the creation of different models keeping, however, a logical and coherent relation with the established procedures. Thus, through refinements, it is possible to construct all archetypes required for a broad range of applications that are expected to be supported.

The formalization approach can be explained with a concrete example related with a self-administrated depression assessment instrument.

Several depression evaluation instruments are available. The 9-item Patient Health Questionnaire (PHQ-9) has been studied as a useful screening tool of depression and the Portuguese version is a public domain clinical instrument. It is suitable to make criteria-based diagnoses of depressive disorders and it is reliable and valid to measure depression severity. Furthermore, PHQ-9 is increasingly used as a patient report outcome to assess the level of depression severity (for initial treatment decisions) as well as an outcome tool (i.e. to determine treatment response) [15].

The major components of the PHQ-9 instrument are the 9 questions to evaluate how often over the last two weeks the patient have been bothered by a set of problems: i) little interest or pleasure in doing things; ii) feeling down, depressed, or hopeless; iii) trouble falling or staying asleep, or sleeping too much; iv) feeling tired or having little energy; v) poor appetite or overeating; vi) feeling bad about himself/herself; vii) trouble concentrating on things, such as reading the newspaper or watching television; viii) moving or speaking so slowly that other people could have noticed, or the opposite being so fidgety or restless that have been moving around a lot more than usual; ix) thoughts that you would be better off dead, or of hurting himself/herself.

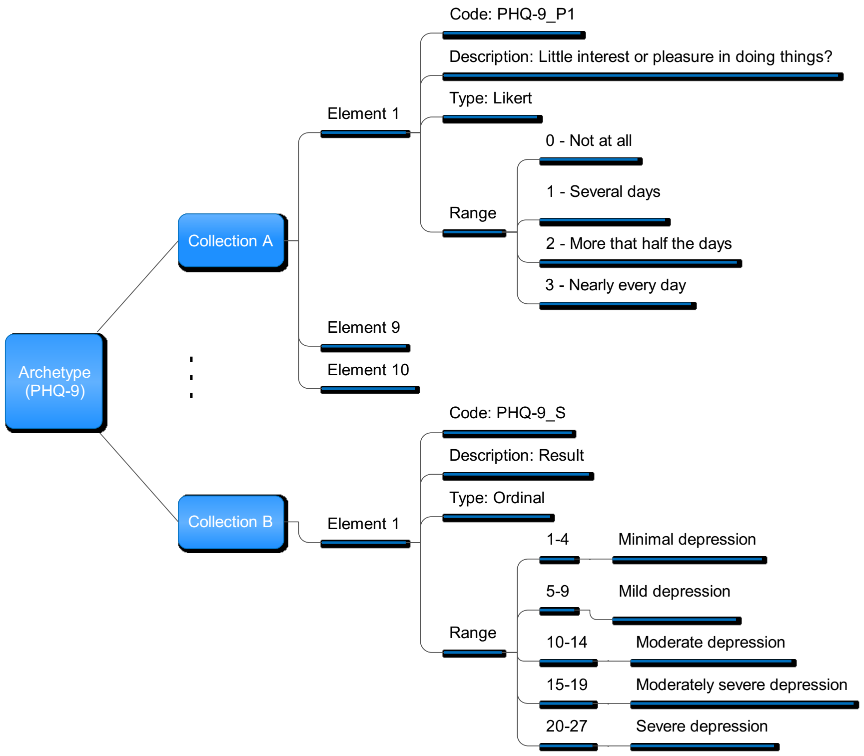


Fig. 3. PHQ-9 representation

All these questions may have answers like not at all, several days, more than half the days and nearly every day. Furthermore, there is a question to check if there are any problems that impact the way the person works, takes care of things at home, or gets along with other people. Finally, the different types of answers are used to determine the partial scores (related with each type of answers) and the total score.

From the total score is possible to identify various clinical situations: i) 1-4 minimal depression; ii) 5-9 mild depression; iii) 10-14 moderate depression; iv) 15-19 moderately severe depression; v) 20-27 severe depression.

To formalize the archetype is first necessary to identify which collections and elements are required. Fig. 3 presents part of the formalization of the archetype PHQ-9 concerning the ten questions and the final score. Therefore, the collection A is composed by ten elements, although only the first one is represented, since the remaining nine elements have similar structures. Collection B formalizes the score.

The information that is stored in relational tables of RIM depends on the resulting archetypes. There is a dynamic field that defines the corresponding data archetype structure in XML format. A complementary XML Schema Definition (XSD) is generated to validate the data field to be persisted (part of this XSD is presented in Fig. 4).

```
<?xml version="1.0" encoding="UTF-8"?> <ns:data
xmlns:ns="http://healthy.oobian.com"
xmlns:xs="http://www.w3.org/2001/XMLSchema-instance"
xs:schemaLocation="http://healthy.oobian.com
http://healthy.oobian.com/data/RIM-PHR-OBSERVATION.phq9.v1.xsd">
      <ns:at0020>
          <ns:value>-1</ns:value>
          <ns:symbol>
              <ns:value></ns:value>
              <ns:definingCode>
                  <ns:terminologyId>
                      <ns:value></ns:value>
                  </ns:terminologyId>
                  <ns:codeString></ns:codeString>
              </ns:definingCode>
          </ns:symbol>
      </ns:at0020>
      <ns:at0021>
```

Fig. 4. PHQ-9 XSD representation

Given the extension of the file we just focus on the value of the first question, which is the default value (-1). Through a Java or Javascript application it is possible to invoke the require services to ensure the information persistence.

5 Conclusion

The proposed information persistence architecture follows an open management information paradigm, where the knowledge model adjusts the information model to the requirements of each specific application domain. This can contribute to the integration of health and social evidence together with personal and AAL generated information.

Considering that was necessary to verify the adequacy of the information persistence architecture, the authors created different scenarios to evaluate if all the information objects required by these scenarios were supported by the developed architecture. These scenarios emphasized the use of PHR, ESR and AAL generated information. The validation has shown that the RIM based information container was able to support a wide range of information structures.

The REST style implementation allows technology independence, and promotes efficiency and scalability. Furthermore, a performance evaluation demonstrated an upper bound of 0.42 seconds to access a set of 20 instances of the PHQ-9 archetype.

In the near future, the full platform will be evaluated in real conditions, supporting a wide range of applications related with health and quality of life domains.

Acknowledgements. This work was supported by COMPETE - Sistema de Incentivos à Investigação e Desenvolvimento Tecnológico, Projectos de I&DT Empresas em co-promoção, under QREN TICE.Healthy.

References

1. Boye, N.: Co-production of Health Enabled by Next Generation Personal Health Systems. In: 9th International Conference on Wearable Micro and Nano Technologies for Personalized Health, Porto, pp. 52–58 (2012)
2. Queirós, A., et al.: AAL information based services and care integration. In: 6th Conference on Health Informatics, Barcelona, pp. 403–406 (2013)
3. ISO/HL7 21731: Health informatics - HL7 version 3 - Reference Information Model - release 1. International Organization for Standardization, Geneve (2006)
4. Dias, A., Queirós, A.: Integração e Continuidade de Cuidados - Plano Nacional de Saúde, 2011-2016. Alto Comissariado para a Saúde, Lisboa (2010)
5. Hayrinena, K., et al.: Definition, Structure, Content, Use and Impacts of Electronic Health Records: A Review of the Research Literature. International Journal of Medical Informatics 77(5), 291–304 (2008)
6. Eichelberg, M., et al.: A Survey and Analysis of Electronic Healthcare Record Standards. ACM Computing Surveys 37(4), 277–315 (2005)
7. Glass, T., McAtee, M.: Behavioral Science at the Crossroads in Public Health: Extending Horizons, Envisioning the Future. Social Science & Medicine 62(7), 1650–1671 (2006)
8. Queirós, A., Silva, A., Alvarelhão, J., Rocha, N., Teixeira, A.: Usability, Accessibility and Ambient-Assisted Living: a Systematic Literature Review. International Journal of Universal Access in the Information Society (2013)
9. Bos, L.: pHealth. In: 9th International Conference on Wearable Micro and Nano Technologies for Personalized Health, Porto, pp. 3–13 (2012)
10. Schloeffel, P.: openEHR Archetypes: Putting the Clinician Back in the Driver's Seat. In: Health Informatics Conference, Sydney (2003)
11. Rigby, M.: Integrating Health and Social Care Informatics to Enable Holistic Health Care. In: 9th International Conference on Wearable Micro and Nano Technologies for Personalized Health, Porto, pp. 41–51 (2012)

12. Teixeira, A., et al.: A New Living Lab for Usability Evaluation of ICT and Next Generation Networks for Elderly@Home. In: 1st International Living Usability Lab Workshop on AAL Latest Solutions, Trends and Applications, Rome, pp. 87–97 (2011)
13. Lenz, L., Kuhn, K.: Towards a Continuous Evolution and Adaptation of Information Systems. International Journal of Medical Informatics 73, 75–89 (2004)
14. Yli-Hietanen, J., Niiranen, S.: Towards Open Information Management in Health Care. The Open Medical Informatics Journal, 42–48 (2008)
15. Lowe, B., et al.: Measuring Depression Outcome with a Brief Self-Report Instrument: Sensitivity to Change of the Patient Health Questionnaire (PHQ-9). Journal of Affective Disorders 81(1), 61–66 (2004)

A Web System Based on a Sports Injuries Model towards Global Athletes Monitoring

Patrícia Macedo[1], Rui Neves Madeira[1,3], André Correia[1], and Marco Jardim[2]

[1] Escola Superior de Tecnologia de Setúbal and
[2] Escola Superior de Saúde,
Instituto Politécnico de Setúbal, Campus do IPS - Estefanilha, Setúbal, Portugal
[3] CITI, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa,
2829-516 Monte da Caparica, Portugal
{patricia.macedo,rui.madeira}@estsetubal.ips.pt,
marco.jardim@ess.ips.pt.

Abstract. This paper presents the project *iReport SportsInjuries*, which is a system with a focus on a Web application directed to sport health professionals, supporting the acquisition, analysis and dissemination of sports injuries information. This software will allow health professionals register and analyze sports injuries among sports populations. The application provides a reporting module that includes tables and charts to individually analyze injuries of a specific sports organization, a specific sport across organizations or a specific athlete, independently of where s/he practices. The system also offers a module for a global analysis, which allows the *iReport SportsInjuries* to obtain the incidence and prevalence values, besides the socioeconomics costs, regarding sports injuries at a national level. In order to offer these features, it is based on a global sports injuries model with the goal of standardizing data related to the sports injuries subject.

Keywords: Health Informatics, Sports Injuries Model, Injuries Incidence and Prevalence, Socioeconomic Impact, Web Systems, Health Systems.

1 Introduction

Participation in sport and physical activity is considered a vital component of an active and healthy lifestyle, reducing the risk of various diseases and contributing to better social and physical performance [1]. However, as more population participates in such activities, we can consequently expect an increased number of sports injuries and simultaneously a substantial expense of financial healthcare resources and, possibly, absenteeism from work [2]. Actually, sports related injuries has been recognised as a major public health problem in many countries [3] and an accurate picture about incidence, prevalence and socioeconomics costs remains a challenge for all communities.

As far as we know, in Portugal there is not any consistent data about the number of sports related injuries at a recreational or even at different competitive levels, as well as, the costs related with amount of sports and working time lost. According to

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_36,

"The Burden of Sports Injuries in the European Union Report", the main objective is to assess the magnitude of sports injuries and their health and economic burden in all European Union countries. Additionally, this document suggests, as pre-requisite to such calculations, the development of a methodology for monitoring sport injuries based on widely available health statistics with the intention to provide data about frequency, severity, and cost of sport injuries as well as about its distribution over the different sports and population groups [4].

Motivated by the aforementioned, the project *iReport SportsInjuries* grew out of a partnership between the Informatics & Systems Department and the Physiotherapy Department of the Polytechnic Institute of Setúbal. It was established that this project should consider two important aspects: i) the development of a sensitive sport injuries report interface, based on a global model, including the most important variables to characterize and assess sports injuries and socioeconomics costs, considering a standardized methodology and terminology based on literature recommendations [5] that would enable data comparison from different sports clinical contexts, sports populations or sports modalities [6]; ii) provide the data model and final reporting through an easy access system that allows users to register, collect, analyze and disseminate data through a quick, safe and effective method [7, 8]. It is also essential to have a flexible system, able to be upgraded or expanded according to the healthcare professionals and sports community usual and regular needs [7].

Regarding these fundamental requirements, we believe that *iReport SportsInjuries* would take advantage from a development into a Web platform. The Web application under development will allow health professionals to individually monitor and manage sports injuries information from sports organization, or even recreational athletes, and simultaneously contribute for an accurate picture at a national level. The project is further described in the following sections. In section 2, an overview of Web applications for sports injuries reports is presented. Section 3 presents the *iReport SportsInjuries* conceptual model that supports the information managed by the system, which is a global model to be followed by all organizations. Section 4 presents the developed prototype, which is a work in progress. We present conclusions and address further work in the last section.

2 Web Applications for Sport Injuries Reports

In literature, we can find similar Web applications for management and monitoring of sports injuries. Most of these Web applications have been developed in the scope of research projects in order to obtain data required to calculate the injury incidence and prevalence levels [9, 10, 11]. Typically, these applications are composed of databases that aggregate injury data gathered from distinct fonts in order to provide a set of indicators [12]. Furthermore, some of these Web applications include athlete injury recovery software, allowing to register the injury and the treatment of each athlete. Some of this kind of software allows exporting data to main applications of monitoring and management sports injuries. As far as we know, the existing systems present a lack of information about the socioeconomics costs related to sports injuries.

NCAA Injury Surveillance System, actually supported by the Datalys Center, is one of the most cited sports injuries system in literature. This system was developed to support the NCAA Injury Surveillance Program [10], which maintains a program since 1982 for intercollegiate athletes. Recently, the system has been improved with an important engine that allows health professionals to export data from some known athletes training injury management software. On the other hand, Reporting Information Online [13] is a Web-based system developed at the Children's Hospital in Columbus, Ohio, conceived to register injuries on national bases. This software was one of the first systems allowing the registering of time loss due to the injury. One potential weakness reported on literature arises from the fact that this system was not integrated with NCAA Injury Surveillance System, resulting in potential duplication of registers in the USA.

Another example is the Fairfax County Public School Injury Surveillance Database [14]. This system was implemented to monitor sports injuries in the scope of sports activities in public high schools of Falls Church, in Virginia, USA. This system has proved to be robust enough to support more than a decade of use. However, it was conceived specifically to school sport, therefore it is not able to support the registration of injuries in a wider panorama that includes sports organizations and sports events. SIMS [15] is an advanced injury documentation and management software application designed to report athletes' participation status for training sessions or competition. It is a non-freeware software and it is conceived for the USA context. Another work is NExTT [16], an injury tracking software directed to athletic departments, which provides Web and mobile accesses directed to a set of features related to management of athletes' training sessions. This software also allows the management of athletes' data injuries as well as clinical history and training sessions. It has an engine to export data to some known surveillance systems.

Additionally, a common problem identified on almost all national injuries report systems is the data reliability [17]. Data reliability depends of several factors: (i) how concepts are understood by users that register the injury; (ii) how the existence of non-duplicate data or nonsense data is guarantee; (iii) how the access to change data is limited to authenticated users. Therefore, the design of a Web application software for sport injuries reports has to consider the implementation of mechanisms to guarantee that provided injuries information is reliable. None of the analyzed systems provide an easy and integrated way to individually support health professionals to report sports injuries at our national level. We believe that *iReport SportsInjuries* will represent a step forward for all Portuguese health professionals, and society at large, due to the current lack of information on sports injuries and their socioeconomics costs.

3 iReport SportsInjuries' Conceptual Model

Injury surveillance consists of an ongoing standardized collection of data describing the occurrence of, and risk factors associated with, an injury in large population groups. It provides the who, what, where, when, and how of information, which is the basis of descriptive epidemiology. Most of the times, data from injury registration is broad in scope, often missing uniformed detail, and sometimes lacking context [12]. This information needs to be tempered with a clinical mindset in order to be useful for distinct research approaches and to be used to answer important health related questions. Therefore, the design of the surveillance system's database is a crucial issue to the overall performance of the system.

We propose a global conceptual model for the sports injuries reports in order to guarantee that all data objects required by the database are completely and accurately represented and distinct stakeholders of the project can easily review and verify data. The UML Diagram that specifies the model is presented in Figure 1.

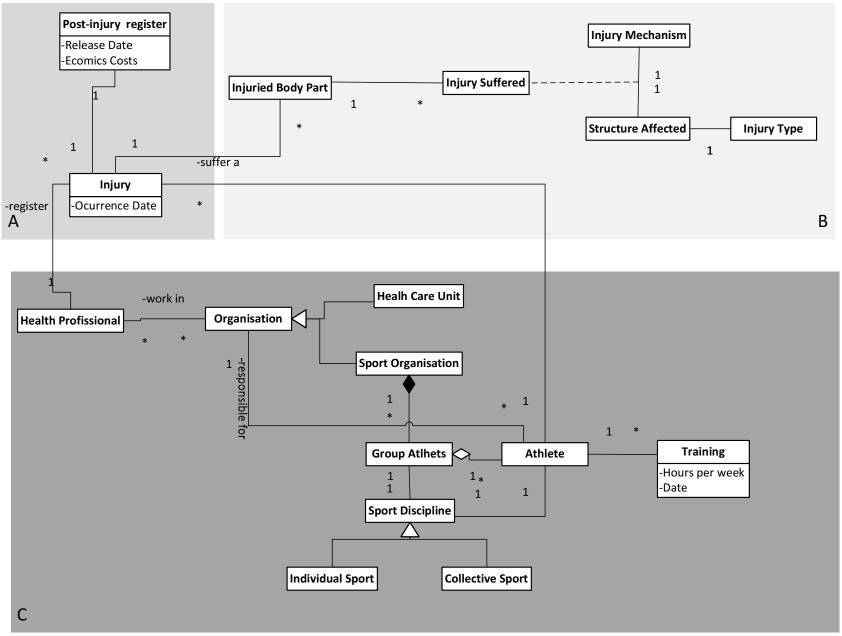


Fig. 1. UML Diagram of *iReport SportsInjuries'* Conceptual Model

Three main groups of concepts can be identified in the proposed conceptual model:

- A - Injury Registration: Includes the injury data and the pos-injury data concepts. Both registers include the date of the occurrences. The dates are used to compute the period of time in which the athlete is inactive. The pos-injury register also includes information about the economics costs related to the injury.
- B – Injury Characterization: Defines the specification of the injury, which comprises the specification of the injured body parts, and for each body part is defined the structures affected (muscle, ligament, bone, etc.); for each structure is specified the respective injury mechanism (luxation, concussions, etc.), and the injury type.
- C - Stakeholders: This sub-model is responsible for the characterization of the main entities related to the injury: (i) the athlete that suffers the injury; for each athlete is also characterized the training schedule (hours per week, days per week, etc.) and the sport discipline. (ii) the health professional that registers the injury and treats the athlete. (iii) the organization in which the athlete is treated; in this

case we can have two situations: the professional is associated to a sport organization, and in this case s/he follows a specific group of athletes that practice a specific sport discipline, or the professional treats the athlete in a health care unit without being associated to a specific sport organization.

4 iReport SportsInjuries System

The *iReport SportsInjuries* project has its origin in a previous project called SMLD, started in 2010 inside the Physiotherapy Department of our Institution. This project developed a survey to characterize injuries among Portuguese sports population. The survey was composed of two main sections: the sport injury registration and the socioeconomic impact registration. The first is composed of several dimensions including the athlete, the sport and the injury characterization. The second is centered on the injury socioeconomics consequences, such as the amount of sports, the working time spent and the monetary costs. The survey was implemented using the Google Docs platform whereas data was obtained through aggregated Excel facilities. The SMLD was validated and is being used in sports injuries prevention projects associated to football and futsal teams from Setúbal and Lisboa.

As mentioned before, the main goal of the *iReport SportsInjuries* project is to provide a secure, trustful and simple way to manage and analyze sports injuries data at a national level. In order to encourage injury data registration, a set of software services for injury reporting is provided to health professionals. These services will allow the health professionals to access anywhere, at any time, the injury history of their athletes, and additionally obtain a set of indicators about their athletes' injuries. In order to fulfill these features, the system has to support three types of users: the Health Professional (e.g., physiotherapists), the iReport Manager, and the Guest (which should apply to be accepted in the application as heath professional, waiting for the iReport Manager's approval), as documented in the UML user case diagram shown in Figure 2.

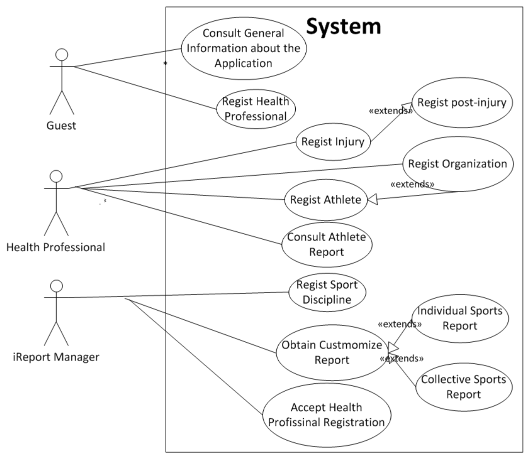


Fig. 2. iReport SportsInjuries System's User Case Diagram

The system supports the registration of a sport injury by a sport health professional (see Figure 3, left), where the registration comprises the two set of questions specified in the SMLD survey tool. The system guarantees the registration of information by authenticated users in order to fulfill data security requirements. Sport health professionals can consult the history of their athlete's injuries and obtain a set of indicators about injury incidence and prevalence (see Figure 3, right).

The iReport Manager profile gives access to unidentified data (the name of the athlete and his physiotherapist is encoded) about all the injuries registered on the system. The system also provides an analysis tool that allows project administrator to obtain customized reports (e.g., period of time, team sports discipline or an individual sport can be selected) on sports injuries incidence and sports injuries prevalence at a national level, since the system implements a global model that ends being followed by everyone, standardizing data.

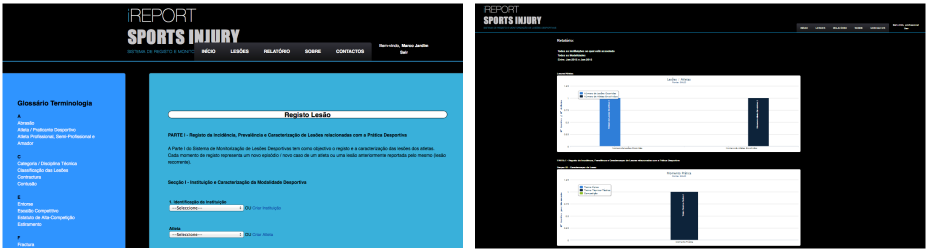


Fig. 3. *iReport SportsInjuries* Web Application' screens: (left) Injury Registration; (right) Report Injuries

The system was implemented following a thin-client server architecture. There is a central database based on SQL and the server logic implements the MVC pattern using Microsoft .net framework 4.

5 Conclusions and Future Work

The *iReport SportsInjuries* Web application is already available for a set of sports health professionals that are testing it. It is expected that this Web application satisfies the identified concrete needs concerning sports injuries information and their socioeconomics costs. This application will be available for free for all sports health professionals in Portugal. Therefore, it is expected that a massive use of the system will provide accurate data about sports injuries in diverse clinical contexts and sports organizations.

A mobile application for injury registration is planned to be developed. This new interface for mobile devices will allow sports health professionals to register injuries directly in their smartphones or tablets.

References

1. Beijsterveldt, A.M.C., Krist, M.R., et al.: Effectiveness and cost-effectiveness of an injury prevention programme for adult male amateur soccer players: design of a cluster-randomised controlled trial. Injury Prevention 17 (2011)
2. Cumps, E., Verhagen, E., et al.: Injury rate and socioeconomic costs resulting from sports injuries in Flanders: data derived from sports insurance statistics 2003. British Journal of Sports Medicine 42, 767–772 (2008)
3. Timpka, T., Lindqvist, K., et al.: Impact of social standing on sports injury prevention in a WHO safe community: intervention outcome by household employment contract and type of sport. British Journal of Sports Medicine 39, 453–457 (2005)
4. Kisser, R., Bauer, R.: The burden of sports injuries in the European Union. Research report D2h of the project Safety in Sports, Vienna, Austrian Road Safety Board (2012)
5. Finch, C.F., Ozanne-Smith, J., Williams, F.: The feasibility of improved data collection methodologies for sports injuries. Monash University Accident Research Centre (1995)
6. Meeuwisse, W.H., Love, E.J.: Athletic injury reporting. Sports Medicine 24, 184–204 (1997)
7. Zemper, D.: Epidemiology of Athletic Injuries.: ACSM's primary care sports medicine. Wolters Kluwer Health (2007)
8. Junge, A., Engebretsen, L., Alonso, J.M., Renström, P., Mountjoy, M., Aubry, M., Dvorak, J.: Injury surveillance in multi-sport events: the International Olympic Committee approach. British Journal of Sports Medicine 42, 413–421 (2008)
9. Emery, C.A., Meeuwisse, W.H., Hartmann, S.E.: Evaluation of risk factors for injury in adolescent soccer implementation and validation of an injury surveillance system. The American Journal of Sports Medicine 33, 1882–1891 (2005)
10. Klossner, D., Corlette, J., Agel, J., Marshall, S.W.: Data driven decision making in practice: The NCAA injury surveillance system. New Directions for Institutional Research 2009, 53–63 (2009)
11. Fuller, C.W., Laborde, et al.: International rugby board rugby world cup 2007 injury surveillance study. British Journal of Sports Medicine 42, 452–459 (2008)
12. Mechelen, W.V.: Sports injury surveillance systems. Sports Medicine 24, 164–168 (1997)
13. Darrow, C.J., Collins, C.L., Yard, E.E., Comstock, R.D.: Epidemiology of severe injuries among United States high school athletes 2005-2007. The American Journal of Sports Medicine 37, 1798–1805 (2009)
14. Hinton, R.Y., et al.: Epidemiology of Lacrosse Injuries in High Schoolar Aged Girls and Boys A 3-Year Prospective Study. The American Journal of Sports Medicine 33, 1305–1314 (2005)
15. Lehr, M.E.: Plisky, et al.: Field expedient screening and injury risk algorithm categories as predictors of noncontact lower extremity injury. Scandinavian Journal of Medicine & Science in Sports (2013)
16. Solomon, D.J., et al.: 2 Team Physicians Corner (2012)
17. Thygerson, A.L., Thygerson, S.M., Thygerson, J.S.: Injury Prevention: Competencies for Unintentional Injury Prevention Professionals. Jones & Bartlett Publishers (2008)

ARTE: An Embryo Quality Assessment Tool

João Silva[1], Paulo Fazendeiro[1], Fernando J. Prados Mondéjar[2], and Soraia Pinto[3]

[1] Instituto de Telecomunicações (IT), Department of Informatics,
University of Beira Interior 6200-001 Covilhã, Portugal
phargoth@gmail.com, fazendeiro@ubi.pt

[2] IVF Laboratory, Hospital Madrid-Montepríncipe 28660-Boadilla del Monte – Madrid, Spain
fernandojprados@gmail.com

[3] Maternidade Doutor Alfredo da Costa, Rua Viriato 1069-089 Lisboa, Portugal
soraiapinto.pt@gmail.com

Abstract. This paper proposes an assessment model of putative embryos for in vitro fertilization (IVF) based on a triangular norm. One of the most common difficulties of IVF treatments is multiple pregnancy. Therefore the number of embryos for transfer is of paramount importance considering the need to reduce the incidence of multiple births without compromising overall pregnancy rates in fertility treatments. Consequently the selective embryo transfer is recommended to optimize efficacy and safety outcomes. The embryo evaluation is of enormous relevance, since it directly affects the success of different techniques used in assisted reproductive technologies (ART). The gathering of all the information needed to embryo evaluation, as well as software that can serve as aid in the decision is of great importance. The tool herein presented accomplishes these two objectives. The analysis of the requirements of the assessment process has resulted in a flexible data model, used in the presented prototype, supporting the selective embryo transfer decision-making process.

Keywords: Embryo Assessment, Decision Making, Software Engineering, Assisted Reproductive Technologies.

1 Introduction

Worldwide, it is estimated that one out of every six couples experience some form of infertility problem at least once during their reproductive lifetime [1], and it is also estimated that 11.3% to 26.4% of women, under 50, in more developed countries have a lifetime prevalence of infertility [2]. In 2009 the Portuguese Society for Reproductive Medicine (SPMR), conducted a joint study with Keypoint, aiming to quantify and characterize the type of infertility existent in Portugal. According to this study there where at the time 116.630 couples facing infertility[1] problems, and 7.9% of women from 25 to 44 have problems in conceiving [4].

[1] According to the World Health Organization, Sterility is defined as the "inability to fertilize the ovum with a spermatozoan" while Infertility refers to "the inability to ensure that the fertilized ovum develops sufficiently for the birth of a viable child" [3]. In this paper we will use the word "infertility" to describe both situations since the type of diagnosis associated with reproductive problems will not have an influence on the current work.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_37, © Springer International Publishing Switzerland 2014

In recent years the number of couples who resort to Assisted Reproductive Technologies (ART) has increased considerably [5], consequently the amount of in vitro fertilization (IVF) clinics has enlarged worldwide and the need to improve IVF techniques has become a major endeavor.

The Assisted Reproductive Technologies (ART) grew in importance medically, socially and economically. The number of units available keeps increasing both in the public and private sectors, greatly due to the considerable growth of the number of couples who resort to these techniques in recent years. Although this is a field with enormous potential there are gaps in some areas including that of information technology (IT).

The available ART units are multidisciplinary units that cover three different fields: medical, nursing and laboratory. Despite their importance and complementary character, the laboratory component has proven to be of particular importance [6, 7]. After infertile couples are diagnosed most of the process of becoming pregnant is largely dependent on the embryology lab.

Several ART techniques are applied in the embryology lab, namely *In Vitro Fertilization* (IVF) [8] and *Intracytoplasmic Sperm Microinjection* (ICSI) [9], [10]. Daily view of oocytes and embryos under an inverted microscope is required in both techniques for a period ranging between 3 to 5 days. Several morphological parameters are assessed during these observations in order to determine the embryo quality and implantation rate. These parameters are later used to determine which of the embryos are transferred, frozen or destroyed. Evaluation should be continuous, systematic and reducing the variability between observers [11].

Since it directly affects the chances for success, the different techniques used in reproductive medicine embryo assessment are of enormous importance. One of the main problems concerning embryo assessment is the lack of a single evaluation criterion, although several efforts have been made to obtain a consensus [12]. The need to compare results makes it imperative to achieve some sort of standardization.

A software that allows its user both the possibility of collecting and storing all needed information as well as assisting in the process of decision making would prove not only to be a valuable resource, but also it would help to fill a gap in terms of specialized software in the field of human embryology.

This paper is organized as follows. Section 2 presents a brief review of related works. In Sec. 3 it is proposed a model for the embryo quality assessment. Section 4 presents the functional requirements and the data conceptual model of the application. In Sec. 5 are referred some major design choices for the developed prototype and finally Sec. 6 draws the conclusions and discusses future work.

2 Background

The available software consists mainly of small applications usually created from generalist tools, like MSAccess, Filemaker, or even MS Excel files, for personal (laboratory) use. Moreover, the existing commercial programs are “heavy” and best suited for the medical component. These are programs designed for patient management, assisting in the patient diagnose, medical history, gamete bank management, lab data storage and produce several kind of reports. The fact that these are multifaceted programs makes them of little use in the laboratory, being used more often than not as a simple tracker of the results obtained in the several cycles that were preformed during a period of time.

Nevertheless there are good integrated equipments commercially available. Arguably the most distinguishable one is the EmbryoScope™ Embryo Monitoring System, developed by Unisense FertiliTech, a built-in gas incubator that is controlled by an embedded PC [13]. This system allows uninterrupted observation of embryo development in a stable controlled environment. It features a camera that collects images of the embryos. The software associated, the EmbryoViewer™, allows the review, annotate and compare the development of selected embryos using data files acquired by the EmbryoScope; it also allows a retrospective analysis of the embryo development, and helps the embryologist to easily assess several parameters for embryo selection [14]. Unfortunately these are costly systems, too expensive for most national laboratories or even considering an Iberian scale.

There is also software that, using computer analysis of images of planes at various depths within the sample (z-stacks) allows the observation of the 3-dimensional morphology of an embryo, particularly a 4-cell. This Image Analysis software operates on a series of digital images acquired at different focal planes (acquired either by means of a common microscope equipped with a digital camera or by specialized hardware) and finds the boundaries of blastomeres as closed contours with regular shapes [15].

The Artemis™ from Medialogic, is a software that helps to manage, track and report several reproductive techniques. More turned to the aspect of reporting production and storage of data related to the patients and the embryo cycle itself [16].

Besides these software there are also studies being made, on the prevention of multi-pregnancy, taking in consideration the embryo quality, as the main objective is always the ability to select the best embryo for transfer [17,18], thus avoiding the transfer of more than one embryo, that could lead to an undesirable multi-pregnancy. In some countries the current legislation is very strict on transferring and freezing embryos, for example in Germany no cleavage embryos can be frozen, so all embryos that reach this stage must be transferred, this leads of course to a high rate of twins and triplets [19].

The main objective of our research is to build an application that using parameters previously defined by an administrator, performs an a-priori embryo classification (predictive value). The purpose will be the elective selection for embryo transfer or cryopreservation. Additionally, it must allow the comparison of different existing classifications, providing insight on which classification produces more accurate results on certain situations. In addiction it should also allow data analytics studies based on actual classifications.

By allowing the parameters to be managed by the administrator a major obstacle is overcome, the constant evolution in the scientific research that brings new parameters to light, and questions the importance of other parameters previously gathered. Also different evaluations can take in consideration different parameters and even with different correlation and/or importance between them.

Another objective is to deploy a tool that will help to uniform the parameters, their more common names, as well as their evaluation.

Summing up we can say that this work empowers the administrator with the tools to tune the evaluation, creating the parameters and defining them, as a programmer creates the set of properties on classes. Then he/she will also be able to "program" the way these parameters will interact to return a result, as a method. Having this liberty the administrator will have no need to ask for an upgrade every time a change happens.

3 Algorithm to Calculate the Quality of the Embryo

One of the key concerns of this work is the elicitation of a formula that will allow the evaluation of the embryo and give it a percentage score, this formula must be flexible enough so that it will not be dependent of any parameter in concrete; this is because all the parameters will be inserted by the administrator of the application, allowing this way a flexible use that will not require the intervention of a programmer to rebuild the project just because a new parameter must be inserted.

Another point is that the evaluation on a certain day of a certain embryo must be conditioned by the previous evaluations of that same embryo. In the following it will be presented a plausible mathematical formula that is flexible enough to be used on such vague environment, but on the same time is precise enough to return reliable results.

In order to simplify the explanation let us start by assuming that the ASEBIR recommendations [20] are enforced making one reading per day, recording the parameters that are indicated in the examination of that specific day. The readings are taken on day 1 after 16-19 hours of the insemination; on day 2 after 44-47 hours; on day 3 after 67-71 hours, etc.

All parameters will have an importance, and the different values that the parameter can take are also discriminated, the system administrator defines all these values.

Let's take in consideration all the parameters P, that must be evaluated on a given day d; and that a certain parameter p on that group has an importance i_{pd}, already defined, so we can say that the relative weight of that parameter p in the score for the day is:

$$w(p,d) = (i_{pd} / \textstyle\sum i_j) \text{ for all } j \text{ in } P. \tag{1}$$

i.e. the relative weight of the parameter is equal to its importance divided by the sum of the importance of all the parameters of that day; this way the relative weight of a parameter will be a value between 0 and 1.

Each parameter p will have a set A_p of associated answers (or range of answers in case of numerical answers) each one of those answers will have a certain grade G_{ap}. The normalized score of each answer, $S(p, d)$, is constrained to the unit interval [0, 1] so that the maximum score possible in any day is lesser or equal to 1.

The score of a parameter p, in a given day d, is obtained through the following formula:

$$DS(d) = \sum S(p, d).\ w(p, d) \text{ for all } p \text{ in } P. \tag{2}$$

The final score must be obtained by considering all the days, but not forgetting that the score of a previous day must always affect the score for the next day.

$$\text{FS} = \prod_{d} DS(d) \tag{3}$$

Where d is a placeholder for the elements of the set of days passed after the insemination.

Notice that in this way it is assured that the score on a given day of a given embryo will never be higher than the score of that same embryo the day before, as is a consensual assumption that an embryo is not likely to evolve to become a better embryo than what was expected based on previous evaluations. Instead of the product norm (the standard semantics for strong conjunction) another triangular norm, such as the minimum norm, could be used to compose a given day-score with the previous ones.

This composite score is returned as a value between 0 and 1 that can be easily converted to any classification table.

4 System Requirements

For a question of organization we divided the project in 3 major modules, these modules are relatively independent, they each are focused on a field of work, although the last module depends on the other two. The first of these modules is an administration module; where it will be possible the management of the users of the application and organize them by laboratories. The second module, model management, here are defined and managed the several parameters, organized them by stages, and organized this stages by readings; also is where it is defined the several types of externals that may be used and interact with the embryo. The last, but not least is the laboratory module, where the real data is managed, the couples, the referred processes and all the information that is related with it, also the medications that will be used and the externals elements.

4.1 Administration Management

The administration is composed by the class relative to the Personnel, these can be Embryologists or Physicians; they can also be registered as users or not. The personnel are associated with a Laboratory. Both of them can have addresses and registered contact. There is a class, Entity, which gathers the common data between the Personnel and the Laboratories. Also, there are several degrees of authorization that a User, registered personnel, can have. These authorizations will indicate what operations a User can or cannot perform.

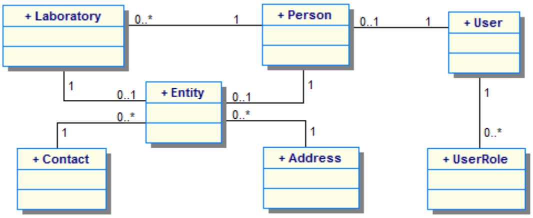


Fig. 1. Administration management class diagram

4.2 Model Management

The model encodes the theoretical part of the application. It is composed by the Parameter class the Stages where this parameter can be evaluated, and the several possible results. Also it has the definitions of the readings that can be made – what Stages will be examined, and what parameters in these stages will be taken into consideration. Since the several parameters can be used in several stages, and the several stages can be used in several readings the association between them is a many-to-many relationship.

There is also the external types class, where it is be defined what types of externals interacting with the embryo that will be registered (i.e. incubator, CO2 Gas, etc.)

The Category is used to register different categories that an embryo may assume, these are status that can be used to register the step where the embryo is and/or its destiny.

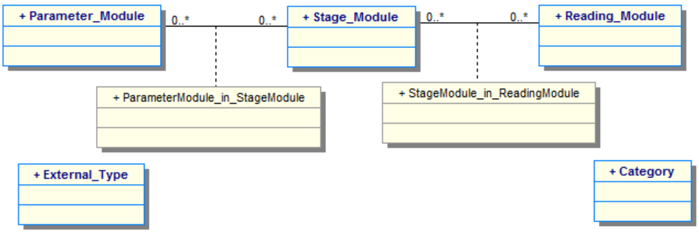


Fig. 2. Model management class diagram

4.3 Laboratory Management

This is where the main data is registered. The Couple class has addresses and contacts therefore, likewise the Personnel and Laboratory, the Couple is also associated with the Entity class. These couples are of course registered in a certain laboratory. These couples have processes that have Embryos associated to them, there is also a class to classify and register the sperm that will be used. These processes have a collection of embryos associated to them, which in turn have a collection of stages. These stages also have a collection of resulting parameters.

These are the results that will be taken into consideration to calculate the evaluation of the Embryo. In order to relate the result of the evaluation with the pregnancy there is also a class that processes the data related to the transfer and the evolution of the pregnancy. A process can imply several transfers because there is the possibility of cryopreserved embryos to be transferred later; also a single transfer may imply more than one embryo in it.

The Sperm is in a different class as a possibility to be studied in future. In this section there is also place for an external class that will register the specific externals that exists and, may be used, in each laboratory.

There is also a class to register the several medication that can be used in the process, this will allow an easier way to fill the information and reduce the chance of the same medication be inserted with different nomenclature.

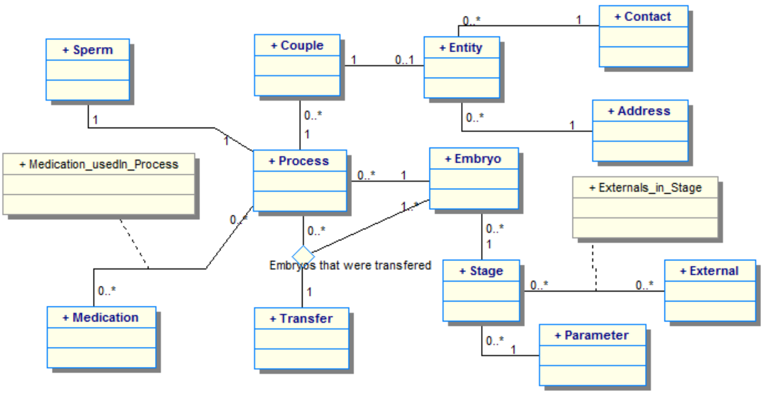


Fig. 3. Laboratory management class diagram

4.4 Evaluation

The evaluation should not be an embedded part of the solution, as one of the objectives is the option to create evaluation rules, and be able to share them easily with other installed versions. So the option was to create a template, for json files, that must be followed by all evaluation files. The files created this way can be deserialized by the application and be used to give the result of a certain reading, or set of readings, of an embryo, or group of embryos.

5 Implementation

The implementation of the application has followed the modulation that was defined on the analysis section. It was decided that, in its current release, it would be a local application, with the possibility of a central database. The main concern during the implementation was always the scalability of the application due to two main reasons: Firstly, as the field of study is very complex and with several sub-fields, each one of them with its own nuances and specifications, and maybe its own application if deemed appropriate. All these sub-fields are determinant for the final result, a successful pregnancy. Secondly, it is expected that individual components of the application may be reused on future works.

5.1 Database

The database has to respect the laws of standardization, and to be constructed in a way that will allow scalability for future implementations on aspects that are not being addressed at the present moment, for example, a more thorough study of the oocyte or the sperm, or recording more information about why a couple should use ART.

The database, of course, must be implemented respecting the rules of normalization trying to reduce the redundancy and dependency as much as possible, isolating the data so that the basic operation of creation, modification and removal of a field can be made on a single table and then, using defined relationships, propagated through the rest of the database.

5.2 Solution

This solution is implemented as an Object-Oriented approach following a basic 3 Tier architecture. Data Access Layer: that should be implemented following the CRUD methods directives (Create, Retrieve, Update, Delete) and that is responsible to communicate with the Database. Business Layer: the heart and soul of the solution, where most of the operations must take place, and where are defined relations between the different classes and how they interact. Presentation Layer: responsible for the communication with the users

These 3 layers must be implemented independently, this way any change or upgrade on a layer will have a minimum impact on any of the others layers, also it will allow that an entire layer can be reused on future developments of the project without need of re-implementing the architecture, or mixing implementations that may not be directly related; for example the implementation of several User Interfaces, or a need of change in the used database engine.

6 Conclusion

The selective embryo transfer is necessary in order to optimize efficacy and safety of the different techniques used in assisted reproductive technologies. The accurate embryo evaluation and the up-to-date laboratory workflow management are essential parts of a successful outcome for this task.

This work produced, so far, a tool that allows the administrator to determine what should be studied during the several observations of the embryo development and record these data. After that, these data are available to be analyzed; also it already returns the grade of the embryo(s) according to the ASEBIR criteria [20], even in the cases where the available information is incomplete (not all the necessary parameter measurements are collected). Additional evaluation criteria are also being studied and their interplay being modeled.

The distinguish factor that is being aimed is to develop a tool that is ready to keep up with the evolution and continuous change on a field of active research and constant innovation. Also independence from a particular criterion will allow the authorized users to choose and tune-up their favorite evaluation criteria.

The suggested evaluation model, based on a triangular norm, is plausible for the accurate assessment of the embryo quality. The implemented algorithm produces evaluation results heavily dependent on the somewhat subjective importance weights provided by the domain specialist. As a future work, based on the initial values, information collected from the different evaluation criteria, and, more important, real life results, we intend to apply machine learning techniques that based on the initial values, information collected from the different evaluation criteria, and, more important, real life results, it can tune the current algorithm (or a reformulated one) correcting the parameters' importance and its answers' grading system.

Another possible improvement is to extend the embryo assessment process with the inclusion of the data related with the sperm used and the oocyte quality, whenever available. We believe that this information can help to increase the accuracy of the evaluation process.

From the functional side, laboratory management of the external elements can also be integrated on this tool, helping to control the stock of consumables and disposables.

It could also be interesting the implementation of several User interfaces for different levels of access, with different objectives of use of the application (web or mobile devices).

Acknowledgement. This work was partly supported by Fundação para a Ciência e Tecnologia (FCT) under the project PEst-OE/EEI/LA0008/2013.

References

1. ART fact sheet, `http://www.eshre.eu/Guidelines-and-Legal/ART-fact-sheet.aspx` (retrieved from HYPERLINK)
2. Boivin, J., Bunting, L., et al.: International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. Human Reproduction 22(6), 1506–1512
3. Rosa, P.: World Health Organization. Regional Office for Europe and International Children's Center. Family health and family planning. Paris, International Children's Centre, and Copenhagen, World Health Organization, Regional Office for Europe (1979)
4. Carvalho, J.L., Santos, A.: Estudo Afrodite – Caracterização da infertilidade em Portugal, `http://static.publico.clix.pt/docs/sociedade/AfroditeInfertilidade.pdf` (retrieved from HYPERLINK)
5. National (2011 - 2009) and European (2009 - 2007) reports of ART activities, `http://www.spmr.pt/pma.php#pma` (retrieved from HYPERLINK)
6. Sunde, A., Balaban, B.: The assisted reproductive technology laboratory: toward evidence-based practice? Fertility and Sterility 100(2), 310–318 (2013)

7. Gardner, D., Weissman, A., Howles, C., Shoham, Z. (eds.): Textbook of Assisted Reproductive Techniques, 4th edn. Laboratory Perspectives, vol. 1. CRC Press (June 2012)
8. The IVF Informant: What is IVF? - Molly Robinson, `http://www.ivf-info.org/what-is-ivf` (retrieved from HYPERLINK) (November 25, 2012)
9. The IVF Informant: What is ICSI? - The IVF Informant, `http://www.ivf-info.org/what-is-icsi-1` (retrieved from HYPERLINK) (April 28, 2012)
10. The IVF Informant: Variations on IVF - GIFT and ZIFT - Molly Robinson, `http://ivf-info.org/variations-on-ivf-gift-and-zift` (retrieved from HYPERLINK) (November 27, 2012)
11. Bączkowski, T., Kurzawa, R., et al.: Methods of embryo scoring in *in vitro* fertilization. Reproductive Biology 4(1), 5–22 (2004), `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.112.4553&rep=rep1&type=pdf` (retrieved from HYPERLINK)
12. ALPHA Scientists in Reproductiove Medicine and ESHRE Special Interest Group of Embriology: The Instambul Consensus workshop on embryo assessment: proceedings of a expert meeting. Human Reproduction, 1–14
13. Embryoscope® Time-Lapse System, `http://www.fertilitech.com/en-GB/Products/EmbryoScope-reg-Time-Lapse-system.aspx` (retrieved from HYPERLINK)
14. Embryoviewer® Software, `http://www.fertilitech.com/en-GB/Products/EmbryoViewer-reg-Software.aspx` (retrieved from HYPERLINK)
15. Giusti, A., Magli, M.C., et al.: Observing the 3-dimensional morphology of 4-cell embryos using computer analysis of image z-stacks. Human Reproduction 26(suppl. 1), i62–i63; Abstract Book of the 27th Eshre Annual Meeting – O158
16. Artimis on Medialogic website - HYPERLINK, `http://medialogicindia.com/artemis/`
17. Wathlet, S., Adriaenssens, T., et al.: New candidate genes to predict pregnancy outcome in single embryo transfer cycles when using cumulus cell gene expression. In: Fertility and Sterility (August 2012)
18. Stylianou, C., Critchlow, D., et al.: Embryo morphology as a predictor of IVF success: an evaluation of the proposed UK ACE grading scheme for cleavage stage embryos. In: Human Fertility, Cambridge, England (March 2012)
19. One at a time – better outcomes in fertility treatment, `http://www.oneatatime.org.uk/372.htm#6` (retrieved from HYPERLINK)
20. Ardoy, M., Calderón, G., et al.: ASEBIR criteria for the morphological evaluation of human oocytes early embryos and blastocysts, 2nd edn., pp. 43–49

Health System Efficiency and eHealth Interoperability – How Much Interoperability Do We Need?

Karl A. Stroetmann

Empirica Communication and Technology Research,
Oxfordstr. 2, 53111 Bonn, Germany
karl.stroetmann@empirica.com

Abstract. The generic concept of interoperability is briefly reviewed, before describing the approach taken in defining eHealth interoperability, based on a reflection of the diverse needs for exchanging patient, other healthcare and health system data with the aim to improve the overall quality and efficiency of healthcare provision. Next, the method applied for collecting global good practice evidence from national or district eHealth platforms is outlined. The information collected led us to revisit the concept of eHealth interoperability levels; a modified concept is presented identifying four distinct domains, and the identification of three levels for the data and access domain. The results are discussed under the perspective of 'Who needs why how much interoperability?' - including the impact of healthcare provider incentives, size of patient population, and the benefits from interoperability on the concrete instantiation of an interoperability platform. Finally, conclusions are drawn on the role of eHealth interoperability for better healthcare.

Keywords: eHealth, interoperability, Africa, global good practice, benefits, framework.

1 Analysing Interoperability in a Global Context

The research upon which this paper is based was performed as part of the ISAES – Interoperable eSystems for Africa Enhanced by Satellites – Study [1]. It gathered global empirical evidence, described, and analysed interoperability opportunities and challenges towards designing a generic eHealth platform for sub-Saharan Africa. The final goal is to initiate and support the implementation and sustained operation of national or district eHealth platforms facilitating the access and exchange of patient and other health data towards delivering improved healthcare [2, p. 6]. This is the guiding vision of the Satellite-Enhanced Telemedicine and eHealth for sub-Saharan Africa (eHSA) Programme of the European Space Agency [3]. Towards this end, two further horizontal studies on (1) governance and (2) regulatory issues have been performed, and a fourth study on sustainability and financing aspects has started early in 2014.

The rationale for this work is the observation that "in spite of the potential benefits offered by space in supporting applications in the field of health, until today the health sector has seen neither significant utilisation of space technologies nor systematic

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_38, © Springer International Publishing Switzerland 2014

analyses of needs for space assets. Aside from cost considerations, this may be due to the health professionals' limited awareness of space capabilities on one side, and by limited understanding of user needs and regulatory issues by the space actors on the other side. Without a comprehensive understanding of the healthcare domain, the chances of a breakthrough in the utilisation of space assets are very limited. In this situation, contributing to paving the way for eHealth and telemedicine service penetrations in developing countries is a significant opportunity for promoting space infrastructure for societal benefit" [2, p. 6].

The following sections will first describe the approach taken in defining eHealth interoperability, based on a reflection of the diverse needs for exchanging patient and other healthcare or health system data, and the method applied for collecting global good practice evidence. This led us to revisit the concepts of interoperability levels and domains. The results are then discussed under the perspective of 'Who needs why how much interoperability?' - including the impact of healthcare provider incentives, size of patient population, and the benefits from interoperability. Finally, conclusions are drawn on the role of eHealth interoperability for better healthcare.

2 Defining eHealth Interoperability

2.1 Reviewing Interoperability Concepts

To provide a foundation for defining an appropriate eHealth interoperability framework, a variety of generic interoperability (IOp) definitions were reviewed [4-9] because how the conceptual space of IOp is defined and structured exerts guiding influence on subsequent discussions. Also, eHealth IOp has to be seen in a wider context, taking into account also considerations, e.g., from the eGovernment field [10]. It turned out that the term interoperability is often a fuzzy one - mostly referring to data exchange among technical systems or "application entities," and less to the exchange of information among persons or organisations (respecting business and workflow processes, procedures, and the cultural context of organisations).

Almost all approaches proceed in one way or another to identify functional and/or area-related IOp distinctions and create appropriate sub-divisions of IOp levels as needed in the respective context. But these concepts also mostly focus on technical aspects. When discussing semantic challenges, this does usually not concern human information exchange and understanding, but "machine" interpretation and handling of data.

This clearly necessitates a more comprehensive understanding and definition of IOp in the healthcare and public health system field, where the collaborative provision of healthcare services by health professionals and health workers is at the core of the subject. To allow for a common (European) understanding and focused, goal-oriented discussions, the European Commission sponsored i2-Health study developed and defined a European Union eHealth Interoperability Framework [11], which was adopted by the European Commission and its Member States [12, pp. 24 ff].

This European eHealth Interoperability concept and framework has gained international recognition [13-15]. It was developed in more detail and expanded into the eHealth architectural domain by the European Thematic Network "CALLIOPE - Creating a European coordination network for eHealth interoperability implementation" [16], and influenced ISO TC 215's "Capacity-based eHealth architecture roadmap" [17].

It is perhaps surprising that hardly any national eHealth strategy document – and consequently also the case studies - conceptualizes and discusses at any detail core issues of IOp, like what it really means, why it may be desirable, what degree of IOp in which health system domain should be accomplished, what benefits to expect from the sometimes very substantial investment.

2.2 Defining Interoperability

In its most basic form,

> *interoperability refers to the ability to communicate and exchange data and information such that a common understanding of their meaning is possible.*

Such sharing of information and knowledge has been a basic precondition for the social development of humankind. In the age of ICT systems, IOp tends to be associated with

> *the ability of telecommunications and digital systems - and the processes they support - to exchange data and to enable the sharing of information and knowledge.*

This is a relatively open definition - predominantly referring to exchanges among technical systems, and less to exchanges among persons and their organisations.

When also taking functional needs into account, specific domain perspectives like the perceived need for and benefits expected from the exchange of meaningful data, and organisational, legal, political requirements come into play.

Moving towards a health system specific definition, one may stipulate that in the

> *(e)Health area interoperability refers to facilitating and safeguarding the exchange, understanding and acting on patient and other health information and knowledge among linguistically and culturally disparate medical professionals, patients and other actors within and across health systems in a collaborative manner.*

To better grasp the need for IOp in health systems, i.e. the seamless exchange and sharing of patient and other health data, one may want to conceptualize the health(care) system as a *value system* of a wide variety of cooperating health service providers, each of which has to manage its own health service(s) value chain. Paraphrasing the words of the 'guru' of strategic thinking and competitive advantage, Michael Porter, "Gaining and sustaining competitive advantage depends on understanding not only a firm's [healthcare provider – HCP] value chain but how the firm [HCP] fits in the overall value system... Competitive advantage is increasingly a function of how well a company [HCP] can manage this entire system. Linkages not only connect activities inside a company but also create interdependencies between a firm and its suppliers and channels" [18, p. 34].

At the core of such a conceptual view lies the model of a generic health service delivery system, which consists of interrelated value chains of individual health service providers. Together they ‘produce’ and sustain health by promoting good health and well-being, supporting disease prevention, undertaking diagnostic and therapeutic interventions, providing acute healthcare, rehabilitation and long-term care services. To enable, facilitate and support such health service delivery in a comprehensive, integrated fashion, supporting processes as well as facilitating tools and services are necessary, connected to and interconnecting these core service processes – be it via ICT-based, electronic means or otherwise. Only as a complex, dynamic ecosystem of interrelated healthcare processes all of these services effectively lead to optimal health.

3 Analysing Global Good Practice

In an attempt to obtain a better grasp of global initiatives and ventures to provide for district or national level eHealth IOp, a global scan of good practice examples of such eHealth platform implementation was undertaken [19-20], including exchanges with global experts and web searches. From a list of more than 50 candidate examples, the following eight were selected for a detailed description and analysis, covering the wide variety of such platforms observable:

1. The Australian Nation-wide eHealth Platform and its Personally Controlled eHealth Record (PCEHR)
2. The Belize Health Information System (BHIS)
3. SIGA Saúde - The City Health Information System of São Paulo, Brazil
4. The Canadian Nation-wide eHealth Platform and its interoperable Electronic Health Record (iEHR)
5. The pan-European eHealth framework and ICT infrastructure for Smart Open Services for European Patients – epSOS
6. Estonia’s national eHealth platform and nationwide EHR system
7. District telehealthcare services platform in Gilgit-Baltistan, Pakistan
8. South Africa Western Cape primary Health Care Information System

The comprehensive case descriptions [1] focused on policy & strategic issues, governance & regulation, and interoperability approaches and implementations. Each case study applied an accordingly structured information gathering instrument for descriptions of national or district efforts. These relied on detailed reviews of public eHealth/eGovernment or Information Society programmes, eHealth strategies, and other published results as well as contributions by local policy and technical experts involved in such endeavours.

4 Case Study Results: Domains and Levels of Interoperability

4.1 Interoperability Domains

The case study results triggered a renewed reflection on the question of ‘How should a healthcare system plan, implement, maintain, and further develop its IOp platform?’ Or, in other words, what are key issues, challenges, and facilitators when investing in such an undertaking.

This analysis led us to question the concept of so-called “levels” of IOp as explicitly named or implicitly assumed in many IOp definitions. A closer review indicated that the respectively named IOp requirement areas are not levels of achieving a certain goal in a consecutive fashion, but rather describe complementary fields which all need to be present to a certain degree in order to achieve even basic IOp. In line with this insight, we suggest to talk about *domains* of interoperability.

Policy Domain

All cases strongly underline the core relevance of clearly defined health policy goals to guide eHealth policy, strategy, and implementation, as well as the need for strong political leadership and support by all relevant stakeholders to indeed realise a national or district platform. Whether the driving promoter and instigator is a national ministry of health (as in Belize), a cooperation of federal and provincial/state governments as in Australia or Canada, a private foundation as in Gilgit-Baltistan, the city department of health as in Sao Paulo, or a district government as in South African Western Cape Province solely depends on the local circumstances and key driving forces responsible for providing healthcare to the general public. The case of Australia also underlines that such initiatives may progress, stop or even fail with a change in government and political/party priorities, as could also be observed some years ago e.g. in Hungary, The Netherlands, the United Kingdom.

In democratic countries, *societal* needs are reflected in policy goals as well as political measures to implement such policies, and eHealth platform *financing* is usually, as all cases indicate, a predominantly government or public health system budget issue (In the resource-poor context of Gilgit-Baltistan it is a private foundation), i.e. also allocated in the policy or public domain.

Governance and Legal Domain

Similarly, governance and legal/regulatory issues represent core success factors for eHealth platform establishment [21, 22]. Key reasons to become active in this domain are to improve and sustain data protection and system security to respond to increasing challenges in this field, thereby strengthening trust and confidence of all stakeholders, particularly patients and health professionals.

Other important aspects are to clarify the links and interrelations between eHealth regulation and the regulation of the healthcare system, to establish a reliable framework for eHealth application services, and to support the development of the eHealth market by encouraging effective competition between eHealth suppliers, and increasing certainty and market stability for them.

A well attended eHealth governance and legal domain assures that countries can expand sustainable eHealth successfully and economically for the benefit of their patients, citizens and the healthcare system. It also helps to ensure effective collaboration with other countries in case a cross-border exchange of patient data is envisaged.

Organisational Domain

Whereas policy and governance issues are often dealt with at a relatively high, generic framework level, all case studies suggest that successful eHealth platform establishment requires dedicated organisational support structures and processes to not only guide and direct eHealth infrastructure investments and controlling, but to also run daily administration and production itself or to commission appropriate private enterprises for selected tasks.

Here again the cases demonstrate the great diversity of organisational approaches and institutions which may be applied in a concrete setting, thereby reflecting the specific context and political-institutional situation. In Canada it is a large organisation, Canada Health Infoway, a public not-for-profit corporation funded by now with more than CAD 2 billion from both the federal and provincial governments; in Australia a mixed set-up, with the Department of Health and Ageing (DoHA) recently assuming the role of systems operator for a two year period and the National E-Health Transition Authority (NEHTA) continuing to guide ongoing development and standards compliance work; in Belize and Western Cape a dedicated department of the Ministry of Health. In the district of Gilgit-Baltistan, Pakistan, it is the eHealth Resource Centre at the Aga Khan University (AKU) in Karachi, in collaboration with the Aga Khan Health Service, Pakistan – all based on private foundation resources, the Aga Khan Foundation.

Data Access and Exchange Domain

In technical and engineering contexts, it is predominantly this domain which is focused on and analysed when discussing IOp issues. This does not only concern the technically correct, univocal transfer of data so that the person (or machine) accessing or receiving the data can indeed see and use them ‘as is’, but also the application of conventions or standards such that the receiver fully understands the structure in which the data, information or knowledge are presented, e.g. header information/metadata, and lastly the conceptual level, i.e. to assure that the data transferred are understood correctly with respect to their ‘meaning’. Accordingly, it is suggested to differentiate in this domain three levels of IOp, which indeed build on each other, but which may be present to different degrees in a given set of data:

- technical interoperability
- structural interoperability
- semantic interoperability

This data access and exchange domain and its three levels are of particular importance in an eHealth environment. In all case studies, issues related to it were explored to the limited extent this was possible. Based on the cases as well as further information from the literature and knowledge of the team, issues and challenges related to the data access and exchange domain are further explored.

4.2 Levels of Interoperability and Data Integration

The three levels of the data IOp domain may be described as follows in the eHealth field:

a) Technical (data) interoperability

As an initial, and in many health system contexts already highly relevant step, basic data IOp is achieved via an exchange of, e.g., electronic 'paper' documents in PDF format or similar, i.e. the document is accessed or transferred as is and cannot be integrated with similar data from another document, except for meta data accompanying the document and allowing for filing it, e.g., by patient name, healthcare provider issuing the document, date of the document, and/or purpose of sending the document – if such information is available as metadata. It seems that, e.g., referral or discharge letters are mostly transferred as they are.

In several of our good practice examples, this is a widely applied procedure.

At the next level of complexity – and also widely applied across national and district platforms -, the exchange of patient data is realised in the form of structural intreroperability.

b) Structural (data) interoperability

Here documents are meant which are structured according to standardised headings, like provided for by the HealthLevel 7 (HL7) Common Document Architecture (CDA) standards, e.g. for prescriptions, continuity of care documents, etc. This allows for regrouping and assembling patient data according to such headings on the contents of the sub-fields, like medications prescribed, symptoms, allergies, and the like.

Such documents may be translated into an XML coded format before exchange, requiring agreement, e.g., on a standard for the contents of the respective field names.

c) Semantic (data) interoperability

A further integration level is achieved when the contents of some fields is fully coded (like diagnoses coded in WHO ICD 10 with three digits; some countries maintain national versions with up to 5 digits) This is necessary in case (patient) data are to be integrated, aggregated and analysed, e.g. for research or surveillance purposes.

Again, several of the cases described encompass at least to some extent coded information, e.g. when gathering patient encounter data for public health surveillance and quality control.

Clinical documentation usually contains narrative text which is difficult, if at all, to standardise, like narrative reports of observations, impressions, and actions related to patient care. To achieve some degree of semantic IOp and to communicate in unambiguous language with collaborating healthcare providers, there exists a large number of "lists" to describe biomedical terminologies, like [13]:

- Controlled Vocabularies – lists of specified items to be used for some purpose, usually in an information system to reduce ambiguity, misspellings, etc.
- System of identifiers (IDs) or "codes" – Controlled vocabularies, and many lexicons, ontologies and thesauri are usually accompanied by systems of

identifiers for their units. Typically, these identifiers act as the primary unambiguous means of referring to the entities in the system for computational purposes, with the text form being used for communication with users. Examples are the "Concept Unique Identifiers (CUIs) of the Unified Medical Language System (UMLS), SNOMED CT IDs, etc.

- Lexicons – Lists of linguistic units that may be attached to a controlled vocabulary or ontology, in a specific language or sublanguage, often including linguistic information such as synonyms, preferred terms, parts of speech, inflections and other grammatical material.
- Ontologies (sensu information system) – Symbolic logical models of some part of the meanings of the notions used in a field, i.e. those things which are universally true or true by definition. Typically ontologies are implemented in logic languages such as Ontylog or OWL, or frame systems such as Protégé-Frames.
- Classifications – Organisations of entities into classes for a specific purpose such as international reporting or remuneration. Examples are the WHO ICD or Diagnosis Related Groups (DRGs).
- Thesauri – Systems of terms organised for navigation with the primary relationship being "broader".

Such biomedical terminologies like the Standardised Nomenclature of Medicine – Clinical Terms (SNOMED CT) try to overcome the mentioned semantic limitations, but making use of SNOMED CT may be quite expensive and/or may require a national licensing agreement with the International Healthcare Terminology Standards Development Organisation (IHTSDO).

As all cases show, achieving full semantic IOp is a long way to go, and may never be achieved fully. To assure some success, one probably should not even attempt to solve all (semantic) IOp challenges, but rather follow a step-wise approach, like making use of term binding to selected archetypes and record structure items.

This is the approach pursued by epSOS services, one of our cases. It represents a nice example and attempt to create and agree upon a small set of documents or rather data exchange models where the full contents of most or even all fields is fully coded such that they can even be safely translated into other languages.

Given a certain level of sophistication of participating platforms and the electronic recordkeeping applications used by healthcare providers, as well as mapping and transcoding middleware services, e.g. from hospital information or electronic medical systems to a central patient summary, and from there to foreign systems allow for assuring some degree of semantic IOp.

In addition, if access to patient data and other health information is not directly to the storage medium and data files, but rather via an exchange of messages, then the application of exchange or messaging standards/specifications like those provided for by IHE (Integrating the Healthcare Enterprise [23]) profiles like XDS/XDS.b, a profile or model for cross-enterprise document sharing (others are XCPD, XDA, XDF, XDR, ...) becomes necessary. Some IHE profiles provide not only for integration, but also for contents/some structure specifications.

5 Discussion: Who Needs Why How Much Interoperability?

One fundamental caveat must be noted, however. Such a conceptual approach towards the need for interoperability relies on the assumption that all actors and providers in the healthcare system have a strong incentive to improve not only their own services (for which they may already invest in internal IOp across their diverse digital systems), but to optimise service delivery and thereby promote better health for all via seamless integration and cooperation across the respective healthcare system. In a centralised national healthcare system (NHS) like the one designed by Lord Beveridge and introduced in the UK for its four home countries in the middle of the last century, this can be ‘ordered’ from the top, implemented by overall design, and centrally controlled. With respect to our case studies, such a situation prevails more or less in Belize, Sao Paulo, Estonia or Western Cape.

On the other hand, in more Bismarck-type (insurance-based), privately organised and mixed systems the incentive situation is different: cooperation and the resulting need for IOp will only arise within a larger organisation, like a hospital chain; within a network of cooperating physicians (perhaps together with hospitals, pharmacies etc.) where this may lead to a competitive advantage vis-à-vis other providers; or when policy actors or insurance companies provide strong incentives, e.g. via specific reimbursement fees, to promote integrated care.

Another route is regulation, e.g. to assure timely reporting of data needed for public health surveillance, or for health system resource allocation and overall management.

The given health system context has surely contributed to the specific challenges and problems encountered by countries like Australia and Canada (and various other larger countries globally). The overall situation in Gilgit-Baltistan - the limited reach of the established telehealthcare system confined to healthcare services provided by Aga Khan Foundation supported institutions – is complicated by its setting in a mixed type of healthcare system.

In a cross-border situation like the one faced by epSOS services, the overall context is even much more complex, and incentives have to come from the political level.

Taking such considerations a step forward, IOp optimisation involves assessing the benefits and costs at a given integration level (e.g. within a single hospital, across several hospitals of a hospital chain, a GP network, a district system, nationally, ...) and the respective need and benefits for healthcare professionals, patients and/or the health system from sharing and accessing in a coherent manner patient administrative, workflow or clinical EHR-like patient data and other records.

Sharing of data in exchangeable, interoperable formats may not at all be in the interest of medical professionals. E.g., for general practitioners, it may mean that they (or their staff) have to do data entry in a coded format – rather than in a narrative fashion - for the benefit of others, like hospitals and specialists to whom they refer patients, or of public health organisations to which they report disease data. There is no immediate benefit for the local doctor to spend extra time, costs and effort to support such IOp. Only if the medical community as a whole embraces the idea and combines it with collaborative working practices, and agrees collectively to accept the, at least initially, additional work, may IOp also become a success factor for the individual actor.

Furthermore, in a fee-based reimbursement system, the re-examination of a patient, another X-Ray, another lab analysis may be more profitable for a physician than relying on earlier results and thereby earning less income. The same may hold for hospitals or other providers, depending on the type of healthcare financing and reimbursement system(s).

In a single (government) provider system like in various national or district health services (cf. Estonia, Belize, Sao Paulo) interoperability may, in theory, not become an issue as long as a single, fully integrated ICT system is implemented which is scalable and provides for all functionalities needed. However, in reality a single system will almost never be sufficient due to various specialised systems like picture archiving and communication systems (PACS), laboratory and pharmacy systems etc. which need to be connected. Usually also the presence of legacy systems, local medical culture and language (like abbreviations in use), and other factors present significant IOp challenges.

Another relevant factor which emerged in large countries (like in Australia or Canada) is size. Seemingly national competent authorities as well as ICT companies and their experts are unable to handle the full complexity of such platforms once the number of persons with records (and the number of healthcare providers involved) surpasses a certain number which may be anywhere between 5 and 10 million patients.

6 Conclusions: The Role of eHealth Interoperability for Better Healthcare

The generic definition given above concerning what interoperability is all about, namely facilitating and safeguarding the exchange, understanding and acting on patient and other health information, hides the issue of who needs which information in what form for what purposes. Assuming IOp is regarded by those politically responsible for the health system as useful, and therefore, e.g., mandated from the top or regulated by law, or seen by healthcare actors as so beneficial that they are prepared to invest in it, then two issues arise:

a) Among which actors should IOp become established?

b) What level of IOp – understood as the degree to which data can be exchanged between eHealth systems and integrated into another one – should be aspired?

Let us first consider a). The wider the integration domain – with respect to the number of people/patients living in the area -, the larger the number of providers and their locations as well as the size of the region covered, the higher will be the IOp efforts involved and the concomitant costs, particularly in a Bismarck-type (insurance based), a private, or a mixed health system, and the lower the individual benefits, because in a competitive system there is usually no inherent incentive for individual stakeholders to share patient data with other providers.

If healthcare is a purely local activity without synergistic benefits from cooperation in a wider geographic area, then IOp will not support better healthcare, and the extra costs are wasted. On the other hand, in a fully centralised system with only a few or perhaps only one central ICT system, illustrated by some of our cases, full IOp will increase the quality and efficiency of healthcare delivery possibly in a linear fashion (up to a certain degree) the wider the enactment area becomes.

With respect to b), the lessons to be learned here from our cases are two-fold: At the generic level, there is no absolute level of IOp, and the concept of "interoperability" as such may even be misleading. As our cases illustrate, in a small to very small country, it may be more efficient to just go for one fully integrated ICT system without any IOp challenges.

Secondly, the given framework conditions of the national and district/provincial healthcare systems (like in Australia, Brazil, Canada) will greatly impact what level of IOp for which levels of cooperation may be optimal. Again, there is no clear-cut answer; rather, it all depends on the given situation, including the resource situation, the presence of legacy systems or not, etc.

Acknowledgements. This research was commissioned and financially supported by the European Space Agency (ESA) through its Satellite-Enhanced Telemedicine and eHealth for Sub-Saharan Africa (eHSA) Programme, specifically the horizontal study on Interoperability. Funding is provided in the context of the Delegation Agreement between the Luxembourg Agency for Development Cooperation (Lux-Dev) and the European Space Agency (ESA) for the implementation of the first phase of the eHSA Programme, sponsored jointly by the EU-Africa Infrastructure Trust Fund (ITF) and the Government of Luxembourg.

Critical review and input by study team members to technical notes prepared for the study are gratefully acknowledged.

References

1. Study on Interoperable eSystems for Africa Enhanced by Satellites, http://www.isaes.org
2. ESA - European Space Research and Technology Centre (ESTEC). Statement of Work: Satellite-Enhanced Telemedicine and eHealth for sub-Saharan Africa (eHSA) Programme - Study on Interoperability, Noordwijk, The Netherlands (2012)
3. Satellite-enhanced eHealth for sub-Saharan Africa Programme (eHSA), http://iap.esa.int/ehsa
4. C4ISR Architectures Working Group: Levels of Information Systems Interoperability. U.S. DoD, OSD (C3I), CIO, Director for Architecture and Interoperability, Washington, DC (1998)
5. Miller, P.: Interoperability. What is it and why should I want it? Ariadne Online (24) (June 3, 2000), http://www.ariadne.ac.uk/issue24/interoperability/
6. Morris, E., et al.: System of Systems Interoperability (SOSI): Final Report. Carnegie Mellon Software Engineering Institute for the SEI Joint Program Office (2004)
7. European Commission. European Interoperability Framework for Pan-European eGovernment Services. Official Publications of the European Communities, Luxembourg (2004)

8. ETSI: SOS Interop. - Standards, Open Standards and Interoperability (A series of Workshops for Industry, Standards Bodies and Policy Makers) (2005, 2006), http://www.etsi.org/standards/how-does-etsi-make-standards/18-services/457-sos-interoperability
9. Ibid.: Hill, J.L.: A Conceptual Model of ICT Interoperability (2005)
10. Novakouski, M., Lewis, G.A.: Interoperability in the e-Government Context. Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA (January 2012)
11. i2-Health - Interoperability Initiative for a European eHealth Area: European eHealth Interoperability Roadmap, Final Public Project Report, Bonn/Brussels (December 2007), http://www.i2-health.eu/
12. European Commission. Connected Health: Quality and Safety for European Citizens. Report of the Unit ICT for Health in collaboration with the i2010 sub-group on eHealth and the eHealth stakeholders' group. Towards A European Roadmap For Achieving eHealth Interoperability, vol. 6. European Commission, Brussels (2006)
13. Stroetmann, V.N., et al.: Semantic Interoperability for Better Health and Safer Healthcare. In: SemanticHEALTH. European Commission, Luxembourg (2009)
14. Mead, C.N.: Committee on the Review of the) Adoption and Implementation of Health IT Standards by the DHHS. Office of the National Coordinator for Health Information Technology (September 17, 2007), http://www.iom.edu/activities/workforce/healthitstandards/2007-sep-16/agenda-meeting-1-review-of-the-adoption-and-implementation-of-health-it-standards-by-the-dhhs-office-of-the-national-coordinator-for-health-information-technology.aspx
15. Stroetmann, K.A., et al.: Towards an Interoperability Framework for a European e-Health Research Area – Locating the Semantic Interoperability Domain (2005), http://www.who.int/entity/classifications/terminology/en/stroetmann.pdf
16. CALLIOPE (CALL for InterOPErability): EU eHealth Interoperability Roadmap. European Commission, Brussels (December 2010), http://www.calliope-network.eu
17. ISO. Health informatics — Capacity-based eHealth architecture roadmap — Part 2: Architectural components and maturity model. ISO/TC 215/SC N 721, draft (2013)
18. Porter, M.: Competitive Advantage. The Free Press, New York (1985)
19. Stroetmann, K., et al.: European countries on their journey towards national eHealth infrastructures. Office for Official Publications of the European Communities (2011)
20. Canada Health Infoway: Category Archives: Global Perspectives, http://infowayconnects.infoway-inforoute.ca/blog/global-perspectives/
21. Greenfield Management Solutions: Satellite-Enhanced Telemedicine and eHealth for Sub-Saharan Africa (eHSA) Programme, Study on Regulatory Aspects: Summary Report, Kimberly (2013)
22. Logica: eHealth for Sub-Saharan Africa (eHSA) – Governance study (2013), http://iap.esa.int/projects/health/ehsa-governance
23. Integrating the healthcare Enterprise, http://en.wikipedia.org/wiki/Integrating_the_Healthcare_Enterprise; http://www.ihe.net/

Predictive Models for Hospital Bed Management Using Data Mining Techniques

Sérgio Oliveira[1], Filipe Portela[1], Manuel F. Santos[1], José Machado[2], and António Abelha[2]

[1] Algoritmi Centre, University of Minho, Guimarães, Portugal
sergiomdcoliveira@gmail.com, {cfp,mfs}@dsi.uminho.pt
[2] CCTC, University of Minho, Braga, Portugal
{jmac,abelha}@di.uminho.pt

Abstract. It is clear that the failures found in hospital management are usually related to the lack of information and insufficient resources management. The use of Data Mining (DM) can contribute to overcome these limitations in order to identify relevant data on patient's management and providing important information for managers to support their decisions.

Throughout this study, were induced DM models capable to make predictions in a real environment using real data. For this, was adopted the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology. Three distinct techniques were considered: Decision Trees (DT), Naïve Bayes (NB) and Support Vector Machine (SVM) to perform classification tasks. This work explored the possibility to predict the number of patient discharges using only the number of discharges veirifed in the past. The models developed are able to predict the number of patient discharges per week with acuity values ranging from ≈82.69% to ≈94.23%. The use of these models can improve the efficiency of the administration of hospital beds. An accurate forecasting of discharges allows a better estimate of the beds available for the coming weeks.

Keywords: Hospital Management, Management of Patients, Management of Beds and Data Mining.

1 Introduction

Organizations collect data from multiple business processes. Based on the idea that large volumes of data can be a source of implicit knowledge, new knowledge can be extracted when using appropriate tools (e.g., data mining) [1]. Hospitals operate within this principle, since their databases contain hidden knowledge that is important, for example, in a patient's clinical status and prognosis [2].

Data Mining (DM) implementation in the health sector can be an asset. In fact, DM can help health insurers, detect fraud or abuse in decision making over the relationship with their clients, doctors can identify more effective treatments, as well as best practices. As consequence patients can receive better health services. DM provides techniques to transform huge data volumes into useful knowledge for

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_39, © Springer International Publishing Switzerland 2014

decision making [3]. Based in this principle it was determined whether it was possible to construct models of DM capable of supporting the process of hospital management. The support is at the level of beds management using the number of patient discharges. The management of beds is currently a real problem for the hospitals due to the increase of the number of patients admitted in some specialities and limited resources. A system capable of predicting the number of discharges of patients can significantly improve hospital quality, optimizing the distribution of beds.

The main goal of this work is to predict the number of patient discharge by week using data mining techniques. To this problem were used classification techniques and a set of configuration applied to four services. The results were satisfactory with the accuracy being around 82.69 % and 94.23%. These results have enough quality to help the decision process and corroborate the initial idea: the extent to which historical data on hospital discharges can be used to forecast futures discharges.

In the second chapter is presented the conceptual framework, DM techniques used, and metrics for evaluating the classification models. The third chapter presents the development of the practical component, guided by the CRISP-DM methodology. In the fourth chapter, the most relevant aspects of this work are analysed and discussed. The fifth chapter focuses on the work conclusions and, finally, possible directions for posterior work.

2 Background

The World Health Organization describes the hospital as being "... part of a medical and social organization whose function is to provide a comprehensive health service and care for the population, both curative and preventive, and whose outpatient services must reach families in the home environment, the hospital is also a centre for training of health workers and biosocial research. " [4].

To manage a hospital a deep understanding about the institution is mandatory. Managers need to have the perfect notion about the rules and routines of services. They have to be able of identifying the strengths and aspects that need improvement. From these aspects managers should outline a clear and organized plan to provide for an efficient and effective hospital management [5].Management consists of a variable set of technical tools and technologies used to ensure organizations' success [6].

In an organizational structure such as a hospital, where the core business is to optimize patients admission in the hospital, minimizing the length of stay and maximizing the quality of treatment should be sought [7]. In this sense, one of the most important features is the distribution of beds among the various services. The way of how the hospital is organized reflects the efficiency and quality of hospital management[8].

Introducing new technologies in hospitals has provided the opportunity for work improvement, e.g. some tasks can be performed faster, more consistently and with lower costs [9]. DM as well as some classic statistic methods have been used in hospital databases since 1990 [10], [11], [12]. Ever since, DM is becoming increasingly popular and essential in healthcare [3].

2.1 Hospital Bed Management

Hospital beds are one of the scarcest resources of hospitals. In most cases the beds are arranged according to hospital specialties in order to provide a better service for patients. Beds management reflects not only the services efficiency but also the quality of hospital management. In order to improve the hospital management focusing on bed management, there are some studies using DM giving special attention on patient admittion in the hospital and considering as main goals: i) to decrease the length of stay of patients; and ii) performing a good management of beds. A project conducted at Chiba University Hospital studied which medical care tasks were directly correlated with patients' length of stay. This study revealed a strong correlation among some variables, presenting coefficients between 0.837 and 0.867, in a range of [-1, 1]. The study concluded that the results obtained have shown a strong correlation between patient discharges and hospital management quality [11].

Another study was conducted at the National University Hospital of Singapore. This study had as main objective: to identify which was the key variable associated to mismatch allocation of beds by department. The models obtained presnted acuity values of 74.1% and 76.5% and concluded that the key variable was the medical speciality. Through this study was possible to determine strategies to the allocation of beds in the hospital [13].

2.2 Data Mining

Technological advances have provided new ways to create and store data. Organizations accumulate data related to their processes (billing, business transactions and accounting) based on the idea that the large volumes of data can be a source of knowledge [1]. From a technical standpoint, DM is a process that uses artificial intelligence techniques, statistics and mathematics to extract useful information and knowledge (or patterns) from large volumes of data. These patterns can be in the form of business rules, affinities, correlations, or terms of forecasting models [14]. The prediction methods goal is to automatically build a behavioural model, obtaining new samples and unknown samples, and being able to predict values of one or more variables related to the sample. The design patterns that enable knowledge discovery are easy to understand and can be easily used as a work base [15].

Classification models aim to identify a function that associates an event to a class within several discrete classification classes, i.e., classifiers map the input space into predefined classes. For example, the class patient has attributes that describe the patient; if a particular person meets the classification properties of the patient, then that person can be classified as patient [1].

For this work the implementation of DM was achieved through the statistical environment R. R presents itself as a programming language and an environment for statistical development [16]. The library e1071 [17] was used to implement the DM techniques Support Vector Machine (SVM), Decision Tree (DT), and Naïve Bayes (NB), and for troubleshoot forecasting - classification. To conduct the DM models evaluation the rminer library has been used [18].

3 Knowledge Discovering Process

The DM process is complex but when conducted in a methodological context becomes easier to understand, implement and develop. CRISP-DM methodology was followed to carry out the present study.

3.1 Business Understanding

Data used in this study were retrieved from Centro Hospitalar do Porto (CHP), Porto, Portugal. Models use as input values the number of patients discharged distributed across multiple services. The supplied sample comprised 62 302 records. CHP has no mechanism to predict the discharges flow. This evidence triggers an entire workflow process in accordance to the objective of this study, to predict the number of patients weekly discharged, primarily targeting the bed management improvement.

3.2 Data Understanding

The provided sample comprises the period between 01.01.2009 and 31.12.2012, referring to 1461 days. The years 2009, 2010 and 2011 have 365 days and the year 2012 has 366 days, which is the only bissextile year from the sample provided. From the referenced timeline were collected 62,302 records, corresponding to discharges from ninety one hospital services. Each record consists of three fields:

- Date: corresponds to the date (day, month and year) of patient's discharged;
- Service: is the service associated to the patient in the date;
- Discharge number: contains the number of patients who were discharged. This field is directly related to the date and the hospital services, i.e., records were grouped by date and service.

As already mentioned, one of the goals is to predict the hospital discharges per week. From the ninety one existing services only four presented records for all week days during the period analysed. Based on this evidence, data was only exploited from four services: Orthopaedics, Obstetrics, Childbirth and Nursery.

The Table 1 presents a characterization of the data associated to each one of the services.

Table 1. Data characterization

	Orthopaedics	Obstetrics	Childbirth	Nursery
Maximum	78	91	8	92
Minimum	19	40	0	34
Average	≈48.6	≈62.9	≈2.4	≈57.3
Coefficient of Variation	≈22.634%	≈17.011%	≈70.833%	≈19.197%
Standard Deviation	11	10.7	1.7	11
Total Discharges	10108	13080	495	11924

3.3 Data Preparation

In this study, as already mentioned, the objective it to make predictions of weekly discharges. So it was necessary to group the daily records into weekly records. By convention, one week begins on Sunday and ends in the next Saturday. Following this principle the discharges were grouped into 52 weeks and the respective years 2009, 2010, 2011 and 2012. Once these records were grouped in weeks it was necessary to use methods capable of determining intervals of values for hospital discharges, in order to make predictions using the classification approach.

The selection of the number of classes or intervals does not constitute a rigorous and scientific method, nor is there any selection method that can be considered appropriated [19]. Therefore, some methods were implemented to create classes: average, quartiles, average and standard deviation and sturges rule. Through the set of methods described, the data tables were created according to the number of weeks and the respective years, i.e., the rows are the number of weeks (52 in total), the columns (4 in total) correspond to the years, 2009 to 2012. This data representation is defined as being conventional. Having in consideration this data it was also explored a different approach using a sliding window data representation. However, this paper was focused in presenting the first approach – conventional.

3.4 Modelling

The techniques considered to induce classification models were: Support Vector Machine (SVM), Decision Trees (DT) and Naïve Bayes (NB). The selection of Data mining techniques was based in three distinct aspects: interpretability of the models, engine efficiency and the suitability for small datasets. SVM achieved the second and third goals. DTs and NBs met the first two goals. In the case of NBs, they also can present good results using a reduced number of data.

In order to implement a mechanism for model testing, two sampling methods have been selected: 10-folds Cross Validation (10-folds CV) and Leave-One-Out Cross Validation (LOOCV). The 10-folds CV was adopted due to the good results that it has demonstrated on multidisciplinary data [20]. LOOCV it was used because it is more suitable for databases with only a few dozens of tuplets [21] - since the data tables only have 52 or 205 records. All the techniques were submitted to the tune function. This function comes with the e1071 package. As main objective it executes network searches of hyper parameters intervals previously provided and consecutively identifies the best model and the respective hyper parameters.

For the SVM technique two distinct kernels were used, Radial-Basic Function (RBF) and Linear. Given the different kernels used it was necessary to perform different parameterizations because their hyper parameters differ from kernel to kernel. Depending on the kernel used by SVMs, a range of values for parameter C was defined. Its range has been defined by the values obtained by the power $2^{(1,..,4)} = [2, \ldots, 16]$, in which $C > 0$. The cost parameter C introduces some flexibility separating the categories in order to control the trade-off between errors in training or stiffness margins [22].

The hyper parameter Gamma (γ) was defined in the same way as C. The range was determined according to the values obtained by the power $2^{(-1,0,1)} = [0.5, 1, 2]$.

Its parameterization was used in the RBF kernel. The γ value determines the curvature of the boundary decision [23].

The DTs were used to perform predictions through the classification approaches. The implementation of this technique was achieved through the CART algorithm. Two methods for attribute selection or splitting rules were used, Information Gain (IG) and the Gini Index (GI). The attribute selection measure IG determines the attribute with the highest information gain and uses it to make the division of a node [24]. The GI is determined by the difference between the original information requirement (i.e., based on only the ratio of classes) and the new requirement (i.e., obtained after partitioning A). This difference is expressed as: $Gain(A) = Info(D) - Info_A(D)$. The attribute A that has the highest information gain, Gain (A) is the division attribute of node n [24]. The objective of GI is to calculate the value for each attribute using the attribute for the node with the lowest impurity index [1]. The GI index measures the impurity of D, using a data partition or a training set of attributes $Gini(D) = 1 - \sum_{i=1}^{m} pi^2$, where pi corresponds to the probability of an attribute D of a class Ci. This value is estimated by $|Ci \cap D|/|D|$. The sum is calculated as a function of m classes [24].

Finally, the algorithm NB was used to perform predictions using the classification approach (was not carried out any manual parameterization for this algorithm). This algorithm used the tune function to identify the sampling method that should be used. These methods were previously determined.

The developed models can be represented by the following expression:

$$M_n = A_f + S_i + C_x + MRD_z + TDM_y + MA_k.$$

The model M_n belongs to the approach (A) classification and is composed by a service (S), a type of class (C) a method of data representation (MDR), a DM technique (TDM) and a sampling method (SM):

$$A_f = \{Classification_1\}$$
$$S_i = \{Orthopedics_1,\ Obstetrics_2, Parturition_3, Nursery_4\}$$
$$C_x = \{Average_1,\ Quartils_2, Average\ and\ Standard\ Deviation_3,\ Sturges_4\}$$
$$MDR_z = \{conventional_1, Sliding\ Window_2\}$$
$$TDM_y = \{SVML_1, SVMRBF_2, DTGI_3, DTIG_5, NB_5\}$$
$$SM_k = \{10 - folds\ CV_1, LOOCV_2\}$$

For instance, the model (M_1) uses the technique SVM with RBF kernel and the sampling method 10-folds CV and it is expressed by: $M_1 = A_2 + S_2 + C_5 + MDR_1 + TDM_2 + SM_1$.

3.5 Evaluation

This task is dedicated to models' evaluation. To evaluate the results achieved by the DM models it was used the accuracy or precision measure.

As the results presented by the DM models depend on the division of the mutually exclusive subsets, two procedures have been implemented: 10-folds CV and LOOCV. For the splitting procedures, ten executions were performed for each of them.

Around 100 experiments were performed for each test configuration with models that use the 10-folds CV procedure. 520 experiments were performed to test models using LOOCV procedure. Table 2 depicts the obtained values of accuracy for the best classification models.

Table 2. Evaluation of classification models

Model	Service	Technique	Sampling Method	Class	Accuracy
M_1	S_1	TDM_2	$SM_{1,2}$	C_4	≈82.69%
M_2	S_2	TDM_2	$SM_{1,2}$	C_4	≈82.69%
M_3	S_3	TDM_2	$SM_{1,2}$	C_4	≈94.23%
M_4	S_4	TDM_2	$SM_{1,2}$	C_4	≈90.38%

The DM technique that presented the best results was the SVM RBF, the two sampling methods used did not show to be crucial for the results, being equivalent. From table 1 it can be seen that rule sturges used in the creation of classes was the best method. Then, figure 1 presents the classes predicted by the models and their respective frequency for 2012 by week. Classes are represented by their maximum (green point) and at its minimum (blue point) value. If the expected values of the year 2012 (red dots) are between the maximum and minimum means that the class predicted was properly carried out, i.e., the actual values were classified in the predicted class limits. The presented classes were obtained using the sturges rule.

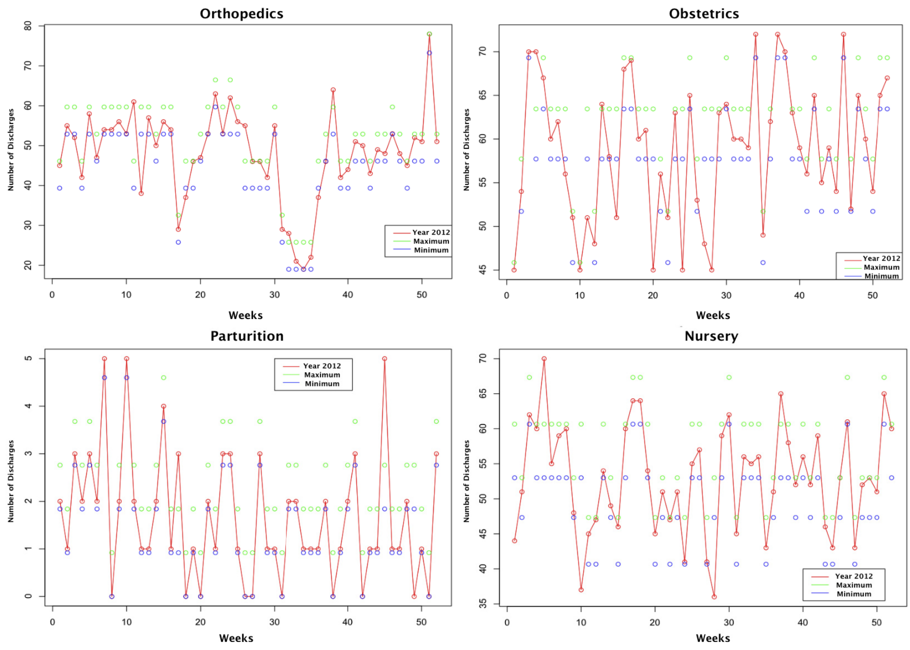


Fig. 1. Classes predicted

4 Discussion

The presented results are quite acceptable due to the prediction evaluations made. For the classification models, the best predictions resulted in acuities greater than ≈82%. Orthopaedics and obstetrics are the services with the worst results claiming ≈82.69% of acuity. The results obtained for the services of childbirth and nursery are ≈94.23% and ≈90.38% respectively. The predictions made for the childbirth and nursery service present results sufficiently satisfactory to support decision making. Table 3 presents the number of times DM models hit and missed.

Table 3. Results of Predictions

Model	Wrong	Correct
M_1	9	43
M_2	9	43
M_3	3	49
M_4	5	47

From the two sampling methods used 10-folds Cross Validation (10-folds CV) and Leave-One-Out Cross Validation (LOOCV), 10-folds CV is the most suitable for this study, not because the results since the differences are very small and often are the same, but because the execution time of the models. Models that use the sampling method LOOCV take a long time to process.

Since the used data is real, the inclusion of these models in a Decision Support System (DSS) becomes expectable. Thus the knowledge generated from the usage of DM can be useful so that it can influence the operational efficiency, facilitating high-level decisions and service providing.

5 Conclusion

This work proved that is possible induce classification data mining models to predict weekly discharge of patients in a hospital using real data of daily discharge. A study was carried out considering data from CHP, corresponding to four years of activity. Good results were obtained in terms of precision for four services: ≈82.69% for orthopaedics and obstetrics; ≈94.23% and ≈90.38% for childbirth and nursery, respectively.

Two different methods of data representation were explored: conventional and sliding window. The second method is computationally more expensive and does not improve the results. Conventional representation in association with sturges rule, the sampling method 10-folds CV and the SVM technique, demonstrate to be the most suitable for this type of data.

Finally, we can conclude that by using classification techniques and past discharges data is possible to forecast future discharges.

6 Future Work

Future research will take into consideration the following aspects:

- To explore different types and configuration of Data Mining techniques;
- Incorporate new variables in the predictive models, such as gender and age of patients;
- Determine if the predictions could be more detailed. The introduction of the entry number of patients, bed occupation time and ratio should be taken into account to create a better bed management system;
- Repeat the experiments for more hospital services with new data.

Acknowledgements. This work has been supported by FCT – Fundação para a Ciência e Tecnologia in the scope of the project: PEst-OE/EEI/UI0319/2014.

The authors would like to thank FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EIA/72819/ 2006 (INTCare) and PTDC/EEI-SII/1302/2012 (INTCare II).

References

[1] Santos, M., Azevedo, C.: Data Mining Descoberta do conhecimento em base de dados. FCA - Editora de Informática, Lda (2005)
[2] Santos, M., Boa, M., Portela, F., Silva, Á., Rua, F.: Real-time prediction of organ failure and outcome in intensive medicine. In: 2010 5th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6 (2010)
[3] Koh, H., Tan, G.: Data mining applications in healthcare. J. Healthc Inf. Manag. 19(2), 64–72 (2005)
[4] WHO, Expert Committee on Health Statistics, 261 (1963)
[5] Santos, I., Arruda, J.: Análise do Perfil Profissional dos Gestores dos Hospitais Particulares da Cidade de Aracaju- SE. Revista Eletronica da Faculdade José Augusto Vieira Na-77 (2012)
[6] Proença, J., Vaz, A., Escoval, A., Candoso, F., Ferro, D., Carapeto, C., Costa, R., Roeslin, V.: O Hospital Português. Vida Económica-Conferforum (2000)
[7] Neves, M.: Os Médicos vão ter de ser os motores da reforma do sistema. Revista Portuguesa de Gestão & Saúde (5) (2011)
[8] Yang, G., Sun, L., Lin, X.: Six-stage Hospital Beds Arrangement Management System. presented at the Management and Service Science (2010)
[9] Dwivedi, A., Bali, R., Naguib, R.: Building New Healthcare Management Paradigms: A Case for Healthcare Knowledge Management. presented at the Healthcare Knowledge Management Issues, Advances, and Successes (2006)
[10] Bose, R.: Knowledge management-enabled health care management systems: capabilities, infrastructure, and decision-support. Expert Systems with Applications 24(1), 59–71 (2003)
[11] Tsumoto, S., Hirano, S.: Data mining in hospital information system for hospital management. presented at the ICME International Conference on Complex Medical Engineering, CME 2009, pp. 1–5 (2009)

[12] Tsumoto, S., Hirano, S.: Towards Data-Oriented Hospital Services: Data Mining-based Hospital Mangement. In: presented at the The 10th IEEE International Conference on Data Mining Workshops, Sydney, Australia (2011)
[13] Teow, K., Darzi, E., Foo, E., Jin, X., Sim, J.: Intelligent Analysis of Acute Bed Overflow in a Tertiary Hospital in Singapore. Springer US (2012)
[14] Turban, E., Sharda, R., Delen, D.: Decision Support and Business Intelligence Systems, 9a Edição. Prentice Hall (2011)
[15] Maimon, O., Rokach, L.: Introduction to Knowledge Discovery and Data Mining. In: Data Mining and Knowledge Discovery Handbook, 2a Edição. Springer (2010)
[16] Torgo, L.: Data Mining with R: Learning with Case Studies. CRC Press - Taylor & Francis Group (2011)
[17] Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F.: Misc Functions of the Department of Statistics (e1071) (2012)
[18] Cortez, P.: Simpler use of data mining methods (e.g. NN and SVM) in classification and regression (2013)
[19] Reis, E.: Estatística Descritiva, 7a ed. Edição Sílabo (2008)
[20] Witten, I., Frank, E., Hall, M.: Data Mining Pratical Machine Learning Tools and Techniques, 3a Edição. Morgan Kaufmann (2011)
[21] Refaeilzadeh, P., Tang, L., Liu, H.: Cross-Validation, Encyclopedia of Database Systems, vol. 5. Springer (2009)
[22] Kantardzic, M.: Data Mining Concepts, Models, Methods, and Algorithms, 2a Edição. Wiley - IEEE Press (2011)
[23] Ben-Hur, A., Weston, J.: A User's Guide to Support Vector Machines. In: Carugo, O., Eisenhaber, F. (eds.) Data Mining Techniques for the Life Sciences. Humana Press (2010)
[24] Han, J., Kamber, M., Pei, J.: Data Mining Concepts and Techniques, 3a Edição. Morgan Kaufmann (2012)

How to Use Activity Theory Contradiction Concept to Support Organization Control

Antonio Goncalves[1], Marielba Zacarias[2], and Pedro Sousa[3]

[1] Instituto Politécnico Setúbal, Campus do IPS, Estefanilha, 2914-504 Setúbal, Portugal
[2] Universidade Algarve, Faro, 8005-139 Faro Portugal, Portugal
[3] Instituto Superior Tecnico, Campus Alameda, Alameda, 1049-001 Lisboa, Portugal
antonio.goncalves@estsetubal.ips.pt, mzacarias@ualg.pt, pedro.sousa@dei.ist.utl.pt

Abstract. Despite the existing organization control model, it isn't clear how people choose the appropriate measures, norms and the control values and how to relate them with operation of organization. To solve this challenge we propose to use Activity Theory contradictions concept to help to choose and monitor useful measures, norms and viability values.

Activity Theory is a framework that helps describe human work in an organization. It focuses on the premise that to examine human work, one has to reflect on the individual actions and the dynamic aspects of cooperative works between people. We take advantage of overlap between the DEMO and Activity theory concepts to find relevant control elements from activities gathered from organization DEMO model. We will use the proposed solution via a case study of a service (e.g., True-Kare) that allows someone to provide a remote assistance to another person by using a mobile phone.

Keywords: Activity Theory, Ontological Model, Cooperative work, Dynamic Work, Context.

1 Introduction and Motivation

Activity Theory can be described as a way of how people work in an organization, such as an individual or group (***subject***) is assisted by artifacts (***tools)*** to attain an objective (*object*) and may accept regulations ***(rules)*** to work with other persons (***community)*** that contribute to the ***objective*** through a ***division of labor***. From such activity there is an outcome.

Contradiction is a relevant Activity Theory concept [1]. It can be regarded as historically accumulated structural tensions within (i.e., elements of an activity) and between activities that generate problems, failures, conflicts and break down, but at the same time become the ability of an activity to develop it self. Contradictions are continuous within an organization, and their awareness and resolution is a regular practice of organization development. For example if in an organization activity, Microsoft Windows is used as a tool and someone can only use Macintosh, the collective activity may not operate properly and may even break down. Detecting

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_40,

contradictions is a hard task, normally, contradictions are perceived based on the reflection and analysis of past work of the people in the organization, i.e. their actions and operations.

The purpose of DEMO organization control is to manage allowed states for certain norms of an organization, which guarantee its viability. Organization control model keeps a record of all norms of an organization that are measured and their respective viability values that restrict them to guarantee organization working in a proper way.

We think that there is a lack of methodology that addresses how to choose the norms and how to find an exception. Traditionally several researchers have addressed the issue exception recognized that organizations have continuously kept on solving them. Tradition solutions comprise to store information about exceptions and on how to solve them. This avoids the expenditure of an added effort in handling the continuous treatment of the same kind of exceptions [2].

Our proposal aims to deal with exception that caused the organization dysfunctions using and DEMO ontological model that deals with activities contradiction detection.

The document is structured as follows: Section 2 and 3 presents the theoretical foundations. Section 4 presents the proposed method. Section 5 presents a case study that is used to validate the solution. Lastly, section 6 approaches the problems encountered, conclusions and future work.

2 DEMO Organization Control Model

The DEMO methodology [3] provides an firm description of an organization through an ontological model, which emphasizes the description of the organization core business. DEMO is grounded on a stable Ψ theory [3]. According to the Ψ theory, through their collective interactions, people engage in social obligations relating to actions to be taken and agree on the results of these actions (i.e. facts).

Aveiro [2] developed a DEMO organizational control model that manage the aspects of organization changes, as a result of treatment of exceptions. To this end Aveiro proposed use of resilience strategy and microgenesis as a mechanism to deal with new kind of exceptions. In short, Aveiro proposes that organizations should explicitly design and deploy their organization control with mechanism of resilience and microgenesis dynamics.

Fig. 1 presents the control object fact diagram that is part of state diagram of DEMO control model of an organization proposed by Aveiro. It describes the conceptual model of allowable state space and can be used to capture the history of organization change. State model defines the main relations between objects and relevant facts that should be recorded in an organization such as: A dysfunction was observed (F1) and a dysfunction was fixed (F2). F3 is a fact that record exception discoverer that causes the dysfunction. The solution of a dysfunction consists in discovering the right strategy (F4) that should be used to resolve the exception and finally applying the strategy (F5). When it is not possible to identify the kind of exception, it is considered that we are dealing with an unexpected exception (i.e., a novel one) and an organization engineering process has to be started (F6) as part of microgenesis dynamics.

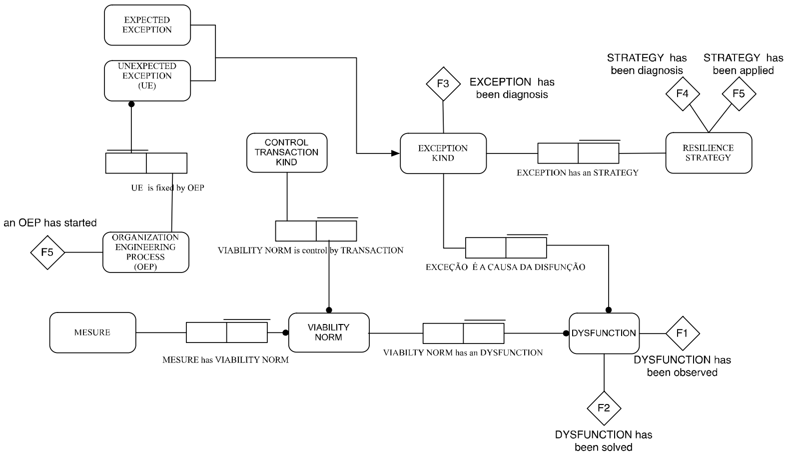


Fig. 1. DEMO Organization Object Fact Diagram

3 Activity Theory and Contradictions

According to Activity Theory [4] people's work in an organization is a complex structure, which is performed by a group of people, consisting of individual actions, according to a division of labour, framed in a historical context. An activity can be seen as a transformation of objects into a result, material or immaterial, which satisfies a certain need or desire of user who executes. A need can only be satisfied when it results in obtaining a final object that is the reason for the activity. A scenario described by Leont'ev explains these relationships: A person is hungry (needs to eat) and can satisfy this need by seeking food (object). He is motivated to the activity of finding food when he feels the need to eat and when idealizes an object that can satisfy him.

Engeström, departed from the theoretical basis of Activity theory, proposes a triangular activity diagrams that includes various components [1]. An example of activity diagram is present in Fig. 2.

This diagrams describes the activity of a person to become a membership of a tennis club named Volley as described in Dietz book [3] pags (15-32). The Triangular Activity Diagram suggests the possibility of multiple relationships within the triangular structure activity; however, the main task is always to understand the entire context rather than their separate connections, since work cannot be understood or analysed outside the context in which it occurs.

According with activity theory contradictions should be provided as tensions or imbalances manifested by failures, problems or errors. We can detect the manifestation of contradiction by analysing the people work and speech in an organization [5], expressed in the actions and operations performed by a person under an activity. Contradictions can be analysed from an isolated element or between the elements that constitute the triangular activity diagram or from the relations between elements of an activity [5].

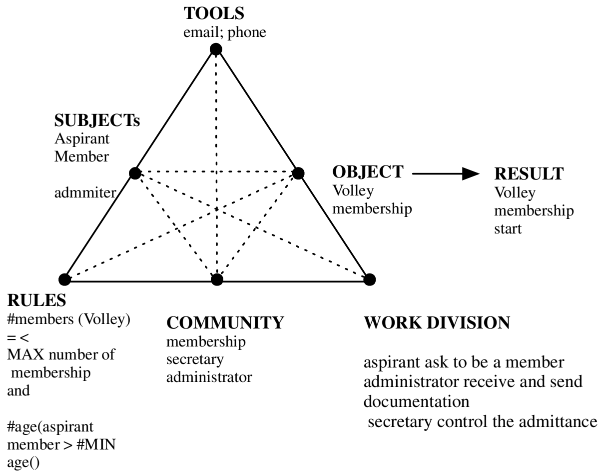


Fig. 2. Engeström Activity Diagram

It can be typified as being the first, second, third and fourth type The ***first order contradictions*** correspond to problems found in an internal element of a given activity. It occurs when you can isolate the manifestation of the occurred contradiction, diagnosing that it is due to a particular element of the Activity. The ***second order contradictions*** occur because the problem cannot be isolated and are related to the interaction between two or more elements of the activity. They are between the corners of the triangle and occur between the components of the activity system. The ***third order contradictions*** arise when conflicts can limit the development of the current activity in relation to a hypothetical activity, which is culturally more developed. Finally, the ***fourth order contradictions*** occur between the central activity system and the surrounding activity systems on the systems network and emerge from interaction of the central activity with peripheral activities. Most of the tensions occur in this situation, where usually a given activity is dependent on a result constructed by another.

4 Proposed Solution

From the viewpoint of the ontological, realization and implementation models of an organization, we follow the proposal made by Dietz [3] (e.g. mainly theorem of the organization and the distinction axiom in the Ψ theory) where the ontological model abstracts from all realization and implementation issues.

However the realization and implementation are bound to the ontological model. The realization starts with the business aspects of an organization (i.e., the ontological model) and comprises the detailed integration, through the layered nesting, of information and document necessaries to operation of organization.

According to our suggestion, we can analyse the implementation of an organization with the use of activity diagrams and use activity contradiction to find useful measures, viability norms and dysfunction. Our position is that the organization implementation is a result of an engineering process that can be analysed by a system of activities. This system can be used to understand technology (i.e., people, rules, division of work and tools) that is part of organization operation. For that we propose ontological model to capture the essential structure of activities from ontological organization model. The construct model of the model is stated in fig. 3.

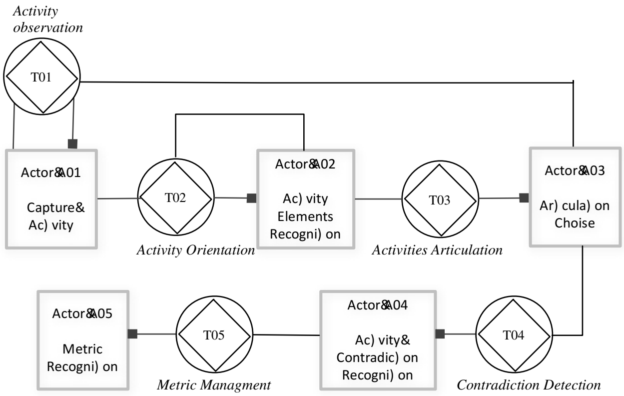


Fig. 3. Construct Model of proposed method

The model is composed of five transaction: T01, T02, T03, T04 and T05 and the associated executer actor: A01, A02, A03, A04 and A05. The actor role A01 has the responsibility to recognize all the transactions of the organization and map each activity. We start from the definition of an activity. This is a working unit with one objective. People involved organize their way of working in order to achieve the goal of the activity.

The actor A02 has the responsibility to capture detail of each activity, i.e. he/she identifies the elements of an activity, such as: actions, operations, subject, tools and business rules and social aspects. This task is iterative. It continuously tries to capture the missing or changing elements. Activity actions and operations constitute the "division of labour". It includes all actions and operations of an activity. The actions have the target to achieve a goal that is planned. An action to be carried out, people must take into account of the resources available in the environment and their affordances and constraints. These physical executions of actions are named operations, and they must be comprehended by the conditions given at the moment of execution.

We use DEMO concepts of phase and steps to structure actions and operations in an activity. We proposed that an activity has three main phases that correspond to the three phases of a transaction (Order-phase, Execution-phase and Deliver-phase). Each phase has a unique step with a well-defined goal. We link each goal action to step, therefore, the goal of an action is to fulfil the step. Operations represent the task that we do to achieve the action.

The actor A03 has the responsibility of identifying relationships between activities, notably defines different types of articulations considered to relate activities. To connect the activities we contemplated two types of interconnections that represent two types of relationships between activities. They are: a sequential relation and an inclusion relation. When activities have a sequential relationship, it means that the result of an activity is the object of another activity. In this case there is a temporal relationship between the first and the second activity and the second can only happen after the first produces its result.

4.1 Introduction Contradictions in the Organization Control Model

The introduction of activity in organizational control is done through contradictions discovered in transaction T04. The result of the contradictions discovery leads to identify the metrics that should be monitored, as well as the feasibility of these rules. This work is done by transaction T05 and executed by actor A05.

Contradictions will be discovered by analysing sets of elements of the activity. We propose to use three sets that encompasses relations between following elements of an activity: subject/tools, subject/object and norms/object. Each set will be associated to a metric mentioned below: i) Metric between subject and tools. This metric measures the contradictions that express the misalignment between tools and subjects to access the object of activity. It also express the support of execution of actions and operations (for example, what is the tool that is used to make access the updated list of the organization's products, and what people think it is his problems); ii) Metric between subject and object. It measures misalignment between subject and object of an activity by counting the cancelation of coordination acts. It is translated into calculating of cancellation of promise acts and cancellation of delivery acts and

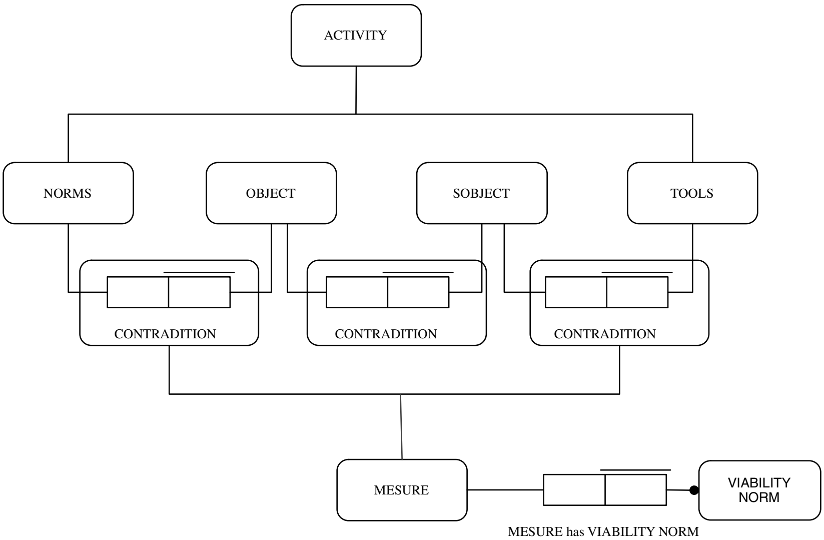


Fig. 4. New organization object fact diagram

iii) Metric between norms and subject, will measure the feasibility of achieving the result of the activity. This metric is the type usually associated to the business objetives (e.g., total revenue per month minimum activity will be 5,000 Euros).

Fig. 4 presents the proposed Organization Object Fact Diagram of organizational control. This model integreates the contradiction measue into organizational control presente in Fig. 1.

5 Case Study

We exemplify our method by using it to analyze a service named True-Kare. The main purpose of the True-Kare service is to facilitate the support given by the family or institutions to the elder person with some level of dependence. The service has two main steps: the purchase of mobile equipment and the service subscription, after the purchase of the equipment. To purchase the equipment a customer has to fill out a purchase form using the True-Kare portal. In effect, it is necessary to mention some personal information, including the equipment delivery address and billing information. When the customer receives the equipment, he must activate the service.

The starting point of our analysis is the Ontological Model of the organization, which was built using the DEMO methodology. Fig. 5 provides a general view of transactions, which in this example are transactions T1 (Equipment Order) and T2 (Equipment Payment). Both transactions involve the actors A1 (Client) and A2 (Organization). Transaction T1 is initiated by actor A1 and executed by actor A2 (i.e. the equipment Order transaction is initiated by the Client and executed by the Organization). Conversely, transaction T2 is initiated by actor A2 and executed by actor A1 (i.e. the Equipment Payment transaction is initiated by the Organization and executed by the Client).

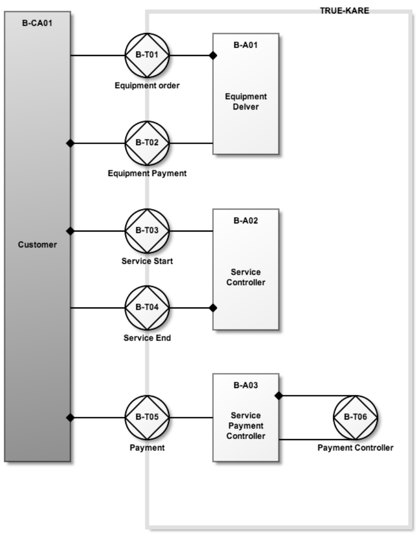


Fig. 5. Construction model of TRUE-KARE

5.1 Mapping Activity Diagram from DEMO Model

The agenda of actor A01 is to capture the relevant Activities of the services offered by TRUE-KARE. The DEMO Constructed model is analysed and for each transaction an activity is created. The result will be a mapping each transaction to one activity because: 1) the result of each transaction, i.e., it productions fact, contribute to bring about the mission of the organization and it can also me related with the outcome of one activity and 2) A transaction is a set of production acts/facts and coordinates acts/facts that occur in generic recurrent patterns that we consider it to discovery the relevant actions as operations inside each activity. With this proposed rule we redefine the concept of activity as: "An activity is a context necessary for understanding people work in an organization. By performing actions and operations, people make plans and fulfilled commitments to bring the output of an activity, i.e., the production acts". As per the definition we attain the following activities, describe in table 1.

Table 1. Discovery Activities

Activity	Name	Output	
AT-01	Equipment Order	CLIENT has order	EQUIPMENT
AT-02	Equipment Payer	CLIENT has paid	EQUIPMENT
AT-03	Service Start	CLIENT has start	SERVICE
AT-04	Service End	CLIENT has end	SERVICE
AT-05	Service Payer	CLIENT has paid	SERVICE
AT-06	Service Execution	CLIENT has used	SERVICE

After identifying the significant activities, we have to recognize the elements of each activity: subjects, tools, rules, community and division of work. Fig. 6 presents for each activity the elements founded, as the result of work done actor role A02 (see Fig. 3). For this we attach each production acts a goal that will be planned by actions and executed by operations.

Actor role A03 has the responsibility to articulate the activities. Following articulations are proposed, after analysing the relationship between the outcome and the objects of each activity: i)AT-01 /AT-03: sequential relationship. It means that to start a service client has first to order equipment; ii) AT-03 /AT-06: sequential relationship. It means that after the start of service, The client can use the service offered by TRUE-KARE ; iii) AT-05 /AT-06: inclusion relationship. It means that for using the service the client has to pay for it; iv) AT-06 /AT-04: sequential relationship. It means that during the execution of service the client can end it; and v) AT-02 /AT-01: inclusion relationship. It means that the equipment order, depend of the payer of equipment, i.e., client has to pay for the equipment before he/she can get it.

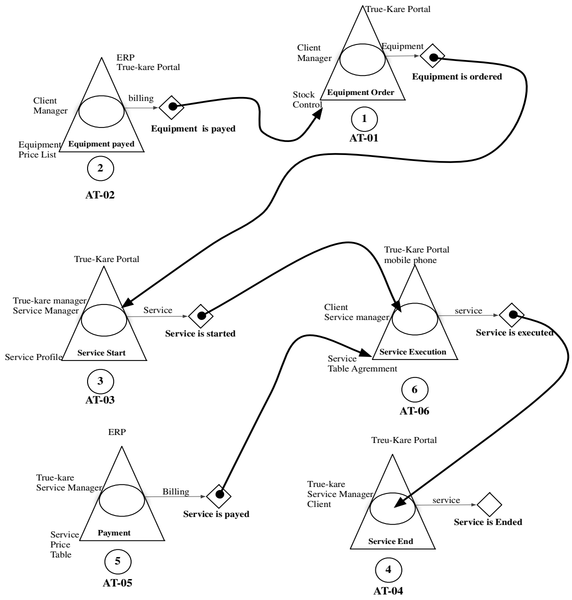


Fig. 6. True-Kare Activities Diagrams

5.2 Discovery Measure through Contradiction Analysis

During the analysis of activities we have identified different types of contradictions. Following proposal presented in section 4.1, they are grouped into several kind of norms: competence (which is measured in accordance with the capacity of the subject and the object of activity), tools (where it is analysed the mediation of subject and object through the use of tools) and finally tensions related to the objective of the activity (which is measured by the ratio of the rules with the purpose of the activity). Table 3 summarizes some of the manifestations of contradictions revealed while analysing the activities.

Table 2. Contradictions analysis

#	AT	CONTRADICTION	DESCRIPTION
C1	AT-01	SUBJECT-OBJECT	Contradiction between vendor and sales
C2	AT-01	RULE-OBJECT	Contradiction between price and equipment service.
C3	AT-01	SUBJECT-TOOL	Contradiction between True-Kare portal and client.

The contradictions analysed in table 2 propose rules and feasibility control of those rules in order to monitor the operation of the organization. These are presented in Table 3.

Table 3. Feasibility control

AT	RULE	CONTRADICTION	FEASIBILITY CONTROL
AT-01	N01 Number of proposed sended to clientes	C1	N2/N01 > 15% N01 > 10
AT-01	N02 Number of proposed accpeted		
AT-01	N03 work analysises	C1	N03 >= 3 (where 0 id bad and 5 is excellent)

6 Discursion and Conclusion

This article discusses how to use activity diagrams into organization control model. It is suggested a method to discover relevant Activity diagrams avoiding the steps of proposed and withdraw activities, based on analysing ontological transactions, from DEMO Ontological model. DEMO assists to find relevant activities because it helps to delimit the area of operation of an organization, through the concepts of components, environment and structure, presented in the models..

After capturing relevant activities diagrams from DEMO, it was possible to use them to take the manifestations of contradictions present when analyzing organizations. The results were to observe the activities and capture the type of rules that should be observed in the organization control field. The type of rules relates people capability, tools capability and feasibility organization rules. People capability has focuses on continuously monitoring people commitments to the organization objectives. Tools capability measures de misalignment between what tools offers and what people expect from tools. The goal is to be able to contribute for tool creation business value. Finally organization rules feasibility measures alignment between business and people in the sense that people are able to fulfil the outcome of an activity with the existing business rules.

In conclusion, with this propose we are able to define activities from the ontological model, which provides a basis for an initial analysis of people´s practices within an organization and contributes to finding relevant organization control measures since DEMO separates an organization into the ontological, the realization and the implementation aspects, it allows establishing a general model that is very stable that can support different kinds of activities.

References

1. Engeström, Y., Brown, K., Christopher, L.C., Gregory, J.: Coordination, Cooperation, and Communication in the Courts: Expansive Transitions in Legal Work. ISeminal Papers from the Laboratory of Comparative Human Cognition, pp. 369–388. Cambridge University Press (1997)
2. Aveiro, D., Silva, A., Tribolet, J.: Modeling the Function Perspective in ODE: Improving Unknown Exception Handling and Change, pp. 1–16 (2009)
3. Dietz, J.: Enterprise Ontology: Theory and Methodology. Springer (2006)
4. Vygotskiĭ, L.S., Cole, M.: Mind in society: the development of higher psychological processes. Harvard University Press (1978)
5. Engeström, S.A.: Discursive manifestations of contradictions in organizational change efforts: A methodological framework. Journal of Organizational Change Management 24, 368–387 (2011)

Pedagogical and Organizational Concerns for the Deployment of Interactive Public Displays at Schools

José Alberto Lencastre[1], Clara Coutinho[1], João Casal[2], and Rui José[2]

[1] Institute of Education, University of Minho, Portugal
[2] Centro Algoritmi, University of Minho, Portugal
{jlencastre,ccoutinho}@ie.uminho.pt,
{joaocasal,rui}@dsi.uminho.pt

Abstract. The integration of digital technology into schools is facing substantial challenges. In this study, we address the roles that public displays in communal spaces at the school can play as an additional vehicle for learning practices and, particularly, as a medium for sharing educational videos created by students themselves. Using a participatory design methodology, three hands-on workshops were conducted with teachers of a secondary school from the north of Portugal in order to elicit expectations towards the system design and its integration into school practices. Based on recorded video, direct observations, paper prototypes and open ended questionnaires created during the workshops, we conducted a systematic content analysis of teachers' concerns associated with the deployment of public displays at schools. Findings suggest that teachers are receptive to the technology and were able to generate scenarios that take advantage of the possibilities offered by digital public displays to stimulate learning processes. However, to accommodate the constraints of learning practices and organizational realities, the design of the system needs to take into account a broad range of sensitivities.

Keywords: Technology-enhanced learning, public displays, videos.

1 Introduction

Even though the potential impact of Technology Enhanced Learning (TEL) is widely recognized, the integration of digital technology into schools is known to face substantial challenges at multiple levels. Previous research has already identified a number of well-known issues regarding the adoption of new technologies at schools, and even suggests that most of the intended innovations related to the use of ICT in schools have not really achieved a substantial impact on school practices [1–3].

In our research, we are specifically addressing the roles that public displays in communal spaces at the school can play as an additional vehicle for learning practices. Goodyear [4] points to the fact that there is a shift in our sense of the spaces and contexts in which education takes place, as different learning activities are becoming more commonly distributed across a variety of contexts. Digital displays can provide a simple and effective way to generate shared experiences in public

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_41,

spaces, and when these displays are also interactive, they can be used to foster user-generated content that may enrich learning activities. More specifically, we are particularly concerned with the use of displays as a medium for sharing educational videos created by students themselves as part of learning activities. Our work is part of an on-going research project, called Juxtalearn, which aims to promote students' curiosity in science and technology through creative film-making, collaborative editing activities, and content sharing. The idea is to identify their learning difficulties or “threshold concepts” [5], i.e. concepts that constitute major learning barriers, and facilitate the learners understanding through the creation and sharing of explanatory videos. These videos, together with additional data, such as quizzes, and the subsequent engagement with viewers is what we call a video performance. Digital displays in the public space of the school aim to extend those video performances to a new learning context, promoting curiosity, offline social interactions and serendipitous encounters with those performances.

In this study, we seek to identify the main implications of integrating these video performances and the public displays into the school and its pedagogical practices. To address this goal, we have conducted hands-on workshops with a selected group of influential school teachers to elicit expectations towards the technology and explore possible features for the system design. In the end, we have coded all the output from the workshops to identify the range of concerns expressed by teachers. These concerns were analysed along what we believe to be the main challenges for a successful integration of the technology: the integration into the school as an organization; and the integration into the pedagogical practices of teachers. The resulting contribution is a systematic elicitation of the key concerns associated with the deployment of public displays at schools and their integration into learning practices and organizational realities. In the remaining of the paper we will briefly present related work, some initial research challenges, the methodology followed in the empirical study as well as the results obtained from data analysis.

2 Related Work

Digital technologies have been used in classrooms for more than 20 years. Research demonstrates that the effective integration of technology into classroom instruction can positively impact student motivation, engagement and interest in learning [6]. Akhras [7] studies have shown that incorporating interactive technology has positive effects on student performance. This integration must take into account the fact that online communication and digital technologies are increasingly an ubiquitous part of children’s daily lives, thus creating many new opportunities but also adding an additional layer of complexity regarding the increased complexity and diversity of the involving eco-system. While many students already have web-capable smartphones [8], Sun [9] observed they no longer seem them as a novelty. However, incorporating them into class is still a potentially effective way to increase student attention and improve their learning skills. Swan et al. [10] reported that students found that using mobile devices is fun and can make schoolwork easier.

The use of public displays at schools, outside the classroom, has also been explored in a number of ways. The Dynamo display system [11] is a large multi-user interactive display for sharing, exchanging and showing multimedia content in a communal room of a high school. Dynamo provides a GUI like interface accessible from various interaction points (wireless mice and keyboards) so that multiple users can interact with it at the same time. Students used Dynamo to display and exchange photos, video and music; to create a pool of public media that anyone could use; to stage performances to the audience in the communal room; to post notices to other users; to leave content as gifts to specific people; and to engage in group discussions and interactions. Instant Places [12, 13] has been deployed in a university bar and in a communal space of a secondary school. It allows users to contribute to the content that is displayed by specifying keywords on the Bluetooth names of their mobile phones. In our own previous work, we have already designed a conceptual prototype of a display system for video sharing at schools using data from these same workshops [14]. In this work, we abstract away from the specific features of the display system, and focus instead on the key organizational and pedagogical issues emerging from the process of integrating and exploring the potential of these technological devices in public spaces of basic and secondary schools.

Videos can also play a worthy role in the learning process. The benefits of digital video as a constructivist learning and pedagogical tool are widely emphasized in literature, ranging from motivational and learning-to-learn benefits to providing strategies of collaborative knowledge building and effective video-based learning. The versatility of this multimedia content, regarding its use in learning activities, has been highlighted by Chambel & Guimarães [15], namely through its potential for the integration with other media and the active role of students in the learning process. A video-based story is employed by Hmelo-Silver [16] as the context of a problem based learning, where that multimedia content is pointed out as a support for comprehension. Studies by Light & Polin [17] indicated the use of YouTube and similar websites as exciting tools to be used on classrooms. The same authors also mentioned that the video production made by students is a form of reflecting knowledge built. Koumi[18] states that video has the ability of stimulating the will to learn and inciting viewers to act, changing habits and attitudes towards learning. In this line of work, Casal [19] made a pedagogical intervention using video production as a process of reflexive learning with Portuguese secondary school students and verified its benefits in the promotion of motivation and autonomy in learning. Barthel, Ainsworth, & Sharples [20] proposed the use of shared multi-path video as a form of promoting collaborative and informal learning, and envisioned that the full potential of ubiquitous video platforms may follow the growing of novel applications and user interfaces such as tablet computers and smartphones with recording and sharing features.

In our work, video creations are the central element of a learning process designed to help students overcome specific learning barriers. However, they are not per se the main outcome of that process. The aim to promote the reflection and collaborative dialog that should emerge from first creating the video and then sharing it on a dedicated web space and on the public displays at schools.

3 Research Design

The use of public displays at schools for the specific purpose of complementing video performances represents a novel concept that requires further characterization and a clear framing of the respective challenges. In this section, we start by describing this context as well as the specific challenges that were presented to teachers and the set of participatory design sessions that were conducted to understand the pedagogical and organizational sensitivities of teachers to the overall concept.

3.1 Participatory Design Workshops

To gain a broader understanding of the issues involved with the deployment of the technologies described in the school setting, we followed a participatory design approach where we invited stakeholder teachers to participate on 3 design workshops, each lasting about two hours. In these workshops, teachers were lead to collaborative conceive a conceptual prototype of the displays for the school and to reflect about the implications of the technology for the school and or their teaching practices.

In the first workshop, the key concepts of the project were presented to teachers and the development of known public display features to the school context was discussed. In a hands-on activity, teachers were asked to express their understanding of the technologies' potential and their expectations regarding the project. Several ethical issues were raised and teacher expectations were assessed. The practice and desired pedagogical impact of the tool to be developed started to be determined. In the end, a questionnaire was distributed to enquiry about the envisioned uses of such system within the school and possible expectations regarding outcomes of its usage. More specifically the questions covered the following themes: reflect about the plausibility of the introduction of digital public displays in schools, consider the advantages and disadvantages of such technology, potential ethical issues, the use of mobile technologies and the design process underway.

In the second workshop, teachers refined their initial ideas. In small groups, they created a scenario based on the paper prototypes, followed by a brief period of presentation and discussion of the prototypes and scenarios created. Once more, a questionnaire was handed out at the end to enquire about the workshop activities and possible changes regarding the envisioned uses of such a system within the school.

In the final workshop the team presented for evaluation a conceptual prototype designed from the results of the previous two workshops. In small groups, the teachers were asked to comment to what extent the prototype corresponded to their ideas, as expressed in the previous workshops. This was followed by a discussion about ways to initially engage students with the system. In the end, a questionnaire was handed out to enquire about how teachers envision the acceptance of and contribution to the platform by the school community (colleagues and students) and what educational outcomes might arise with the use of the system. All workshop sessions were videotaped and supplied evidence for data content analysis.

The data collected in the workshops includes: (a) notes from the conversations taken by the researchers present at the workshop; (b) the scenarios and paper prototypes developed; (c) the transcripts from the videos; and (d) answers to the questionnaires. To analyse this data, we followed a content analysis process [21]. We started with a "brief reading" of the transcribed text for general evaluation and then use used the MAXQDA software to code any references to organizational or pedagogical concerns present in the material. These codes transform raw data by clipping, sorting and aggregation, thereby establishing categories according to the most common topics found. In an attempt to reduce the error of the encoding process, a second researcher who knew the context and the content ratified the analysis consistency (intercoder reliability) [22].

4 Results

This section summarises the key ideas emerging from the data analysis, aggregated under the two dimensions of our study: the integration into pedagogical practices and into the school as an organisation.

4.1 Pedagogical Practices

In general, teachers were positive about the pedagogical potential of the displays. In particular, public displays were seen as a technology that could complement traditional forms of learning by creating new opportunities to stimulate overall curiosity about science topics. One teacher also referred that the display would be an opportunity to show what goes on in the classroom, like an extension of the classroom: "*It's almost like an exhibition where we show what we're doing (...) and usually students enjoy to see their work exposed and viewed by colleagues and by the educational community in general*" [PROF2 (13:41)]. Teachers were also positive about video creation as a playful activity for students. The example of a Parkour video to reflect about the laws of physics was seen as good example [Prof1 (27:54)]. Still, the use of video creation and sharing in this particular context has also raised several concerns, with the most salient being: students' autonomy and fairness.

Students' Autonomy. Regarding video creation, we clearly explained to teachers our assumption that the production side was to be done entirely by students and that teachers were not even expected to instruct students on video editing techniques. Their role would be to define an instructional scaffolding to support the learning process in which the creation of the video was just a tool. There were, however, different perspectives to the meaning of this scaffolding and the level of autonomy given to students. Some teachers focused more on the certification of scientific quality : "*The videos must be made by the students, but the scientific guarantee has to be given by the teacher*" [Prof10 (10:09)] or "*Validate the content should be the responsibility of the disciplinary groups*" [Prof10 (24:27)], while others focused more on the learning context in which the video was produced: "(...) *when we ask a student to do a work and we follow him in the development process, in some way we are already guiding him in order to have an adequate scientific product*" [Prof4 (13:35)].

Regarding video sharing, and assuming that the videos were created as part of a learning process, teachers seemed to perceive that responsibility could be given to students, also as a way to increase their engagement with the process: "*experience tells us that when we give responsibilities to the students (and they realize the limits), they succeed well*" [Prof2 (10:20].

Fairness. Another pedagogical concern was the potential of the technology to avoid or reduce a digital divide between students. If the pedagogical videos are something that all students 'have to do' this could be a problem, because the teacher would then have the responsibility to ensure equal experiences to all, i.e., provide appropriate training and equipment (video cameras) for all of them and that is something that the school and the teachers were not prepared. A teacher said: "*My concern is with the fact that not all students have computers or smartphones to interact with the display. (...) Why not use touchscreens so that all students can use the system*?" [Prof4 (21:17)]. While the introduction of a touchscreen as a complimentary interaction device was appreciated, the teachers also recognised that it would not solve the problem because it would not have the same features. Also, the students' smartphones are increasingly a viable tool for shooting video clips without the need for any dedicated camera, but once again not all students have access to appropriate devices and those who did not would be clearly in disadvantage.

4.2 Organizational Concerns

The possibility to integrate public displays into the school was perceived as being positive to the school itself. There was a clear association between the use of innovative technologies with the image of the school as being dynamic and modern. Teachers also emphasized that this type of technology could help to capture students' attention and refocus some of their leisure time on school activities. Still, it was also clear from the workshops that the introduction of a high visibility digital artefact, such as a public display, into the school setting could be a disruptive process for existing organizational rules and power balances. In particular, three main topics have emerged from the analysis of the teachers observations: shared display ownership, responsible use and identity/membership.

Shared Display Ownership. To some extent, a public display that is deployed in a communal space of the school is perceived by most people as being endorsed by the school board. This perceived ownership creates the expectation that the school board should control or endorse content shown on that display. However, this institutional control of the displays would conflict with our need to share the displays with teachers and students so that they can integrate them as a tool for their learning processes. During the workshop, three different types of stakeholders were identified as potentially interested in having some control over the displays and forming three interest groups [Prof1 (01:50)]: "*there is one regulated by the school board in order to use the display to provide institutional information (classes, school rules, schedules, and so on); there is a pedagogical group with teachers using the display as* *a pedagogical tool; and there is the students group, with a more ludic interest*". However, teachers have also recognized that managing these groups and the content they would drive to the displays would be very challenging as they would need the collaboration of many teachers who are increasingly overwhelmed with work. The teachers talked about workload: "*when the assigned group is too small, it does not work because this group does not have enough time. But if you are going to allocate a teacher or two ... never again, the work will never be complete*" [Prof1 (3:03)]. Again, this hints at solutions in which some students assume the role of content managers at least in selecting pre-defined content and according to established rules.

Responsible Use. While student engagement and interaction is seen as key to the success of the system, it also opens the door to the publication of offensive content, not just as part of the videos themselves, but also as comments or some other form of feedback. The possibility to comment the videos has sparked considerable discussion on the merits of an open approach. On the one hand, the richness of open comments could be one of the motivating factors for sharing the videos and gaining feedback. On the other hand, many teachers feared that allowing anyone to make comments without moderation, even if these were not shown on the display, would lead to offensive language, spamming or even be used as a vehicle for bullying. The possibility to vote (positively with a 'like' or negatively with a 'dislike') was also seen as possibly creating frictions between school groups. The teachers felt that the school had a very strong ethical obligation to avoid those situations and ensure that no susceptibilities were hurt [Prof3 (21:45); Prof8 (22:17)]. This ethical obligation was also extended to privacy. Very significant segments of the school community are young children and it is not easy (or permitted) to advertise images without direct authorization from parents or legal guardians [Prof8 (22:17)]. Teachers emphasized the need of a disclaimer in the school's website warning about this. Another ethical issue had to do with the need of protecting minority groups and facilitating diversity of opinions. Teachers pointed out the need to explore mechanisms to protect the diversity of the schools' populations and try to promote the expression of different opinions, topics, likes and dislikes. The idea of a set of guidelines and rules of good system usage was encouraged, along with the possibility of banning from the system students that would not comply with those rules.

Identity and Membership. A key requirement identified during the workshops was the need to associate strong identities with user profiles and assert their membership as school students. For the school board, this is fundamental as it is the only way to make people accountable and ensure compliance with any rules that may be established to manage the shared use of the displays. Identifying those who are members of the school community is also of the utmost importance for multiple reasons, in particular to allow access to the content and display resources.

It was clear that groups should also be an integral part of the system functionality. Students are not just seen as individuals, but also as members of formal or informal communities with strong impact in interaction patterns and organisational models, e.g. class, school year. The public displays should stimulate the perception of belonging to a community and should leverage upon communities as a way to drive engagement. One teacher said [Prof2 (6:44)]: "*I think the community spirit is more important than the individual*" (many teachers agreed). However, as a shared medium, the display should increase mutual awareness between those communities and provide an opportunity to break away from self-imposed content bubbles. In particular, for a school community it could play an important role in breaking traditional discipline or year boundaries and create a broader sense of community. For example, a student of literature sees a good post on electronics and may feel motivated to discuss it with colleagues: "*Back in my school the personnel from electronics made a spectacular thing! I have to read the book Lusíadas while others do an extraordinary electronic work...*" [Prof7 (14:26)]. This strengthening of social ties allows "*each student to belong to more than one group*" [Prof1 (00:30)], what is good for the school environment.

Managing identity and membership has revealed to be a complex issue because of the complexity to manage enrolment. This can be a very cumbersome issue when schools do not already have a formal digital identity service in place. Many ideas have been discussed to create a process that enabled all school students to have a strong identity in the system that also asserted their membership to groups. The use of institutional mail addresses was a possibility, but their real usage by students is very low. School enrolment processes could be extended to support these features, but they have many dependencies with other processes and are difficult to adapt. A specific endorsement process could recognize external identities as members of the school community, but that would be a process very complex to manage. Essentially this has remained as an open issue for which no clear solution was identified.

5 Conclusions

The results of this study suggest that digital public displays might have an important role in enticing learners to explore themes and share learning content within schools' public spaces. The teachers we engaged in the study with were able to envision the pedagogical potential, in particular as a way to develop student´s autonomy and sense of responsibility but showed some concerns on issues of equality of access to video recording technologies to as a way to transcend the digital divide.

At the institutional level, besides recognising the innovative image the deployment would bring to the school some organizational issues were addressed extending previous research on adoption of digital displays in K-12 schools, in particular associated with ownership and control of quality and adequacy of the contents shown on the display. For some participant teachers the possibility to comment the videos arise concerns on the challenges and opportunities of an open approach. Ethical issues were always latent in teachers´ words and the need to explore mechanisms to rule and protect the diversity of opinions in order to enhance the feeling belonging to a community where students were seen not as individuals but as stakeholder members.

The workshops were useful to generate ideas and make the design team aware of the specific challenges related to the school's context. Furthermore, the workshops were also able to create a sense of partnership between teachers and the design team towards the design and development. We believe this partnership will be crucial to foster wider adoption by the schools' communities and we now plan to move towards the deployment of a functional prototype in the schools. We will be able to collect data about the adoption and use of the different features included in the prototype (logs of the system, questionnaires and interviews). The data collected in conjunction with further design workshops with different groups of relevant users (teachers, students, and other potential stakeholders) will lead to further iterations of the design.

Acknowledgments. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 317964 JUXTALEARN. We would also like to sincerely thank all the teachers from Escola Secundária de Alberto Sampaio (Portugal) that agreed to participate in the participatory design workshops.

References

1. Buckingham, D., Willet, R.: Digital generations: Children, Young People, and New Media (2006)
2. Coiro, J., Knobel, M., Lankshear, C., Leu, D. (eds.): The Handbook of Research on New Literacies. Erlbaum, NJ (2008)
3. Snyder, K., Panichi, L., Lindberg, O.: On the issue of quality of experience in technology supported learning. Acad. J. 3 (2010)
4. Goodyear, P.: Emerging Methodological Challenges for Educational Research. In: Methodological Choice and Design, pp. 253–266 (2011)
5. Barradell, S.: The identification of threshold concepts: a review of theoretical complexities and methodological challenges. High. Educ. 65, 265–276 (2012)
6. Reiners, P., Renner, K., Schreiber, J.: The Effect of Technology Integration on Student Motivation, Engagement and Interest (2005), `http://kr012.k12.sd.us/Portfolio/GroupResearchPaper.pdf`
7. Akhras, C.: Interactive Technology: Enhancing Business Students' Content Literacy (Original Research Article). Procedia - Soc. Behav. Sci. 83, 332–336 (2013)
8. King, L.: The Teaching Revolution (2012), `http://www.taolearn.com/articles/article62.pdf`
9. Sun, J.: Influence of polling technologies on student engagement: An analysis of student motivation, academic performance, and brainwave data. Original Research Article. Comput. Educ. 72, 80–89 (2014)
10. Swan, K., Hooft, N., van't Kratcoski, A., Unger, D.: Uses and effects of mobile computing devices in K-8 classrooms. J. Res. Technol. Educ. 38, 99–113 (2005)
11. Brignull, H., Izadi, S., Fitzpatrick, G., Rogers, Y., Rodden, T.: The introduction of a shared interactive surface into a communal space. In: Proceedings of the 2004 ACM Conference on Computer supported Cooperative Work, CSCW 2004, p. 49. ACM Press, New York (2004)
12. José, R., Otero, N., Izadi, S., Harper, R.: Instant Places: Using Bluetooth for Situated Interaction in Public Displays. IEEE Pervasive Comput. 7, 52–57 (2008)

13. Otero, N., José, R., Silva, B.: Interactive Public Digital Displays: Investigating Its Use in a High School Context. In: Proceedings of the Workshop on Social and Mobile Computing for Collaborative Environments, SOMOCO (2012)
14. Otero, N., Alissandrakis, A., Müller, M., Milrad, M., Lencastre, J.A., Casal, J., José, R.: Promoting secondary school learners' curiosity towards science through digital public displays. In: Proceedings of International Conference on Making Sense of Converging Media, AcademicMindTrek 2013, pp. 204–210. ACM Press, New York (2013)
15. Chambel, T., Guimarães, N.: Learning with Video in Hypermedia, Lisbon (2001)
16. Hmelo-Silver, C.E.: Problem-Based Learning: What and How Do Students Learn? Educ. Psychol. Rev. 16, 235–266 (2004)
17. Light, D., Polin, D.K.: Integrating Web 2.0 tools into the classroom: Changing the culture of learning, New York (2010)
18. Koumi, J.: Designing Video and Multimedia for Open and Flexible Learning. Routledge - Taylor & Francis Group (2006)
19. Casal, J.: Construtivismo tecnológico para promoção de motivação e autonomia na aprendizagem. XII Congresso Internacional Galego-Português de Psicopedagogia (2013)
20. Barthel, R., Ainsworth, S., Sharples, M.: Collaborative knowledge building with shared video representations. Int. J. Hum. Comput. Stud. 71, 59–75 (2013)
21. Bardin, L.: Análise de conteúdo. Edições 70, Lisbon (1977)
22. Coutinho, C.P.: Metodologia de Investigação em Ciências Sociais e Humanas: Teoria e Prática, Almedina, Coimbra (2012)

Application of Virtual and Augmented Reality Technology in Education of Manufacturing Engineers

Jozef Novak-Marcincin, Miroslav Janak,
Jozef Barna, and Ludmila Novakova-Marcincinova

Technical University of Kosice, Faculty of Manufacturing Technologies,
Bayerova 1,08001 Presov, Slovakia
{jozef.marcincin,miroslav.janak,jozef.barna,
ludmila.marcincinova}@tuke.sk

Abstract. Method of dialogue of person with computer is named interface and Virtual Reality (VR) is newest of row this interfaces. Applications of VR in area of computer games was first and in this time is rise need to exercise this technology in another areas: industry and education. Nowadays VR are used by many universities, education and industrial companies as well, because they offer lower operating costs than the real situation. In other disciplines where training is necessary, simulations have also offered big benefits. Augmented reality (AR) system generates a complex view where the virtual areas are covered by real environment and offers the basic working place for the user. VR and AR are established itself in many disciplines of human activities, as a medium that allows easier perception of data or natural phenomena appearance. Therefore the education purposes seem to be the most natural ones.

Keywords: virtual reality, augmented reality, education.

1 Introduction

With the advent of high-resolution graphics, high-speed computing, and user interaction devices, virtual reality (VR) has emerged as a major new technology in recent years. An important new concept introduced by many VR systems is immersion, which refers to the feeling of complete immersion in a three-dimensional computer-generated environment by means of user-centered perspective achieved through tracking the user. This is a huge step forward compared to classical modeling and CAD/CAM packages, which inherently impose major imitations on intuitive user interaction. VR technology is currently used in a broad range of applications, the best known being flight simulators, walkthroughs, video games, and medicine (virtual surgery). From a manufacturing standpoint, some of the attractive applications include training, collaborative product and process design, facility monitoring, and management. Moreover, recent advances in broadband networks are also opening up new applications for telecollaborative virtual environments in these areas. In addition to these areas, the paper explores some connections between VR and computer vision (especially camera self-calibration and stereo vision) in the context of depth recovery

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_42,

in virtual manufacturing. Some of the automation techniques resulting from these concepts can potentially reduce a lot of time and boredom for users involved in manually creating CAD-based virtual environments. Lately, with the emergence of complementary areas of virtual reality such as augmented reality (AR), one can address crucial problems of registration between virtual and real worlds [1, 2].

2 Virtual and Augmented Reality Presentation

At the beginning of 1990s the development in the field of virtual reality became much more stormy and the term Virtual Reality itself became extremely popular. We can hear about Virtual Reality nearly in all sort of media, people use this term very often and they misuse it in many cases too. The reason is that this new, promising and fascinating technology captures greater interest of people than e.g., computer graphics. The consequence of this state is that nowadays the border between 3D computer graphics and Virtual Reality becomes fuzzy. Therefore in the following sections some definitions of Virtual Reality and its basic principles are presented.

Virtual Reality (VR) and Virtual Environments (VE) are used in computer community interchangeably. These terms are the most popular and most often used, but there are many other. Just to mention a few most important ones: Synthetic Experience, Virtual Worlds, Artificial Worlds or Artificial Reality. All these names mean the same [3, 4]:

- Real-time interactive graphics with three-dimensional models, combined with a display technology that gives the user the immersion in the model world and direct manipulation.
- The illusion of participation in a synthetic environment rather than external observation of such an environment. VR relies on a three-dimensional, stereoscopic head-tracker displays, hand/body tracking and binaural sound. VR is an immersive, multi-sensory experience.
- Computer simulations that use 3D graphics and devices such as the DataGlove to allow the user to interact with the simulation.
- Virtual reality refers to immersive, interactive, multi-sensory, viewer-centered, threedimensional computer generated environments and the combination of technologies required to build these environments.
- Virtual reality lets you navigate and view a world of three dimensions in real time, with six degrees of freedom. In essence, virtual reality is clone of physical reality.

Although there are some differences between these definitions, they are essentially equivalent. They all mean that VR is an interactive and immersive (with the feeling of presence) experience in a simulated world – and this measure we will use to determine the level of advance of VR systems [5].

Augmented Reality (AR) is a growing area in virtual reality research. The world environment around us provides a wealth of information that is difficult to duplicate in a computer. This is evidenced by the worlds used in virtual environments. Either these worlds are very simplistic such as the environments created for immersive entertainment and games, or the system that can create a more realistic environment has a million dollar price tag such as flight simulators. An augmented reality system generates a composite view for the user. It is a combination of the real scene viewed by the user and a virtual scene generated by the computer that augments the scene with additional information. The application domains reveal that the augmentation can take on a number of different forms. In all those applications the augmented reality presented to the user enhances that person's performance in and perception of the world. The ultimate goal is to create a system such that the user can not tell the difference between the real world and the virtual augmentation of it. To the user of this ultimate system it would appear that he is looking at a single real scene [6, 7].

The discussion above highlights the similarities and differences between virtual reality and augmented reality systems. A very visible difference between these two types of systems is the immersiveness of the system. Virtual reality strives for a totally immersive environment. The visual, and in some systems aural and proprioceptive, senses are under control of the system. In contrast, an augmented reality system is augmenting the real world scene necessitating that the user maintains a sense of presence in that world. The virtual images are merged with the real view to create the augmented display. There must be a mechanism to combine the real and virtual that is not present in other virtual reality work. Developing the technology for merging the real and virtual image streams is an active research topic.

Milgram describes a taxonomy that identifies how augmented reality and virtual reality work are related. He defines the Reality-Virtuality continuum shown as Fig. 1.

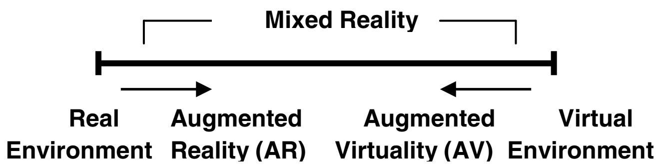


Fig. 1. Milgram's Reality-Virtuality Continuum

3 Some Examples of Virtual Reality Application in Education

For a long time people have been gathering a great amount of various data. The management of megabytes or even gigabytes of information is no easy task. In order to make the full use of it, special visualization techniques were developed. Their goal is to make the data perceptible and easily accessible for humans. Desktop computers equipped with visualization packages and simple interface devices are far from being an optimal solution for data presentation and manipulation. Virtual reality promises a more intuitive way of interaction [8].

The use of flight simulators has a long history and we can consider them as the precursors of today's VR. First such applications were reported in late 1950s, and were constantly improved in many research institutes mainly for the military purposes. Nowadays they are used by many civil companies as well, because they offer lower operating costs than the real aircraft flight training and they are much safer (see Fig. 2). In other disciplines where training is necessary, simulations have also offered big benefits. Therefore they were prosperously applied for determining the efficiency of virtual reality training of astronauts by performing hazardous tasks in the space. Another applications that allow training of medicine students in performing endosurgery, operations of the eye and of the head were proposed in recent years (Fig. 3). And finally a virtual baseball coach has a big potential to be used in training and in entertainment as well.

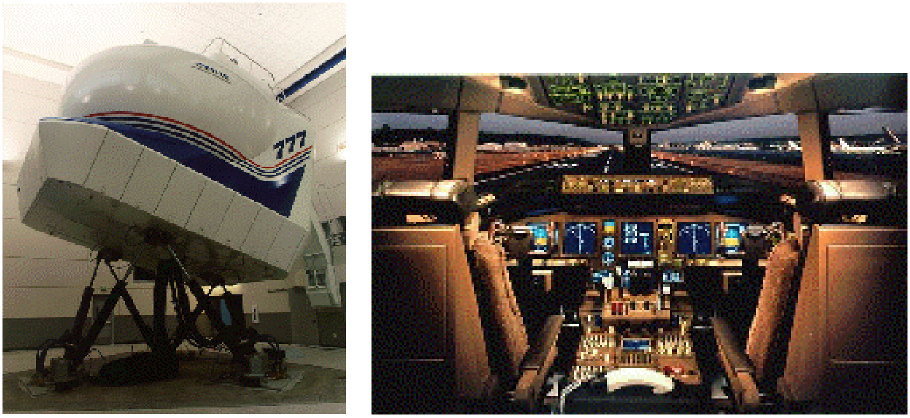


Fig. 2. Advanced flight simulator of Boeing 777: (a) outside view, (b) inside view

One can say that virtual reality established itself in many disciplines of human activities, as a medium that allows easier perception of data or natural phenomena appearance. Therefore the education purposes seem to be the most natural ones. The intuitive presentation of construction rules (virtual Lego-set), visiting a virtual museum, virtual painting studio or virtual music playing are just a few examples of possible applications. And finally thanks to the enhanced user interface with broader input and output channels, VR allows people with disabilities to use computers [9].

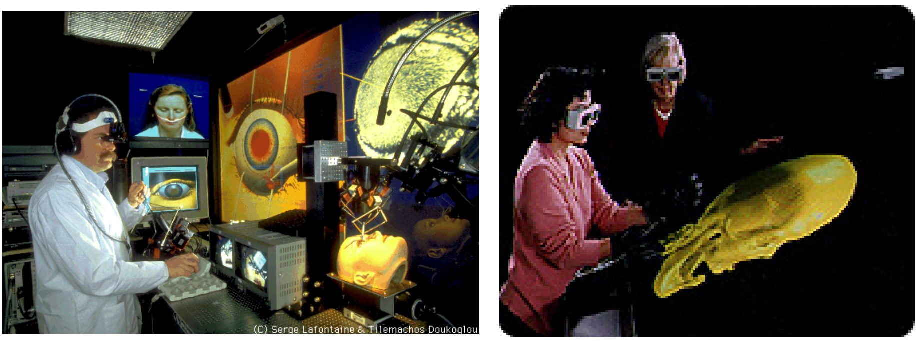


Fig. 3. VR in medicine: (a) eye surgery, (b) head surgery

4 Augmented Reality in Education of Manufacturing Engineers

Augmented reality, as a sub-area of the virtual reality, utilizes hardware and software tools for creation of the mixed environment that combines real scene usually present in the form of video sequence with augmented scene consisting of virtual models of additional objects. There are several techniques among those commonly used in augmented reality that are tried to implement into the system of robot programming education. The central object of newly created environment was robotic device from Swedish producer ABB - compact robot IRB 140. Robot is installed in laboratory of Department of Computer Aided Manufacturing Technologies of TUKE in Presov.

When creating the environment of augmented reality that would be suitable for realization of real-time tasks (for example the programming of the robotic device) is possible to consider the matter of continual space calibration in order to keep the real and virtual scenes spatially aligned in the final form of one consistent working place. For this purpose can be use well known 3D digitization device Kinect in combination with a specialized software tool called Skanect. Kinect firstly allows to obtain the 3D scan of the real environment and to use this real data while creating the virtual one. Entities from the scan serve as the references for generation of spatial links between computer models and real things, such as machines, tables, robot base. Secondly, the ability of sensing the workplace with depth sensors in real time provides of user with direct calibration of the environment and all included devices. This generation of mixed space and co-calibration of its both parts is realized in the software environment of Blender application. This powerful and widely complex graphical solution offers many useful tools, libraries and sub-routines for object programming with excellent level of graphical overview, everything based on the principles of Open Source software philosophy [10, 11].

For real-time detection of position of important objects can be used the technique of color markers. Thin paper stripes of different colors are stick on the surfaces of the robot and other devices (table, milling machine). They are either easily locatable by the camera and also suitable from the viewpoint of robot motion description. Color stripes are monitored in pairs, while each pair consists of two stripes which are one to other in upright direction. This way the stripes of equal color create a graphical marker which can be used for monitoring of exact position of each robot axis. Next the relevant command line called in the Blender environment activates the procedure of color tracking checking the position of real robot and recalculating the coordinates of individual motion axis in software environment. Position and orientation of robot model can be then adjusted or proposed (programmed) on the base of the real one. This means that the programmer can create the sequences of a new program also while proceeding from the positions used in old one or from actual robot position (possible program creation in mixed online/offline mode) [12, 13].

Other possibility to apply the color tracking is to use it for detection of workplace objects and also the task objectives. With a camera suitably located near the end of robot reach zone we can monitor the working area and achieve the location of bodies that are supposed to be handled. Blender integrates the ARToolkit features so it can perform even upgraded level of color tracking - shape detection. With this technique some parts of the programming sequences (basic moves of the robot to desired position) are written automatically, as the virtual model can be led to the objective while avoiding the obstacles detected in its working zone. Accurate finishing of there motions has to be programmed additionally. On the Fig. 4 is presented model of a robot inserted into the live video sequence on the base of the data acquired from combination of Kinect sensing and color tracking [14, 15].

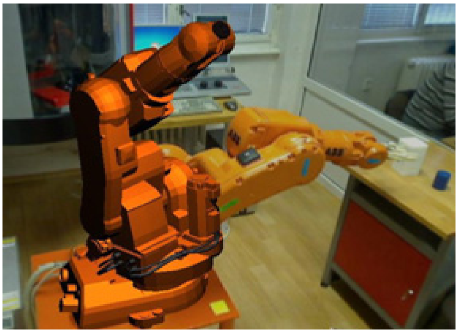


Fig. 4. Virtual model of robot in real environment

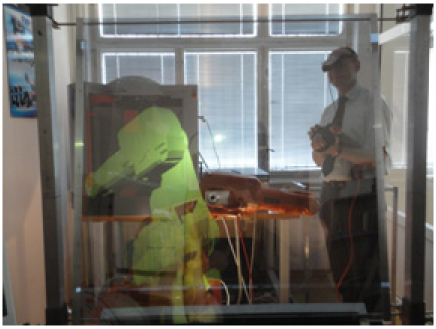


Fig. 5. Displaying of AR environment using the half-silvered mirror

New perspective possibility of displaying the environment of augmented reality is using of special visualization unit, which utilizes the principle of combined glass-mirror medium. The surface of the glass is either half-silvered or there is a half-leaky foil stick on it that creates a reflection and at the same time allows a view to the working environment with no obstacle or decrease of view quality. This commonly available kind of mirror is often used in gaming, medicine or business presentations. By optical connection of two seemingly different views it creates an ideal platform for the creation of a realistic spatial effect. Displaying is a reversed emission of the view to the reflex surface. It can be provided by computer monitor or classical projector placed over the working area. With a development of the projectors and their displaying technologies is possible to use the advantages of LED projecting. In comparison to the classical light projector the LED technology does not generate the luminous cone that would reflect in the form of light spot on the displaying glass. The setting up of whole scene becomes easier as you can mount the devices in the necessary displaying angles without the need to prevent the direct light reflection. On the Fig. 5 is presented use of the half-silvered mirror for presentation of virtual model of industrial robot activity in comparison with real robot workcell activity [16, 17].

5 Conclusions

The combination of information technology (IT), virtual reality, augmented reality and education activities has greatly changed traditional education processes. Many education tasks have been carried out as information processing within computers. For example, students can design and evaluate a new robot programs with virtual and augmented reality presentation without a real prototype of robot. As many activities in education process can be carried out using computer systems, the concept of virtual education (VE) has now evolved.

Acknowledgment. Ministry of Education, Science, Research and Sport of SR supported this work, contract VEGA 1/0032/12, KEGA No. 002TUKE-4/2012 and ITMS project 26220220125.

Agentúra
Ministerstva školstva, vedy, výskumu a športu SR
pre štrukturálne fondy EÚ

We support research activities in Slovakia/
Project is cofinanced from fonds of EU.
ITMS project code: 26220220125

Európska únia

References

1. Austakalnis, S., Blatner, D.: Real about Virtual Reality. Jota, Brno (1994)
2. Banerjee, P., Zetu, D.: Virtual Manufacturing. John Wiley and Sons, New York (2001)
3. Kalpakjian, S., Schmid, S.R.: Manufacturing Engineering and Technology. Prentice-Hall, New Yersey (2001)
4. Ong, S.K., Nee, A.Y.C.: Virtual and Augmented Reality Applications in Manufacturing. Springer, London (2004)
5. Lederer, G.: Virtual Manufacturing - Manufacturers Challenge of the 1990s. CIME - Computer Integrated Manufacture and Engineering 1(2), 44–46 (1996)
6. Vallino, J.: Introduction to Augmented Reality. Rochester Institute of Technology, Rochester (2002), `http://www.se.rit.edu/~jrv/research/ar/`
7. Vallino, J., Kutulakos, K.N.: Augmenting reality using affine object representations. Fundamentals of Weareable Computers and Augmented Reality, 157–182 (2001)
8. Tachi, S.: Virtual Reality in Robotics. In: Proceedings of the 8th International Symposium on Measurement and Control in Robotics, ISMCR 1998, Praque, pp. 29–34 (1998)
9. Lamar, K., Neszveda, J.: Average probability of failure of a periodically operated devices. Acta Polytechnica Hungarica 10(8), 153–167 (2013)
10. Barna, J., Fecova, V., Novak-Marcincin, J., Torok, J.: Utilization of open source application in area of augmented reality assembling processes. Manufacturing Technology 12(12), 2–7 (2012)
11. Novak-Marcincin, J., Janak, M., Barna, J., Torok, J., Novakova-Marcincinova, L., Fecova, V.: Verification of a program for the control of a robotic workcell with the use of AR. International Journal of Advanced Robotics Systems 9, 54 (2012)
12. Novak-Marcincin, J., Doliak, M., Hloch, S., Ergic, T.: Application of the virtual reality modelling language to computer aided robot control system ROANS. Strojarstvo 52(2), 227–232 (2010)
13. Novak-Marcincin, J., Barna, J., Janak, M., Novakova-Marcincinova, L., Fecova, V.: Production processes with integration of CA data in augmented reality environment. In: INES 2012 - Proceedings of the IEEE 16th International Conference on Intelligent Engineering Systems, art. 6249806, pp. 77–80 (2012)
14. Novak-Marcincin, J., Brazda, P., Janak, M., Kocisko, M.: Application of virtual reality technology in simulation of automated workplaces. Tehnicki Vjesnik 18(4), 577–580 (2011)

15. Novak-Marcincin, J., Novakova Marcincinova, L.: Advanced techniques for NC programs preparation. Applied Mechanics and Materials 389, 726–729 (2013)
16. Novak-Marcincin, J., Janak, M., Fecova, V., Novakova Marcincinova, L.: Verification of machine position in production plant with use of virtual reality technology. Applied Mechanics and Materials 308, 171–174 (2013)
17. Novak-Marcincin, J., Torok, J., Novakova Marcincinova, L.: New possibility of visualization in mechanical engineering. Applied Mechanics and Materials 442, 209–215 (2014)

Mobile Technologies Applied to Teaching: Experience in Healthcare

Raquel da Luz Dias, Márcia Cristina Moraes, and Leticia Lopes Leite

Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Brazil
{raquel.dias,marcia.moraes,leticia.leite}@pucrs.br

Abstract. Considering the significant increase in the presence of mobile devices in classroom and the opportunity to use these resources to improve and streamline the process of teaching and learning, this paper proposes a pedagogical use of mobile technologies, tablets, in class. The proposal was developed under the Project LabTEAR in its subproject LabsMóveis, and was applied in a graduating class of the Health area. The use of tablets in the university classroom contributed to the expansion of general and specific cognitive abilities stimulating creativity and leadership, enhancing the initiative to solve problems.

Keywords: education, mobile devices, health professionals.

1 Introduction

According to [1] the number of smartphones, tablets, laptops and phones with internet access will be greater than the number of people in the world by the end of 2013. Indeed, we have observed, every year, a significant increase in the presence of mobile devices in classroom. Considering this scenario, we believe it is necessary to include the use of these resources in differentiated pedagogical proposals that encourage and work with abilities like research skills, leadership, dialogue and collaboration .

Some authors have been working to develop tools for use with mobile devices. Peter et al. [2] present the study, planning and development of an Interactive Geometry tool to be used in mobile devices based on Android. Isotani and Orlandi [3] describe an authoring system for generating lists of exercises that can be accessed via smartphones and tablets.

Other authors have worked on the use of mobile devices in different areas of education. Batista et al. [4] present a study on the use of mobile phones in the disciplines of Calculus I using the plug-in MLE-Moodle. Forte et al. [5] present challenges and considerations about the proposed use of mobile devices to the field of medicine.

This paper presents a pedagogical proposal for use of tablets, which aims to promote collaborative learning and work abilities like research skills, leadership and dialogue. According to Moraes [6] students learn through research, engaging in joint productions with colleagues and thus reconstructing knowledge and practices.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_43, © Springer International Publishing Switzerland 2014

Therefore, it is necessary that students take the role of protagonists in this process and to share their research with colleagues through dialogue. Moreover, according to Fontes and Freixo [7], Vygotsky believes that social relations between individuals allow the development of a number of skills such as voluntary attention, logical memory and abstract thinking. In this sense, from both established interaction between the subjects with the environment, and between the subjects themselves, it establishes a dialogue from which cognitive development occurs.

The proposal to be presented was developed within the LabTEAR Project in its subproject LabsMóveis, and was applied in a graduating class of the School of Nursing, Nutrition and Physiotherapy at Pontifícia Universidade Católica do Rio Grande do Sul.

The paper is organized in four sections. Section two briefly describes the project LabTEAR and subproject LabsMóveis. Section three presents the developed pedagogical proposal and section four presents some final considerations.

2 Project LabTEAR and Subproject LabsMóveis

The expansion of the use of technological resources in university classrooms emerges as a necessity that, when implemented, can contribute to the improvement of educational actions. However, equipping classroom with technological resources is not enough, there needs to be a consistent pedagogical proposal that is in accordance with the Institutional Pedagogical Project (IPP). In this way, LabTEAR Project is an institutional proposal to respond the inherent challenges in the dialogue between technologies and teaching in higher education.

The LabTEAR aims: to know the possibilities of technological resources uses to support teaching; to create Information and Communication Technologies (ICT) laboratories to qualify classes at the University; to develop pedagogical proposals for the use of ICTs, aligned with the IPP and the strategic objectives of the University, and to investigate the impact of pedagogical proposals elaborated owing to its expansion.

In order to achieve your goals and enable the execution of the project and research from it in different areas of knowledge, it was formed a study group with members of the Faculties of: Education, Law, Nursing, Nutrition and Physical Therapy, Physical, Social Communication; and Engineering. This group is responsible for the execution of the first subproject of LabTEAR: the LabsMóveis.

The LabsMóveis aims to develop, implement and evaluate the appropriate use of mobile technologies in the University classroom. To this end, each faculty involved with the study group was awarded with a mobile laboratory. This laboratory consists of a trolley equipped with notebooks or tablets. The choice of devices that are part of the trolley was made individually by each Faculty, considering the different pedagogical practices to be adopted and the resources of hardware and software required to perform them. At the beginning of each semester, professors who are part of the group of studies, select which disciplines will be part of the subproject. The group performs throughout the semester, biweekly meetings to discuss and develop pedagogical practices to be used with mobile devices. After the development, the professors apply the practices in classroom and then present to the group the results of the established practice. These results are evaluated by the group and a systematization of the work is performed.

In 2013, the study group was composed by 22 professors, who selected 18 disciplines to apply the developed pedagogical practices. With the implementation of pedagogical practices in these disciplines, the subproject involved a total of 1213 students, being the proposal presented on this paper applied to one of the selected disciplines of LabsMóveis project.

3 Seminar with Video Production

This experience report was developed in a graduating class of the School of Nursing, Nutrition and Physiotherapy, under LabTEAR Project in its subproject LabsMóveis, more specifically in the discipline of Maternal and Child Nutrition, part of the 5th level of the Nutrition course.

The proposed activity was the "Seminar with video production." In its traditional form, the seminar was conceptualized according Gessinger [8] as:

> "A didactic procedure consisting in the study and discussion of a topic or problem presented by one or more students, under the supervision of professor. Looks to develop in students the investigative spirit, autonomy, critical thinking, reflection, cooperation, ability to express ideas clearly, among others."

Three distinct phases are required for the development of this methodology: preparation, development and final assessment. In this case, such steps have been adapted to the use of tablets in classroom, allowing the use of different languages for discussion in the large group.

The presence of academic content with a dense amount of information and often distant from reality or understanding of its relevance to students, stands out as one of the challenges in teaching. Give meaning to knowledge whose ultimate goal is still far from the previous experiences of the student is crucial to keep the student focused and interested in the activities proposed in class.

> “In order to broaden the discussion of content from many points of view and place the student as an active subject of your learning process, using the seminar as teaching methodology, can be an interesting alternative as it allows the student to perform analysis interpretation, criticism, raising hypotheses, assumptions search, obtaining and organizing data, comparing and applying facts to new situations [9].”

The adaptation of the steps composing this methodology to the use of tablets is described below.

3.1 Preparation

In the first step of the seminar, which lasted about thirty minutes, the professor introduced the new content to the class, the purpose and the methodology to be used for the presentation of the seminar, highlighting the use of technological resources and the use of languages such as audio and video. To do that, 30 tablets of LabsMóveis project were available to students. All tablets had applications such as video capture and wireless internet access at the University provider, inside and outside the classroom.

The 42 students who set the class, and the professor who led the activity, had basic computer skills and worked without the support of experts in the field. The class was divided into 8 groups spontaneously, 4-5 components, and each group was distributed a subject related to the central theme, who at the time was the "Right to Breastfeeding". The topics covered were: working women and breastfeeding; baby-friendly hospital; milking and storage of breast milk; drugs and breastfeeding; composition of breast milk, human milk bank; breastfeeding friendly primary care initiative and maternal nutrition and breast milk. The task was to research the issue and present it to the large group in the form of a video, which would be produced by the students themselves, through the use of video production available on tablet application.

Besides the production, the video should also be shared over the network, via the video sharing site YouTube. This fact brought to an additional commitment to the task execution, expressed by comments from students that highlighted concerns with the quality of material that would be produced.

To perform the activity, was distributed to each group one base text, which would be a starting point for research. Students were free to explore the subject as they wanted and to produce the videos at the location of their choice. As the class was large, many preferred to search for quieter places or to look for a "scenario" more interesting for their productions. After the orientation provided in the first step of the seminar video production, it was established the time to return to classroom for presentation.

3.2 Development

After the initial orientations, the students had 1 hour and 30 minutes to develop the activity. At this step, several groups were out of the classroom and others remained in classroom. This fact brought an initial discomfort by the professor, who realized that classroom was just a physical space and the students could construct knowledge in any other place.

This leads us to reflect on one of the principles of pedagogical mediation in the university classroom using ICT [10]: the principle of teaching presence and learning. According to Freitas et al. [10], "the pedagogical teacher presence, through listening and dialogue, in person or virtually, creates situations of interaction with a view to mobilizing the student for learning in the context of the integral formation of the person." This pedagogical principle gives new meaning to the act of teaching and learning, as the student does not need necessarily the physical presence of the teacher and this, does not need the student presence in the formal classroom space. A common kind of code is established, through the commitment of both parties in knowledge production. This is the initial spark for the mobilization of student learning.

Another characteristic strongly observed in the development of this methodology is mobility, which becomes possible due to the equipment. There are no limits to the creation and the student feels free to explore the different areas of the University and to access information in different ways. This refers to the concept of mobile learning, which according to Pereira [11], is defined as the use of mobile and portable devices used to facilitate access to information on education programs. Also according to Unesco [12] mobile learning involves the use of mobile communication technology to enable learning anytime and anywhere. Figures 1 and 2 illustrate the issue of mobility, showing two different environments used by groups for carrying out their work.

Fig. 1. Working moment of one of the groups around the campus

Fig. 2. Working moment of one of the groups in classroom

The development of the activity by the students occurred in different ways. Some groups chose to read the basic text first, search the major questions and, after mastery of content, starting to build the video. Other groups were concerned with, at first, to master the technology available, exploring the application of video production, establishing the video script and testing the video-sharing site. It was noticed that, regardless of the order of task execution, an important time was devoted to familiarization with technology. Many students had never had contact with the iPad features and others had never shared videos on YouTube. This required an additional effort from the professor to address those questions. However, as the students had more contact with the required applications, intimacy with technology increased. These students helped the colleagues who were still uncertain, with a visible democratizing access to information and high interaction among individuals, characteristics the principle of interaction and interactivity [10].

After the time established for the development of the task, in the preparation phase, all groups returned to classroom with their materials produced and posted on YouTube, ready for presentation. Following, they watched the videos produced and after each video presented, it was taken the discussion about the thematic discussed in each material. Interestingly, in each presentation, students demonstrating shame at being seen by all colleagues and showed a remarkable self-criticism. Comments regarding self-image, the way they speak and also to the content presented in the video were common. The discussion was rich, because the videos were produced with different approaches, simulating real everyday life situations, followed by orientations relevant to the content or creating hypothetical television programs or lectures, allowing viewers a view of content across different languages.

Some of the videos were produced: "Guidelines for milking and storage of breast milk" (avaiable in http://www.youtube.com/watch?v=BpsSa6OrmO8), "Baby-Friendly Hospital Initiative" (avaiable in http://www.youtube.com/ watch?v=GovuGUbT1_w), "Human Milk Bank" (avaiable in http:// www.youtube.com/watch?v=GkoU0c2h8hQ) e "Surgery in breasts and breastfeeding" (avaiable in http://www.youtube.com/ watch?v=8lpwiUMXzbA).

After discussion of the material produced, the activity headed for the last step, described below.

3.3 Final Assessment

After the presentations, a final assessment of the work done was performed with all participants. It was noticed increased student engagement with the content, when compared with the approach of the same content in a traditional way, through a lecture. In assessing students, the activity promoted a higher power of decision on the job and greater autonomy. Another interesting comment made by the students, was referring to the acquisition of new knowledge in the area of technology and the ability to improvise. Many students mentioned that the time for the development of the task was short, and that they could have made a better quality video. On the other hand, less time devoted to the mastery of content and construction material, simulated a real work situation in which the health care professional must be prepared to establish good communication and make appropriate guidelines and, mostly, in a time narrow for the users of the health system.

4 Final Considerations

In this paper was presented a pedagogical practice for use of mobile devices, tablets, in classroom. This practice was developed under the subproject LabsMóveis and applied to one of the classes of the discipline Maternal and Child Nutrition and it was a total of eight videos produced and posted on the YouTube video sharing network.

Considering the observations made by the professor throughout the monitoring activity, as well as the comments of the students, it is understood that the pedagogical work with the use of technologies in the form of seminar with video production contributed, in general, to the expansion of general and specific cognitive abilities stimulating creativity and leadership, enhancing the initiative to solve problems.

Moreover, the mobility afforded by tablets allowed both the students and the professor, understand that the process of knowledge construction would not necessarily occur within the physical space of the classroom. This process can occur in different locations, since the professor, as a mediator of activity, present a clear work objectives and what is expected to result from students, as well as follow students while performing the activity, clarifying doubts, both concerning content and the use of technological resources.

We understand that, although the pedagogical practice has been developed with students of the health sector, it can be adapted and used in different areas of knowledge, being necessary only some adjustments in terms of content to be worked.

According to Mota [13], while education and technology are connected, the teaching and learning using new technologies are still global challenges and without no clear answers. In this sense, this paper integrates the efforts of Computing in Education community to build educational possibilities for the use of technology, especially tablets, in classroom.

In terms of continuity of this work, we intend to expand the use of mobile technologies in the University classroom by adapting other methodologies, such as the case study.

References

1. Cisco. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012–2017. Disponível em (2013), `http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html`
2. Pedro, L.Z., Borges, S.S., Lopes, A.M.Z., Souza, J.P.T., Brandão, A.A.F., Brandão, L.O., Isotani, S.: Projeto e Desenvolvimento de um Aplicativo de Geometria Interativa para Dispositivos Móveis. In: 23º Simpósio Brasileiro de Informática na Educação, Rio de Janeiro (2012)

3. Orlandi, B.H., Isotani, S.: Uma Ferramenta para Distribuição de Conteúdo Educacional Interativo em Dispositivos Móveis. In: 23º Simpósio Brasileiro de Informática na Educação, Rio de Janeiro (2012)
4. Batista, S.C.F., Behar, P.A., Passerino, L.M.: Uso Pedagógico de Celulares: Análises de Estratégias Pedagógicas. In: XVI ENDIPE - Encontro Nacional de Didática e Práticas de Ensino, Junqueira e Martins Editores, Campinas (2012)
5. Forte, M., Souza, W.L., Prado, A.F.: Utilizando Dispositivos Móveis no Suporte ao Ensino de Medicina: Desafios e Propostas. In: XII Congresso Brasileiro de Informática em Saúde. Disponível em (2011), `http://www.sbis.org.br/cbis11/arquivos/797.pdf`
6. Moraes, R.: Aprender e Pesquisar: reconstruções propiciadas em sala de aula e em grupos de pesquisa. In: Diálogos com a Educação, Nilda Stecanela (eds.) EDUCS, Caxias do Sul (2012)
7. Fontes, A., Freixo, O.: Vygotsky e a Aprendizagem Cooperativa, Lisboa (2004)
8. Gessinger, R.M.: Seminário. In: A gestão da aula universitária na PUCRS, Freitas, A.L.S., Grillo, M.C., Lima, V.M.R. (eds.) EDIPUCRS, Porto Alegre (2008)
9. Anastasiou, L.G.C., Alvez, L.P.: Processos de ensinagem na universidade: pressupostos para as estratégias de trabalho em aula, Univille, Joinville (2004)
10. Freitas, A.L.S., Lima, V.M.R., Grillo, M.C.: Mediação pedagógica na aula universitária. In: Ensinar e Aprender com TICs: praticas de capacitação docente na PUCRS, Freitas, A.L.S., Leite, L.L., Lima, V.M.R. (eds.) Edipucrs, Porto Alegre (2012)
11. Pereira, M.: Utilização da aprendizagem móvel no ensino superior: um suporte ao ensino e a aprendizagem. In: X Congresso Brasileiro de Ensino Superior a Distância, Belém (2013)
12. UNESCO. Turning on mobile learning: in Latina American. Disponível em (2012), `http://www.unesco.org/new/en/unesco/themes/icts/m4ed/`
13. Mota, R.: Educação no Contexto das Tecnologias Digitais. In: Educação e Ciência Aberta: o impacto das novas tecnologias no apoio e distribuição de conhecimento, Computação Brasil. Disponível em (2013), `http://www.sbc.org.br/downloads/CB2013/computacao22_jul_2013.pdf`

Mobile Games for Children

Sílvia Martins, Catarina Silva, and Luís Marcelino

Escola de Tecnologia e Gestão, Instituto Politécnico de Leiria, Portugal
2121582@my.ipleiria.pt, {catarina,luis.marcelino}@ipleiria.pt

Abstract. This article presents a study on the use of educational games on mobile platforms, with touch screen, for children with ages between 2 and 3 years-old. The evaluation of the use of the games is based on the PLU model that proposes a three vector approach: Play, Learner and User Requirements. Hence, a set of paradigmatic games was selected that fulfil those three vectors defining the model. Experiments were carried out including the observation of the interaction of children during the use of selected games. The results show that the use of mobile technologies is extremely well accepted by children, and it was possible to reach a set of meaningful and relevant conclusions concerning the relationship between play / learning / user.

Keywords: Mobile platforms, learning, children, games.

1 Introduction

The use of computers and mobile platforms by children at home and in school environments is nowadays common in several countries. The exposure of a child varies according to the type of computer technology available, the child's age and sex, as well as the family social group [7]. Technologies like tablets and iPads have currently high rates of dissemination. Therefore, children are frequently exposed at an early age by their parents, relatives or teachers, using them with some frequency to work and entertain. In fact, measure of exposure of the impact of these devices is related to how parents and teachers present or implement strategies for using these technologies [9].

A potential consequence is that technology is rapidly changing the way children learn and have fun. Thus, the development of applications for children should consider the factors that affect children's cognitive abilities and take an active role in the realm of the methodological process of the design of the interaction [8].

The study presented in this article aims to evaluate the use of games in children aged 2 and 3 years-old. Although there are several studies with older children [10, 11, 13], our purpose is to effectively focus on younger children, since they are being increasingly exposed to this type of technology [12]. To clearly define the evaluation criteria, the work is underpinned in the PLU model, proposed in [2], defining different tasks from a set of selected games, paradigmatic of the three vectors that constitute the model: Play, Learner and user Requirements.

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_44,

The study involves the observation and further analysis of the behavior of 14 children aged between 24 and 40 months while playing different games. The analysis and conclusions presented here are extracted from the results of the interaction with a selected game set that attempts to span over the three PLU vectors, as well as the insights gained by observing the interaction of children.

In the next section we present some of the most significant work carried out with children and interactive technology, which fall in the area of this work, including some that make use of the PLU model. In Section 3 we describe in detail the proposed approach and experimental setup. In section 4 we describe with the implementation of the proposal, in Section 5 we present the results that we analyze in Section 5, and, in Section 6 we discuss the results of the interaction with the children. Finally, Section 7 presents the conclusions and some possible lines of future research.

2 State of the Art

We start by analyzing the abilities and skills of children in this age group, as well as their interaction with touch screen devices. Children of this age are said not to have yet refined skills in the use the mouse [8]. However, children realize that they can have an effect on the objects on the screen. As a consequence, touch screens are most suitable since interaction is performed directly on the screen [1] [4].

2.1 Children's Learning

Jean Piaget's theories were of great importance for understanding how children think and learn; therefore they can be very useful to frame the development of children's software. Children between 2 and 3 years-old generally have limited attention span, they acquire numerous skills, among which we highlight [1]:

- Place individual shapes on "form board" type puzzles;
- Use a pencil to replicate a vertical line;
- Match objects by simple shapes;
- Enjoy taking things apart and putting them together again;
- Ask "why" and "how" questions;
- Anticipate consequences and understand the impact of their own actions;
- Answering simple questions;

Although children acquire these skills at these early ages, competences acquired in the physical world may not have the same effectiveness with touch interactive devices.

2.2 Children and Games

All children love to play, but there are games that bring more interest and more motivation. Therefore, when developing software for children, the evaluation evaluation must take into account not only children's capabilities, but also e principles of children's motivation. Hence, we can make use of these principles as follows [1]:

- *Enjoyment* - children choose activities that they like to do, and avoid activities that are frustrating, static or boring
- *Control* – children avoid activities where they have no control. Software should increase children's sense of control with scenarios where their actions have impact
- *Interest* - children are more likely to engage in an activity when their interest has been sparked
- *Feeling of Competence* - children develop feelings of competence if they think they have a reasonable chance of success

2.3 Children and Touch Screen Devices

As children between 2 and 3 years-old love to play, learn easily, but have a natural difficulty handling the mouse, whilst understanding that their actions can have an effect on a screen, touch screen devices are a good alternative regarding computer interaction. These devices are now quite disseminated, so they can be used as learning tools. We must take account some relevant aspects regarding their use by children [4]:

- In general, all children are excited by such devices, but only this enthusiasm does not guarantee the interest on learning;
- Application interface is as critical as the platform and need to be intuitive to allow easy access to the child;
- Young children explore and learn in ways that are natural to them through touch, repetition, and trial and error;
- The evolution from novice to master is often achieved in the first experiment.

2.4 PLU Model

Conducting evaluations through questionnaires with children is a difficult job, since they often tend to respond in order to please the adults who interact with them, rather than on the question. The various models of product evaluation for children we have analyzed are all based on the observation of children's behavior. This work is underpinned in the Play-Learner-User model (PLU) [2], as this presents a metric with several vectors in an interactive application.

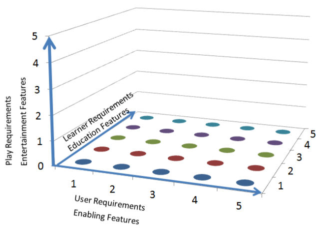


Fig. 1. PLU Model [5]

Almost all evaluation models are intrinsically subjective, while in opposition in this work we propose to define and use a more accountable approach. The PLU model was designed to help understand how children interact with technology. The first version was proposed in 2004 and later updated in 2008 [2] and in 2011 [5]. This model, depicted in Fig. 1, attempts to map the approximate relationship that children have with technology and with three types of interactive products. To reach this goal, technologies are classified according to their features: as entertainment, education and qualification. Each set of features map to one of the three vectors to evaluate:

- Play – Dimension fostered by the entertainment capabilities of the application. In this relationship, the child sees the product as a toy, therefore to fulfill its purpose the product should amuse or entertain. Examples of technologies include games and electronic pets.
- Learner – Dimension that corresponds to the educational component of the application. The interactive product is seen as a "teacher" who teaches, challenges and rewards. Examples include educational software (e.g. arithmetic or writing).
- User – Dimension given by the capabilities to produce results. The child sees the product as an interactive tool that, to be useful, must allow to "do" things. Examples could include drawing applications or calendars.

Based on these ideas, an evaluation framework can be put forward based on the PLU model, named PLU-E (Playing for Learning and Using Evaluation):

1. Decide on the purpose and focus of the product, to project goals and PLU targets;
2. Identify basic users and expert users;
3. Based on steps 1 and 2, set PLU weighs representing the product;
4. Define the moments in the development process evaluations should be done;
5. Based on steps 3 and 4 and on the design constraints (e.g., time and availability of users) evaluations can be planned and carried out.

3 Proposal and Work Plan

Based on the analysis of existing studies, this paper analyzes 3 types of educational games for children between 2 and 3-years old. We used a memory game, a paint game, and a puzzle. We modified both PLU and PLU-E models [3] [5], in order to introduce flexibility and easier results' analysis. Instead of using a normalized weighting (0%-100%), we implemented a qualitative value-based range from 1 to 5 for each vector ("Very Low", "Low", "Medium", "High" and "Very high").

3.1 Assumptions and Expected Results

Taking into account the studies and theories presented in the previous section, we defined several hypotheses to be verified experimentally:

H1. Children that have not yet had previous contact with touch screens devices, will nevertheless have ease in grasping the concepts of touch and drag;

H2. Children will have natural curiosity about technology and games and will take part in the experiment without any external compensation;

H3. Children that are more agitated are more likely to give up when difficulties arise. Given the short duration of the concentration ability at this age, most time consuming tasks may also be the cause of dropouts;
H4. Children prefer games with more game component;
H5. The time that each child spends gambling reflects the interest in the object of the game.

3.2 Selected Games and Tasks

Since the objective of this study is to evaluate children's exposure to technology, "traditional" games were adopted since the rules of the games were already familiar to the children. The types of games considered were memory game (joining pairs), puzzle, painting, search (identify something "hidden" in an environment), association (match or group related elements), and target acquisition (e.g. a game for "smashing balls"). Since the concentration span of children is reduced, only three sets were considered to limit the change of context of children. Physical versions of the chosen games were already present in the children's kindergarten. Several implementations of these games were installed and tested, and the following versions, available on Google Play, were adopted: Animals Memory (Task A), Kids Painting (Tasks B and C), Noddy Puzzle (Task D). In the definition of the tasks the PLU model was used, with the amendments referred in the previous section, as shown in Table 1.

Table 1. Selected games and characteristics

Game	Image	Task	PLU
Memory game with pictures of animals.		A. Identify the pairs of identical pictures and find all pairs;	P = 4 L = 3 U = 1
This game a puzzle of 6 pieces with images from the world of Noddy.		D. Build the puzzle according to the displayed image;	P = 3 L = 5 U = 3
Game to paint designs preexisting, where you can choose the colors or pencil "magic" that paints the correct colors.		B. Paint with the pencil "magic";	P = 5 L = 1 U = 5
		C. Paint with other colors chosen by the child;	P = 4 L = 4 U = 5

The memory game is a game that, in terms of interaction, involves simply the touch in the card to make it turn. It's an easy game to play, with immediate reward (pairs made disappear). For this reason it is considered to have a "high" (4) play value (P). Being a memory game, the game's success involves the use of related cognitive memory, so it was assigned the value "Medium" (3) in learning (L). At the end of the game the child does not get any significant result, therefore the User (U) component has a value of "Very Low" (1). The painting game involves touch (to choose the color) and drag (to paint), and can be played in two completely different ways. The choice of a "magic pencil" allows the "discovery" of the original design, as if the drawing becomes transparent. The choice of colors implies a fine motor coordination, since the lines can go over the outlines. As painting is an activity very well accepted by children the Play (P) value assigned to the game was "Very High" (5) for the game with the "magic pencil" and "High" (4) for the game with pencil paint colors. Learning (L) in the case of the magic pencil is "Very Low" (1) since the result is always the same while in crayons is "High" (4). As at the end of the game, the player has a painted design we considered the usefulness (User) as "Very High" (5). To the puzzle game we assigned a value of Play (P) of "Medium" (3), since it is a game where there is only "movement" as a result of interaction of the child by moving a piece. On the other hand, it is a game that requires observation and comparison of the different parts and fittings of identifying the location of the tiles (learning "Very High"). At the end of the game the child gets the complete picture and, therefore, it was assigned a utility value "Medium". Table 1 presents the games adopted, describes the tasks associated, and summarizes the PLU values adopted.

4 Developed Work

The work was composed of two stages:

(i) A questionnaire to parents and educators to know details about the children, e.g. previous contact with interactive touch devices or their usual reaction to given tasks;

(ii) Experimentation with a group of children aged between 24 and 40 months, 6 girls and 8 boys. Children played the selected games using a Samsung 7 inch tablet.

4.1 Questionnaire to Parents and Educators

In a first step surveys were conducted with the parents of the children, and we found that none of the children had regular previous contact with the type of devices. The main personal characteristics of each child were also registered, in terms of curiosity, calm or agitated behavior, shyness, social behavior, and difficulty in concentrating, based on information obtained from the educator. Of the 14 children that completed the test, based on information from educators: four are reported to have some difficulty in concentrating, which may be considered normal for children under 2 years of age. Two children are also identified as particularly shy and can be a barrier to interaction that accompanies the test.

4.2 Games with Children

In the second phase, there were several sessions of observation of children playing with a 7-inch tablet. This phase of experimentation was conducted in a quiet room, with no other children, but accompanied by the educator. During the session, each child was asked to play each of the games mentioned in the previous section, starting with a brief explanation of how to play each one. Each child has had contact with only one game at a time. During the child's interaction with the game the reactions of the child were recorded on paper, the time s/he took to understand what was asked and the time taken to reach the goal of the game. Additionally, the difficulties encountered by children were also registered.

Given the limited attention span of children at this age, individual sessions took about 15-20 minutes to complete the four tasks with the three games. However, some of the sessions lasted a bit more when the child's interest so justified. The same adult and the educator accompanied the execution of all the tasks. The interaction between the adult and some children was hampered by timidity and/or by the difficulty in concentrating on the requested task. During the experiments the following observations were registered:

- All children answered affirmatively when asked if they had previously known the game. However, none of them had used a tablet before and their knowledge was related to the physical game (cards, paper and pencil to paint and puzzle);
- The shyness in girls and difficulty in concentrating in boys influenced negatively the performance, leading some to give up;
- In the puzzle game, the greatest difficulty was the fact that the pieces placed in its final position may leave the place with an accidental touch, which forced them to put them back;
- The tablet device aroused much curiosity.

5 Results

5.1 Identify Pairs of Images (Memory Game)

In this game, of the 14 children involved, only one gave up, having identified only one pair after more than 1 minute. Fig. 2 shows the time each child took to complete the memory game. On the x-axis, the children are ordered by increasing age. The graph also includes a trend line to aid on the interpretation of the values obtained. One observation that can be made of the results is that the trend line is very flat, reducing only 20 seconds when the age increases by 16 months.

5.2 Paint with the "Magic" Pencil

The ease with which you paint the drawing with the "magic" pencil, where one only needs to swipe the finger, led the authors of this study to suspect that this would be one of the preferred forms of interaction. However, regarding the level of enthusiasm, the children showed a greater interest in the game where they were able to choose the colors, even with a resulting image distant from the original.

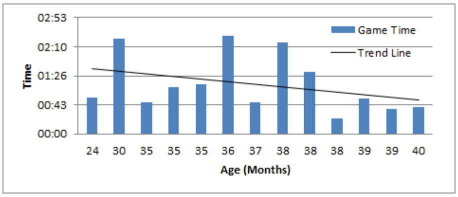


Fig. 2. Times of conclusion of the memory game ordered by age

5.3 Paint with Other Colors Chosen by the Child

Unlike the game of paint with the "magic pencil", the drawing with chosen colors is final only when the child so considers. Since there was no stipulated time for completion of the task, it would be acceptable that this task would take significantly longer than the "magic pencil". However, such scenario only occurred expressively in one of the children, while for all the others the times for both painting activities was similar. Fig. 3 shows the times that the children took in both painting tasks. Notice that the trend line exhibits a reduction of almost 1m30s corresponding to an increase of 4 months of age. Additionally to the playing time during the painting activities, the children's accuracy motility was also registered: a high motility results from caution when reaching contours and also from applying the right colors to specific areas; on the other hand, low motility results from scratching with colors without giving importance to the contours. More than half the children had not yet these skills (nine exhibited better motility skills, while four had still some challenges to overcome). The small difference between the times of painting in these cases can be explained by differences in the resulting paintings (less motility resulted in less accurate paintings).

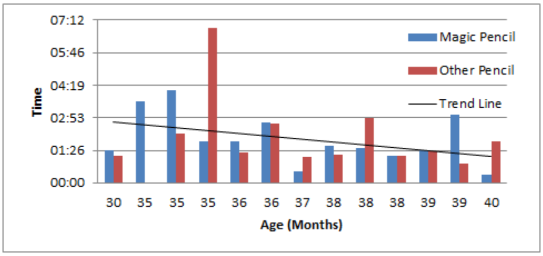


Fig. 3. Times for conclusion of the painting game ordered by age

5.4 Build the Puzzle

The puzzle game is the most cognitively demanding: it requires the identification of the location of the part in the final image and dexterity to put it in that position. Since the pieces placed in the end position can also be moved, sometimes the child inadvertently moved a piece out of place. Given the cognitive demand, it is normal that the trend curve is much more pronounced. As shown in Fig. 4, this line falls almost 3 minutes in 10 months.

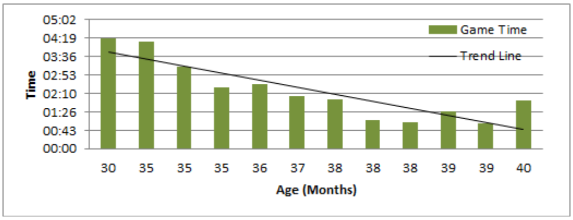


Fig. 4. Times for conclusion of the puzzle game ordered by age

Table 2. Children that gave up each task organized by gender (Withdrawals marked with X)

Game	Female		Male		Total		%
	X	√	X	√	X	√	
Memory	0	6	1	8	1	14	7
Painting A	0	5	0	8	0	13	0
Painting B	1	5	0	8	1	13	8
Puzzle	0	5	3	8	3	13	23
					5	53	<10%

6 Analysis of Results

After the experimentation phase, we reach the following analysis:

- The great majority of children has poor fine motor skills, natural given their age;
- In the memory game, it can be seen that neither gender nor age influenced the child's comprehension of the purpose of the game (see Fig. 3);
- In the painting game we see that age was not relevant to the comprehension or the skills needed (although final drawings were extremely variable). Analyzing completion times by gender, we can see that the completion time varied slightly;
- In the painting game, most children were more interested and thus preferred the task of choosing the pencils of various colors instead of the "magic pencil";
- In the puzzle game, we can observe that children older than 36 months had a better performance than younger ones (Fig. 5);
- We can also observe that while the puzzle game has a higher drop out percentage than the other games, the final percentage is less than 10% (see Table 2).

7 Conclusions and Future Work

Based on the work and analysis of the results obtained, we can draw a set of conclusions. Almost all children, despite having no previous contact with touch devices (H1), learned with ease the concepts of touch and drag. The child's gender did not affect the capacity or the ability to understand the goal of the game, but the preferences of each child were determinant. On the other hand, shyness influences the demonstration of skills and understanding. The difficulty in concentration reduces the acquisition of skills and understanding the purpose of the game. Another interesting finding is related to the age of the children. There are games where age is not a determining factor, e.g. the memory game or the painting game. It would be expected that the older children would have the memory more exercised, but it was found that the performances of smaller children were identical. On the other hand, the age factor has an influence on the puzzle games, since they need more understanding.

Future work includes the use of performance measures in the game to evaluate more objective hypotheses, to make them richer, more reliable and easier to obtain. It is also our goal to develop a game, based on the results with balanced PLU values, considering age and capacity characteristics of children in the specific age group.

Acknowledgments. This work would not have been possible without the collaboration of Jardim Infantil de Ourém and we present our acknowledgements, namely to the President Dr. Paiva, Pedagogical Coordinator Dr. Paula Neto and Educator Sylvia Castanheira. To all the parents of the children, we also thank the availability and the participation of their children. Finally, of course, the children who cooperated in the best way, with their so characteristic simplicity and curiosity.

References

1. Wolock, E., Orr, A.: Child development 101 for developers of interactive media. Active Learning Associates (2006)
2. Markopoulos, P., Read, J., MacFarlane, S., Hoysniemi, J.: Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers. Morgan Kaufmann (2008)
3. McKnight, L., Read, J.: Plu-e: a proposed framework for planning and conducting evaluation studies with children. In: 25th BCS Conference on HCI, pp. 126–131. ACM (2011)
4. Michael Cohen Group LLC. Young children, apps & iPad, U.S. Department of Education Ready to Learn Program (2012)
5. Read, J.C., Bekker, M.M.: The Nature of Child Computer Interaction. In: Proceedings of British HCI 2011. British Computer Society, Northumbria (2011)
6. Kremer, K.E.: Conducting game user experience research with preschoolers, Austin, Texas, USA (May 2012)
7. Straker, L., Pollock, C., Maslen, B.: Principles for the wise use of computers by children. Ergonomics 52(11), 1386–1401 (2009)
8. Agudo, J., Sánchez, H., Rico, M.: Playing games on the screen: Adapting mouse interaction at early ages. In: IEEE Int. Conf. on Advanced Learning Technologies, pp. 493–497 (2010)
9. Couse, L., Chen, D.: A Tablet Computer for Young Children? Exploring Its Viability for Early Childhood Education. JRTE | 43(1), 75–98 (2010)
10. Beyers, N.: A Five Dimensional Model for Educating the Net Generation. Educational Technology & Society 12(4), 218–227 (2009)
11. Lieberman, D., Fisk, M., Biely, E.: Digital Games for Young Children Ages Three to Six: From Research to Design. Computers in the Schools 26(4), 299–313 (2009)
12. Geist, E.: Qualitative Examination of Two Year-Olds Interaction With Tablet Based Interactive Technology. Journal of Instructional Psychology 39(1), 26–35 (2012)
13. Pohradský, P., Londák, J., Cacikova, M.: Application of ICT in Pre-school Education. In: 52nd International Symposium ELMAR, pp. 159–162 (2010)
14. Barendregt, W.: Evaluating fun and usability in computer games with children. PhD Thesis, Technical University of Eindhoven (2006)

Authoring Environment for Location Decision Games for Decision-Making Skills Development

Roberto Guimarães Morati Junior,
Orivaldo de Lira Tavares, and Crediné Silva de Menezes

Graduate Program in Computer Science,
Federal University of Espírito Santo, 01-9011, 29060-970, Vitória-ES, Brazil
{robertomorati,orivaldotavares,credine}@gmail.com

Abstract. Locative games are being used increasingly in larger scale to support the learning process. We are working on a specific subclass of such games, whose main focus is the learning support necessary for the development of decision-making skills. As purpose, we are building the project Clic&Ação-ARLA. This paper presents the conception of an extension of the aforementioned facing Authoring and Realization of Locative Adventures with a focus on decision-making skills development.

Keywords: authoring environment, locative games, location decision games, decision-making skills.

1 Introduction

Every time in our lives we make decisions and it is the most obvious skills in humans [7]. Decisions are simple, such as run or not to take a bus, or more complex as choose a house to live. We can engage with high-risk decisions. For example, the crucial decisions for the outcome of a military mission, or in scenarios where accidents can happen, such as civil construction.

This is present in our daily lives, learn to make decisions is not learned by accident. The humans normally make decisions based on emotions, which may not always be the best way. In this context, the school can assume an important role in building citizenship. So, there are providing opportunities for perform decision-making in simulated environments.

In order to maximize the chance of making good decisions, it is necessary to take into consideration the information available. Such information may change every moment, and depending on how the individual interprets this information, decisions can be more accurate.

With the purpose of the contribution to the development of decision-making skills, serious games have been constantly explored [2], [8]. However, most of these serious games are using different artificial contexts of real contexts known by players. Moreover, it has a limited social interaction, since the interaction between players does not occur naturally [3].

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_45, © Springer International Publishing Switzerland 2014

In order to contribute for the creation of opportunities to promote making decisions, we are working on an environment to allow the construction of a specific class of locative games, that we call LDG (Location Decision Games).

This paper aims to present specific editors with the purpose of the enabling the construction with high-level abstraction of the LDG. Specifically the class ALD (Alice in Labyrinth of Decisions) and other games with flexible contexts. These editors make up the Authoring Environment for Location Decision Games. This environment is developed in the context of the Clic&Ação-ARLA (Adventures Locative Augmented Reality) project that aimed at the creation with Android OS (Operating System) [9].

To describe the purpose, the rest of the paper is arranging as follows: In Section 2 the concept of LDG is presented. Section 3 ALD is presented. Section 4 presents characterization of the problem. Section 5 presents the Authoring Environment for Location Decision Games. Section 6 presents the conclusions of the paper and future work.

2 Location Decision Games

LDG are games where elements of the real world and its characteristics make up the game scenario using GPS (Global Positioning System). Allows the inclusion of elements of the virtual world in the real world, hence the players can visualize this in mobile device. Besides, it allows the interaction between players with more naturally. In this way, we can make compound scenarios and situations that react with the positions of the players, so to allow the game has following characteristics:

Dynamic Environment: means that the elements, i.e., the characters and monsters, should have an autonomous behavior in the scenario of the game.
Time: sometimes the game must impose a particular time for the player collect items or accomplishes a mission. Then, it induces the player to quickly plan his decisions.
Uncertainty: the autonomous behavior of the elements in the game makes the player take decisions with uncertain information. For example, a character that constantly changes his position, that is necessary to better the choices in his previous experiences to make better decisions in the future.
Collaboration: it is important for decision-making skills that players can dialogue between them and make collaborative decisions as well as discuss experiences. So, the game needs enable the players to play different roles with your mobile devices.
Suggestions: each scenario or situation in the game needs one or more suggestions. These are elements of perception the environment that help in players' decisions. The suggestions enable the player contextualize his choices of actions by his decision.
Story: The elements include text, image or audio. These elements are important in the plot of the game, because it allows the player to contextualize and understand better the results of his decisions.

These characteristics of LDG permit the development of decision-making skills. Such skills conceptualize from the model RPD (Recognition-primed Decision Making) [6], in three types: *Situation recognition* - the decision maker acknowledges the situation not only as similar/familiar or not, but also leading to typical actions; *Evaluation of a Set of Options* - in this phase the decision makes a review of the possible actions, in a typical situation. The process of analysis of actions is making by means of a mental simulation. *Mental Simulation* - that is simulating the decision maker by possible steps, results and possible problems. These problems can be found and treaties.

3 ALD - A Game with Emphasis on Decision-Making

Preciously, we built the ALD game based on the classic Hunt the Wumpus, proposed by Gregory Yob [10]. Our version of the game is in the pranks' universe, where the player tries to find a treasure in 3D (Three Dimensional) world with dangers and a scenario called "Labyrinth of Decisions".

In next step, this game is transposing to the world of mobile devices. Where that is called ALDloc (Alice in Labyrinth of Decisions Locative). In addition, the main elements of ALDloc are below:

Limit of the Labyrinth: this object has the purpose of restricting the geographic area in the game;
Cascumpus Monster[1]: if the player finds the Cascumpus, the obstacles and dangers of the game take on a new geographical location. Cascumpus is a monster that stole the Treasury of the Alice.
Healer Character: this character cures the burns player.
Alice: it is the main character of ALDloc who had her stolen Treasure.
Fire Obstacle: It causes burns in player around it.
Treasury Object: the player's mission is retrieving this item that was stolen from the Alice.

The game was also projecting to help the player with objects of indicate a better decision, in that case, those are: Heat, Stench and Shine - These objects indicate the presence of Fire, Cascumpus, Treasure, respectively, near the player. So, the game give one way for the player can think in a good way for find the Treasure and finish the game. And consequently, the ALDloc is played by means of a mobile device, where the player must avoid the obstacles Fire and Cascumpus, for this he will use the reference the obstacles Heat and Stench, and he should find the lost treasure.

4 Characterization of the Problem

Within the competence of environments for the construction of locative games, there are the fAR-Play (A framework to develop Augmented / Alternate Reality Games) and ARIS (Augmented Reality and Interactive Storytelling). The fAR-Play does not currently offer an environment for creation locative games. However, the ARIS was considering a great advancement in the conception of locative games [1], [5].

[1] Cascumpus: It is a smelly green monster that seeks to hinder the player's mission.

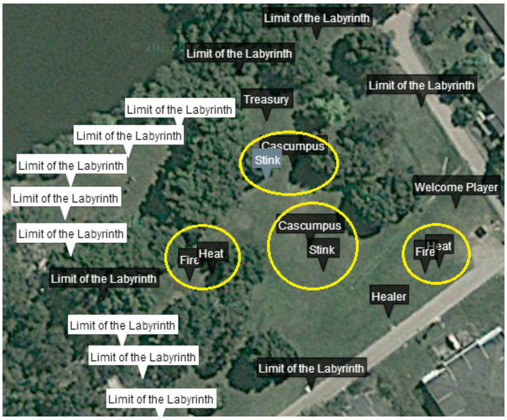


Fig. 1. ALD created with environment ARIS

We analyze the ARIS environment for build ALDloc and we found several difficulties in implementation, as we report below.

ARIS is an environment friendly and easy to use platform. However, it lacks mechanisms to delimit the geographical area of the game. This is bad, because when a player leaves his geographical area, the locative game must inform him.

So, it was necessary to distribute various objects of type "Limit the Labyrinth" to supply such deficiency (see in Fig 1). Other difficulties in the project are below:

The First. Creations of objects with types, i.e., character and obstacles, and limited information for decision making. For instance, the objects have no suggestions to help in contextualizing of the decision of the player and in the perception of danger near.

The Second. Creation of conditions is limited by pre-established rules for the environment. Part of dialogue with a character of the adventure can not make a rule. For example, the dialogue *"Yes, I will find your treasure!"* coming from the interaction with the character Alice in ALDloc can not be used to compose a condition to activate a mission (primary mission).

The Third. The environment allows only the use of static objects. However, in computer games there are objects with autonomous behavior (not static), for example, characters that act in accordance with the axioms of the game. Then, in conception of the ALDloc was necessary duplicate the static objects, and after even add rules to them. Hence, at one point of the game some visible objects become invisible and other objects become visible in the adventure. Then, it is simulating an adventure to the player with dynamic scenario (an attempt of simulate objects with autonomous behavior). Static objects duplicate (Treasure, Fire and Cascumpus) are marked by circles in yellow color in different geographical locations (see Fig 1).

5 Authoring Environment for Location Decision Games

This environment has the primary purpose to allow the construction of games that use real-world scenarios and virtual objects to create situations that requires the decision making. Then, it is necessary that the player becomes aware of the importance of use past information to make mental simulation of their actions and to make effective decisions. These are following the editors that make up the authoring environment of LDG. It is important to emphasize that due to space limitations, the description of the editors and their functionalities were summarizing in this paper.

5.1 Objects Editor

The Object Editor allows the creation of Objects that have an Object Type and Icon that represents the Object Instance on the map when it is creating the adventure. When the Object is dragging into the map, it becomes an Object Instance. In addition, the characteristics of the Objects Instances (as images, sounds, suggestion and dialogues) can be considered as the core of learning. Then, the player through these characteristics becomes able to build his knowledge [4]. So, some Objects can assume different roles in an adventure (game).

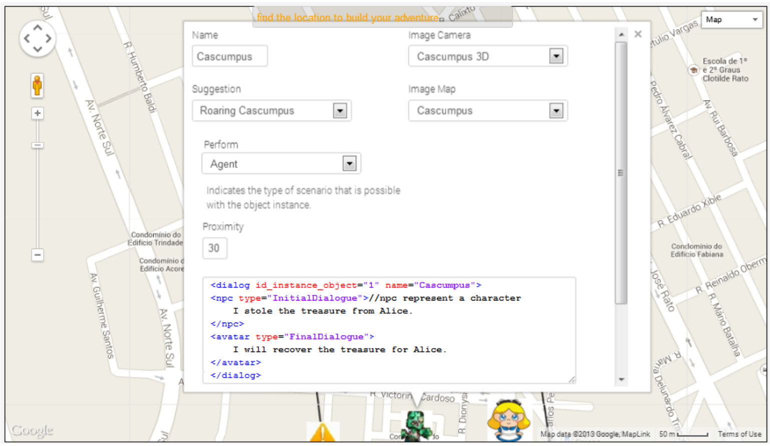


Fig. 2. Configuring object instance Cascumpus with Object Editor

The Object Instance represent (in the view of the player) by images in 2D or 3D, that allow the player can choose to see the adventure through Google Maps or the camera of your mobile device, characterizing a Mixed Reality (fields Image Camera and Image Map in Fig 2).

The Object instantiated (Cascumpus in the Fig 2) in the adventure map has name, geographic location (registered at the time the object is dropped on the map), proximity[2], visibility, perform[3], dialogues (conversation via *script*) and suggestion.

---

[2] Proximity: is the distance in meters (radius) that indicates when the player is close to a danger, for example Cascumpus. Or near the suggestion, for example, listen to the roar of Cascumpus.

[3] Performer: refers to the possibility of the object being controlled by a player with your mobile device.

As a suggestion also has an attribute that indicates the proximity distance required for the player to see the suggestion. The Objects Editor allows the author to create suggestions of three types: auditory, visual and textual.

Helping us to better understand the suggestion attribute, i.e.,: in ALDloc the monster Cascumpus has a suggestion for the player, that can be his Stench, as previously seen, or yet the player can hear the Cascumpus' Roar (as show in Fig 2).

A suggestion can be defined as elements of perception of the environment that influence the various decisions to which the player is subjecting. Therefore, the suggestions allow players remember past experiences to make more effective decisions.

5.2 Players Editor

The environment ARIS does not allow a player with your mobile device may assume roles of different characters. For example, in ALDloc a player can be the adventurer (default avatar) that needs to recover the Treasure and another player can assume the role of Cascumpus who wants to steal the Treasure of the adventurer (default avatar). Thus, the attribute Performer of Object Instance Cascumpus during the editing must be marked as Avatar (see Fig 2).

This flexibility is important for a decision-making environment, because it allows the player to interact with other players in cooperative or competitive way. In this context, the Players Editor allows players join a player to a character of a certain adventure to do the same role. The player (adventurer) register in a particular adventure can have your default avatar, or your avatar can be representing by an instance of a given Object, such as Cascumpus. Additionally, a user may have multiple avatars, so that it can have different characters on a single adventure. Importantly, each Object Instance in a particular adventure has a set of rules, attributing behavior to match.

5.3 Mission Editor

The Mission Editor allows the creation of missions[4] and conditions. Furthermore, the conditions serve to render tasks, i.e., to complete a mission from a set conditions must be fulfilled. The types of conditions are below:

<instanceObject>:<link>:<instanceObject>. The condition may be joining two Objects Instances. For instance, the player in adventure must get a Treasure and Box to keep the treasure. Then, the condition to check if the player kept the Treasure in a Box is (**Treasure combined Box**). The event to validate this condition occurs when the player attempts to combine (**combined** is a link) the two objects.

[4] Mission: when the author creator of Adventure (ALDloc) creates a mission, it is possible add the time necessary to complete the mission.

<avatar>:<link>:<instanceObject>. The condition may be joining an avatar of a player with an action in the object. For instance, in a given adventure the player's mission is to collect a potion to heal the burns. So, the condition can be: (**player picked potion**). The event to validate this condition occurs when the player gets (**picked** is a link) an instance object (i.e., Treasure) in the adventure (game).

<avatar>:<link>:<attributeObject:referenceInstanceObject>. This triple allows parts of a dialogue of a certain character in the adventure could compose a condition. For example, in the ALDloc for player activate the mission to recover the Treasure of Alice, the player can answer the follow dialogue option *"Yes. I can recover the Treasure",* for the Alice (character). It is important to emphasize that the player does not need to verbalize or to write his answers, because the dialog is built in the dialog *script* of the Object Instance (as shown in Fig 2). The condition for the situation described here is: (**avatar talked confirmation: Alice**). The attributeObject ***confirmation*** must have a reference to an Object Instance, in this example Alice. The attributeObject covers the follow situations:

Initial Dialogue. Occurs when a player greets a character, being characterized by dialogue like: "*Hello adventurer*", "*Good morning*", "*Good Afternoon*" or "*How can I help you young adventurer?*" The Initial Dialogue is representing by attributeObject ***InicitalDialogue:instanceObject***.

Final Dialogue: occurs when the player finishes the dialogue with a character, it is characterized by dialogue like: "*Thank you!*", "*Good luck!*" or "*Goodbye*". The Final Dialogue is represented by attributeObject ***FinalDialogue:instanceObject***.

Negation Dialogue: occurs when a player refuses to help a character in the adventure. For example, "*Do you could capture all the loose chickens?*" In this context Negation Dialogue is "*No! Sorry, I'm busy with another mission!*". The Negation Dialogue is represented by attributeObject ***negation:instanceObject***.

Confirmation Dialog: occurs when the player agrees to help a character. For example, "*Can you can recover the stolen treasure?*". In this context, getting the answer is "*Yes! Will be glad to help!*". The Confirmation Dialog is represented by attributeObject ***confirmation:instanceObject***.

Finally, it is important to emphasize that every Object Instance that has a dialogue, has attributeObject defined as **type="**attributeObject**"** in the dialogue script (as seen in Fig 2). When we create a condition, we must provide the name, mission and Story element. For example, a condition can be: Hunter needs to accept the request for help from the Alice character (see Fig 3). Still, it is possible to compose conditions of a mission using logical operators (i.e., AND or OR).

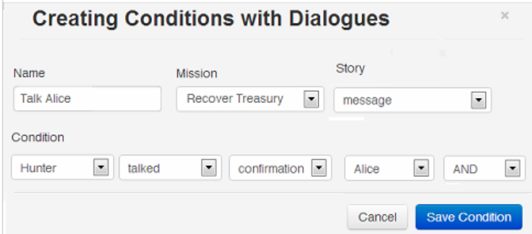


Fig. 3. Creating conditions with parts of dialogue of the Alice character

5.4 Movement's Editor

This Movement's Editor is purpose of allowing the adventure has characteristics of realism using mechanisms that allow Objects Instances (the monster Cascumpus, for example) have a dynamic behavior in the game.

This behavior attributes to Objects of an adventure aims to make the player continue to find information to make more efficient decisions in a given situation. So, avoiding the player to finds information by prematurely, preventing Overconfidence. For example, in ALDloc the Cascumpus may appear near the player several times. Then, the player will need to continue searching for information to develop new strategies of how to get to the Treasure.

For defining the dynamic behavior it is necessary to consider a set of information, such as:

Geographical Position: longitude, latitude and altitude which the object will take on events of the environment.

Behavior: the agent can be aggressive, passive, collaborative or competitive. For example, in the context of the game ALDloc, if the monster Cascumpus is aggressive it means that he will appear next to the player and steal your Treasure. Passive refers to behavior in which the Cascumpus will be nearby player, making it listen to its Roar (Perform) and make quick decisions under pressure. Collaborative means that Cascumpus can cooperate with another agent in the environment, stating that the player is going to one direction **x**. Competitor, is an agent that aims to hinder player adventure, making the same approach of monsters or obstacles in the game.

The agents perform their actions based on events that are generated by the player in the environment or other agents. The shares may be communication between agents, agents and player; actions on the environment (Adventure). For example, it is stealing an item from player, to swap an item or appear next to player (see Fig 4).

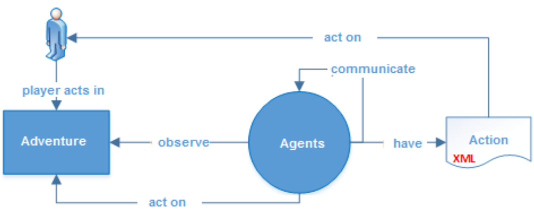


Fig. 4. Agent Architecture for the Movements Editor

5.5 Story Editor

The Story Editor arose from a need to help the player to reflect during the analysis of information originated through the decisions taken. In this context, the Story is taken as the transmission of events into images, words, audio and Object Instances. So, it is compounding the plot of the adventure.

This editor allows the author of an adventure to add images, words or objects in-game events by means of tasks and conditions, these are named as “plot elements”. So, when the player makes a decision and an event is triggering, the player has a vision of a plot element, so as to compose a narrative about the adventure.

The use of these elements allows the player to build knowledge and meaning about their actions in the game. So, it is helping the player has a better understanding of the information related to the results of his decisions. The Stories are relating conditions, so that, when a condition is triggering by a certain event a player receives feedback (a plot) about his decision.

5.6 Environment Development

The current version of the environment is available at link: http://192.241.218.138/autendlg/login/. The authoring environment is being developed with Django, a web framework that allows fast and friendly development.

Additionally, we have used the Bootstrap, it is a framework for application design with HTML 5. Hence, allowing the layout of the environment adapts according to the size of the device screen. Therefore, it is possible to use the environment in devices such as iPad or Smartphone with a large screen. Soon, the environment is developing with API version 3 of Google Maps, it based on HTML 5.

The environment has a top bar (in blue) which shows that the author is authenticating and the name of the adventure that is editing. On the left side has a menu with the editors of the environment. Below the menu, we have the Object Library that is instantiated in the adventure maps. Finally, in the central part of the environment the adventure map is presenting. For example, the ALDloc edited (see Fig 5).

Objects Instances linked by lines are Objects that have more than one geographic location, as mentioned in Section 5.1. Yet, if these Objects have the attribute Perform saved as Agent (as shown in Fig 1), it will have an autonomous behavior in the game by means of geographical positions recorded (Cascumpus, green monster has three geographical locations, as shown in Fig 5). This is configured by Movement´s Editor.

Already, the Object Instance "Exclamation" has the attribute Perform saved as Disabled. Thus, all geographical positions of the Instance Object Exclamation could be visible to the player in the device. This means, then be possible to limit the geographic area of the game.

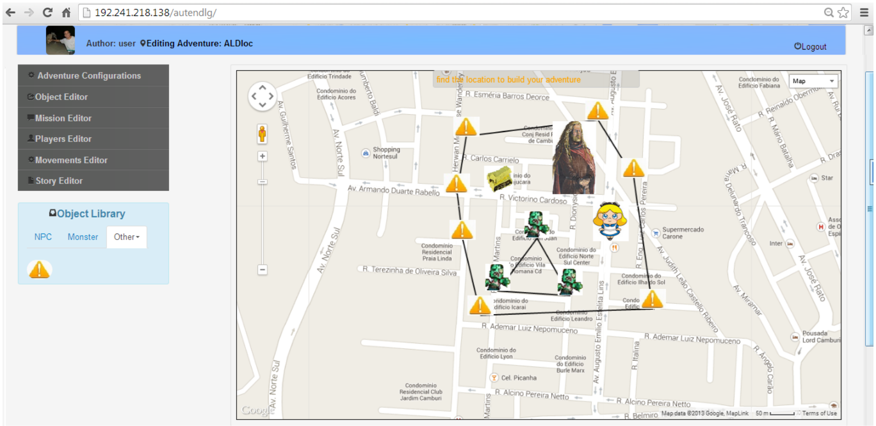


Fig. 5. The main screen of the authoring environment

6 Conclusions

Developing locative games in order to contribute to skills development decision-making stands out as a motivational factor for this research, as the combination of locative games and learn are a major challenge.

The results of the conceptions of the editors show that it is possible to build locative games like ALDloc. It is important emphasizing that editors allow the design of games with similar characteristics, however, with different contexts. So, allow the construction of games to develop decision-making skills in various contexts. An example, of these contexts is the civil construction, in situations where rational reasoning is not enough, and the experience and uncertainties (down material, for example) that scenarios should be consider. Yet, this work shows LDG concept as a new strand of locative games with features essential to the development of decision-making skills.

In the future, we plan to adapt the setting for service-oriented style, to integrate interoperable way with the executor of the games of Clic&Ação-ALRA in Android (Game Engine). Game engine is a core of functionalities that includes a graphical representation, touch detection, inference conditions, sound, script, agents, networking, and memory management. Finally, develop an extension of GDD (Game Design Document) to support the construction of LDG.

References

1. Aris.: ARIS – The Manual (January 2013), `http://three.umfglobal.org/resources/1857/ARIS-themanual.pdf`
2. Caird-Daley, A., Dawson, B., Ciereszko, R., Osborne, B., Parker, I.: Training Decision Making Using Serious Games: Requirements Analysis for Decision Making. In: Human Factors Integration Defence Technology Centre, UK (2009)

3. Choek, D.A., Goh, H.K., Liu, W., Farbiz, F., Fong, W.S., Teo, S.L., Li, Y., Yung, X.: Human Pacman: a mobile, wide-area entertainment system based on physical, social. Personal and Ubiquitous Computing 8, 71–81 (2004)
4. Gee, J.P.: What Video Games Have To Teach Us About Learning and Literacy. Palgrave Macmillan, New York (2003)
5. Gutirrez, L., Nikolaidis, I., Stroulia, E., Gouglas, S., Rockwekk, G., Boechler, P., Carbonaro, M., King, S.: fAR-PLAY: A framework to develop Augmented/Alternate Reality Games. In: IEEE PERCOM, pp. 21–25 (2011)
6. Klien, G., Orasanu, J., Calderwood, R., Zsambok, C.E.: Decision Making in Action: Models and Methods. Ablex Publishing Co., Norwood (1993)
7. March, J.G.: A Primer on Decision Making: How Decisions Happen, NY (1994)
8. Prasolova-Forlanf, E., Fominykh, M., Darisiro, R., Morch, A.I.: Training Cultural Awareness in Military Operation. In: The 46th HICSS, Wailea, HI, USA, January 7–10, pp. 903–912. IEEE (2013)
9. Santos, L.O., Rafalski, J.P., Menezes, C.S.: A Game Engine for Pedagogical Locative Adventures in Augmented Reality. In: SBIE, pp. 396–405 (2013)
10. Morati Junior, R.G., Berger, A.P., Tavares, O.L.: Menezes. C.: Alice in Labyrinth of Decisions: a game to exercise decision-making and planning. RENOTE, Journal New Technologies in Education 10, 60–69 (2012)

Leveraging Web Information for E-Learning

Jawad Berri

Information Systems Department,
College of Computer and Information Sciences,
King Saud University,
P.O. Box 51178, Riyadh 11543, Saudi Arabia
jberri@ksu.edu.sa

Abstract. Reusing web information to create instructional content is an important opportunity for e-learning. Web information is diverse and rich in content and can serve to a great deal the learning field if it is described and annotated semantically to fit educational purposes. Hypermedia web resources are particularly interesting as they promote the development of interactive and non-linear material for learning. Exploiting web resources to generate learning material for e-learning represents the key idea of the *Web of Learning* – a learning ecosystem build on the top of the existing web that makes use of existing web resources to fit learning purposes. The paper presents the Web of Learning architecture and a mobile learning system that creates instructional material on-demand. The system discovers and retrieves hypermedia resources based on semantic descriptions and reuse them to create learning objects that constitute the building blocks of learning material delivered to users.

Keywords: The web of learning, knowledge management, e-learning, learning object, web hypermedia resources.

1 Introduction

The potential of the web as the largest repository of information is investigated in many fields. The web can reveal many interesting facets if information is described and exploited efficiently [1]. This represents one of the semantic web initiative objectives of "having data on the web defined and linked in such a way that it can be used by machines for automation, integration and reuse across various applications" [2]. Information related fields which use the web as a source of data such as information retrieval, information extraction, data mining, and data science process data in order to establish correlations among data, discover recurrent patterns, or simply find web resources among the huge masses of information on the web to fit user needs [3, 4]. In the field of e-learning, instructors make use of web information to enrich and diversify the instruction material. They use Web 2.0 technologies such as social networks, blogging, wikis, and collaboration tools to deliver the material in a fashionable way which promotes cooperation and collaboration among learners. Web information is diverse and rich in content and can serve to a great deal the learning

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_46, © Springer International Publishing Switzerland 2014

field if it is described and annotated semantically to fit educational purposes. Hypermedia web resources are particularly interesting as they promote the development of interactive and non-linear material for learning. Although these resources have been authored to fit a specific purpose or objective, viewing these resources within a specific context represents a form of learning as knowledge is acquired and used in user's lives. For instance a video showing a virtual tour of the city of Istanbul in Turkey which is authored for tourism purposes can be used in a social studies class to explain the different civilizations that have influenced the city in the past centuries. Exploiting web resources to generate learning material for e-learning represents the key idea of the *Web of Learning* – a learning ecosystem build on the top of the existing web that makes use of existing web resources to fit learning purposes [5].

This paper presents the web of learning as an original approach to reuse web information to fit learning purposes. The Web of Learning architecture is then introduced and its components are explained. We present also a mobile learning system which has been implemented according to the Web of Learning principles. The system automatically creates instructional material on-demand for m-learning. It discovers and retrieves hypermedia resources based on semantic descriptions and reuse them to create Learning Objects (LOs) that constitute the building blocks of learning material delivered to users. The course of learning is fully controlled by the user who may navigate through a tree-like structure that organizes the progress of the learning process and allowing the user to deepen and diversify his knowledge. Moreover the system sustains durable learning as it manages the learning process through sessions allowing the user to interrupt and resume learning at any time.

The paper is organized as follows: next section presents a representative set of research works in the field of technology enhanced learning and e-learning. Section 3 exposes the main features of the Web of Learning as a new form of learning which relies on a reorganization of the web information. Section 4 is dedicated to the architecture of the Web of Learning. Section 5 presents a mobile learning system that illustrates a real-life implementation of the Web of Learning for users on the move. The last section is a conclusion of this work along with research directions aiming to enhance it in the future.

2 Related Work

The rich content of the web is attracting instructors who are keen to use it as an information source to support learning and acquisition of knowledge. There are mainly two approaches that exploit web information in the field of e-learning: the first approach consists in using the web as a large source of information. Web information is used as part of the learning material by embedding manually hyperlinks or web resources such as websites, videos, animations, presentations, etc. inside learning objects. The second approach involves Web 2.0 technologies and tools such as social networks, wikis, virtual environments, blogs, etc. to support learning. These new technologies offer collaborative learning environments which motivate learners who become more actively involved in the learning process. Some of these researches are presented in this section.

The special issue on current advances in learning technologies [6] presented a set of contributions that are representative on technologies that are used nowadays in e-learning. Technologies include cloud computing for education, advanced assessment technologies, digital gaming technology, intelligent learning systems, and other technologies used in technology enhanced learning. The diversity of technologies presented in this special issue shows that the e-learning community is keen to exploit the new technology innovations in order to design intuitive instructional material. The work in [7] presents APPRAISALWeb an online interactive platform that analyzes classifies and evaluates web resources that are used in language learning from a pedagogical point of view. Although this work does not show concretely how resources are evaluated pedagogically, it goes in line with our approach that consists in annotating semantically web resources as a prerequisite for their use in e-learning. In [8] Dunlap and Lowenthal present the use of Web 2.0 technologies such as blogging, social networking, document co-creation, and resource sharing to develop student skills such as autonomy, responsibility, and intentionality needed to become effective lifelong learners. Web 2.0 technologies promote engaging students into communities of practice and encourage them to participate in discourse and collaboration. This work is interesting as it illustrates how the new technologies have an impact on the motivation of learners. Brusilovsky and al. propose in [9] an architecture for e-learning systems that is based on reusing adaptive content services. The authors advocate the use of adaptive Web-based educational systems and ontology technologies to create Web-enhanced courses. These technologies help assisting developers and educators organizing, personalizing, and publishing learning content. This work is particularly interesting as it shows that the use of ontologies has an impact on the development and organization of e-learning systems.

The present paper proposes an alternative approach to exploit the web content. The approach consists in reusing hypermedia web resources to generate learning spaces on-demand. Web resources are annotated, retrieved and then packaged into learning objects which are then mapped into a learning web (LW) which is a course structure that organizes learning about a specific topic. Learners navigate into the learning web through learning sessions.

3 The Web of Learning

The Web of Learning is a redeployment of the Web where web information is reorganized semantically and reused to allow users to learn and build knowledge through a disciplined approach. Web users who usually seek information become learners as their queries will constitute a request for learning about a specific topic. The Web of Learning if compared with the web of information is characterized by the following main features:

1. While the search results given by web search engines are displayed as a list of ranked results which are generally restricted to one type of resources (either

websites, videos, images or others), the Web of Learning provides integrated instructional material including the best matches of many types of hypermedia resources serving the topic sought by the user.

2. Users in the web need most of the time to filter out irrelevant results and to refine their search by rewriting their queries and evaluating the results. In the Web of learning users are exposed to learning material that is designed as a course structure which includes a variety of aspects related to the topic of interest. The navigation inside the course structure promotes active and adaptive learning allowing users to construct their personal learning paths. They can go deeper in the course structure by following a branch to acquire specific knowledge or they can follow a normal course flow to acquire breadth knowledge.
3. Web search engines deliver generally punctual sessions to users with no possibility to record their search results. In the Web of Learning learners are able to learn according to their pace and constraints; they can resume learning at any time. Therefore, learning is managed through sessions to give users the opportunity to be exposed to their search results through a sequence of learning episodes.

The Web of Learning is a just-in-time learning which means that the production, delivery and user's knowledge acquisition is done when needed [10]. This paradigm promotes learning on demand allowing users to control the pace and direction of learning.

4 The Web of Learning Architecture

The Web of Learning promotes just in time and just enough type of learning. It is a just in time learning as users can request learning about a specific topic at any time. This paradigm of learning fits very well nomadic learners in context while facing special situations requiring knowledge, a skill or a know-how about a specific topic. It is a just enough type of learning as the leaning content is organized in a way that allows free navigation through the learning web which facilitates personalized learning experiences for every learner. The architecture of the web of learning has been designed into four modules namely the Interaction Manager, the Learning Engine, the Web Interface and the Knowledge Management modules (Fig. 1).

4.1 Interaction Manager

The Interaction Manager is responsible of many tasks: it handles all requests from users, senses and manages the context, adjusts the graphical user interface and manages users' learning sessions. The Context Manager creates an instance of the context for every new learning session. This instance includes data about the user, the device and the environment [11]. The context instance is updated periodically resulting in adjusting the learning courseware and the learning session content displayed to each user. User requests and queries submitted are handled by the Query Handler. It analyzes the user's query by extracting the objective and the topic. The query objective represents the type of information the user is looking for. For instance

"How queries" such as "*How to access the settings in Windows 8?*" are typical queries looking for the method, the procedure or the way something is done. Accordingly the query will be annotated as "Query_objective=method" and the topic sought by the user is "access the settings in Windows 8". The Learning Session Manager maintains the history of sessions and retrieves the last session upon the user's request while he logs in the system.

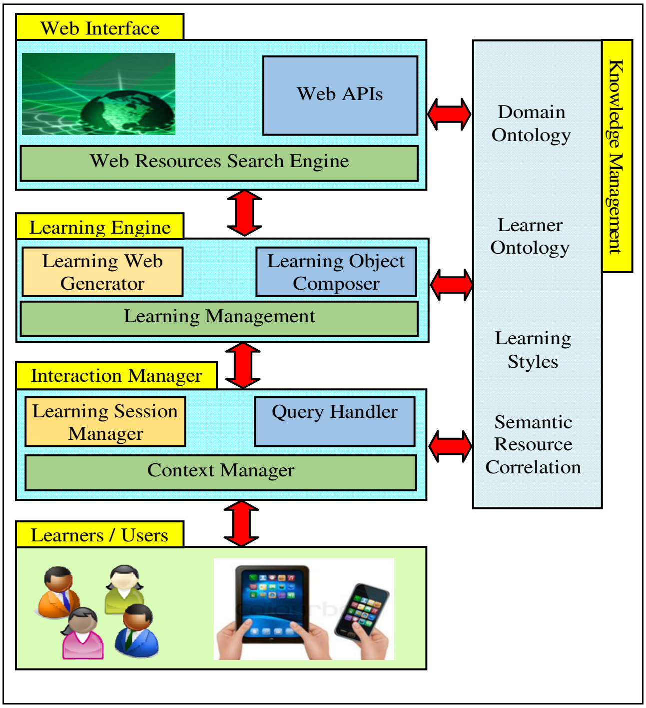


Fig. 1. The Web of Learning Architecture

4.2 Learning Engine

The Learning Engine is the module that creates the instructional material which includes learning objects and the learning web (Fig. 2). LOs are the building blocks of the course. They include a set of correlated hypermedia resources that present a specific concept of the course. LOs are generated by the Learning Object Composer; this module aggregates hypermedia resources gathered by the Web Resources Search Engine and packages them according to a set of preset layouts. LOs are then posted on the LW which represents the course's structure. The LW nodes are the LOs and the links are the possible pathways in which the learner may navigate in. The LW is created by the Learning Web Generator from ontologies or from information that structures a specific topic gathered from the web. The Learning Management module allows an efficient management of the learning process by mapping LOs into the LW and by managing the navigation of the learner through the LW [12].

4.3 Web Interface

The Web Interface is responsible of retrieving resources from the web. This module includes the Web Resources Search Engine and the Web APIs module. The Web Resources Search Engine is a Semantic Search Engine that retrieves web resources according to their semantic descriptions. Hypermedia resources that are used in our system needs to be analyzed in terms of their content, function and relationships with other resources. We have addressed automatic annotation of web resources in [5] as a requirement to share and reuse them in the Web of Learning. Web APIs (Application Program Interfaces) module is a generic interface that allows accessing hypermedia repositories through their API which offer methods to access and use publicly available web resources. APIs are not standard as they have been developed by different web hypermedia repositories providers. Hence, in order to access multiple repositories we need to implement a specific interface that invokes resources using the API of each service provider.

4.4 Knowledge Management

Knowledge Management includes all the knowledge components that are involved in organizing and feeding the different processes to create and dispense the learning material. It includes i) Domain Ontologies which are used to organize the courseware content, ii) the Learner Ontology which specifies the categories of learners and a specification of the information type they need, iii) Learning Styles that denote the differences in acquiring and processing information by learners in learning situations, and iv) Semantic Resource Correlation which is a module that computes semantic similarities between web resources in view of reusing them into the similar learning objects. Knowledge Management is connected to the web to retrieve available domain ontologies to generate the LW. In case an ontology is not available then the system retrieves information that allows the creation of an ad-hoc course structure. The system uses the table of content of Wikipedia (www.wikipedia.com) as the course structure.

5 Automatic Generation of Courseware for M-Learning

One of the implementations of the Web of Learning has been done as a Mobile Learning System [13]. The system has been designed to search and gather information from different websites and packages the resources in learning objects.

The first website used is Wikipedia for many reasons: i) It is the largest and most popular open encyclopedia and knowledge reference in the web. Wikipedia is a comprehensive source of knowledge that grows every day. This is very important for users' query coverage as queries are directed to Wikipedia in order to extract information to be imbedded in LOs, ii) It is developed as a wiki offering a standard template with a table of contents as hyperlinks to all the page sections. The table of contents is used to generate the LW in our m-learning system (Fig. 2), iii) Wikipedia provides an Application Program Interface (API) to access and manipulate the pages contents.

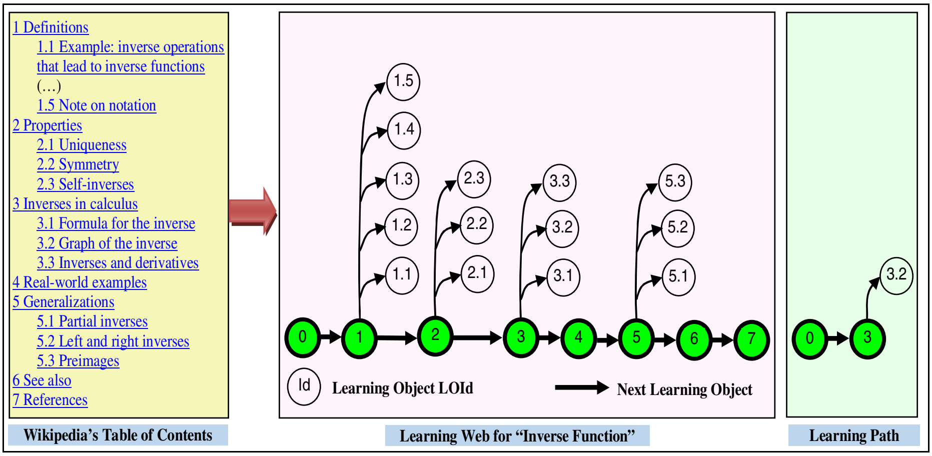


Fig. 2. Generating the Learning Web from Wikipedia table of contents for "Inverse Function"

Videos are gathered from Youtube (www.youtube.com) and images are searched using Yahoo images (www.yahoo.com) search engine. The two above web repositories offer APIs for accessing and retrieving content.

In order to illustrate how the system generates the learning material let's consider the following scenario: Saad is a college student who would like to learn how to invert functions through a short and concise hypermedia presentation on his tablet computer. Saad submits the following query *"How to find the invert of a function"* to the system which is first analyzed and forwarded to the three hypermedia web repositories Wikipedia, Yahoo Images and Youtube in order to search and retrieve relevant hypermedia web resources corresponding to the information requested. Then resources are annotated, retrieved and packaged into a LO as shown in Figure 3.

LO0 is the first generated object displayed to the user. LO0 includes the title on the top which corresponds to the phrase "Inverse Function" that has been extracted from the user's query. On the top left of LO0 a short text description gives the learner an introduction about the topic. This text description is the first paragraph extracted from Wikipedia webpage[1]. The image on the top right is retrieved from Yahoo images[2] based on a match between the image description and the query topic. The video about "finding the inverse function" in the center of the LO has been retrieved from Youtube[3]. The displayed video is ranked the second according to Youtube top list of results. The system selected the second result as it is the best match according to semantic annotations associated to each resource [5]. The learner can navigate inside the LW which is represented in the lower rectangles "Main Topics" and Sub Topics". All topics and subtopics are hyperlinks to all learning objects in the learning web of Figure 2 which can be generated in the same way as LO0. The user can navigate

---

[1] `http://en.wikipedia.org/wiki/Inverse_function` (accessed on 30/11/2013)
[2] `http://images.search.yahoo.com/images/` (accessed on 30/11/2013)
[3] `http://www.youtube.com/watch?v=sbvSBH2Mo20` (accessed on 30/11/2013)

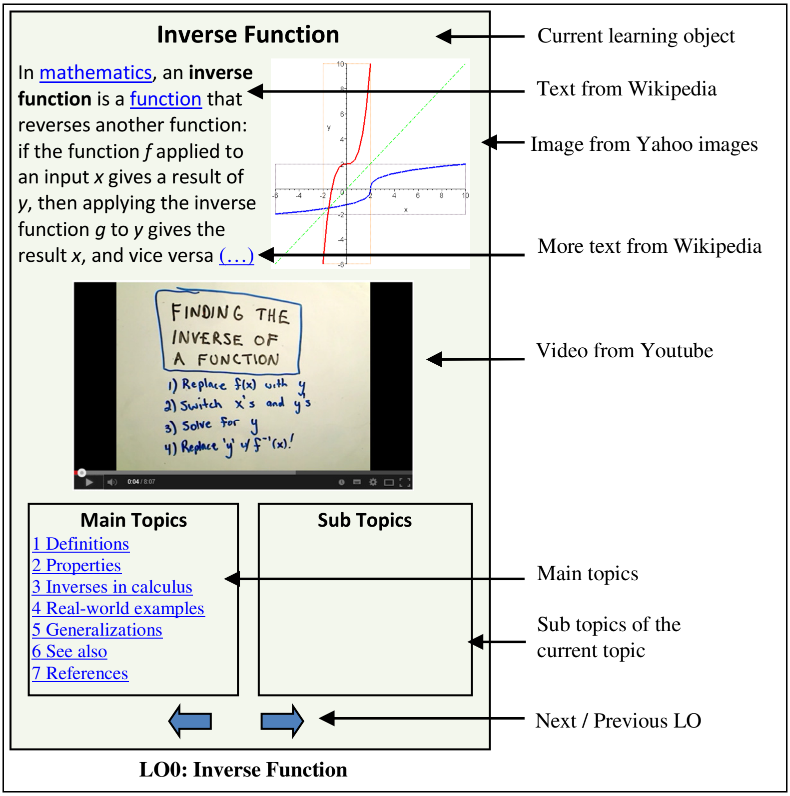


Fig. 3. Learning Object LO0: Inverse Function

freely in the LW; he can follow the top sequence of LOs in the learning web by using the "Next" button (represented as an arrow) or he can navigate randomly by clicking on any topic proposed in LW. Any topic requested by the user results in the generation of the LO corresponding to the topic in question. The two rectangles in the bottom of LO0 are i) "Main Topics" containing the topics that correspond to the main sections of Wikipedia Webpage, and ii) "Sub Topics" displaying any sub sections of the currently displayed topic. For instance, when Saad clicks on the topic: "3. Inverses in Calculus", LO3 is generated and displayed on the screen. In this case "Sub Topics" will include the list of sub topics 3.1, 3.2 and 3.3 (Fig. 2).

The Learning Web represents the course structure that includes all the possible personal pathways of potential learners. The learning path in Fig.2 shows that Saad has visited LO0 then he requested LO3. As he is interested to graph the invert of a function, he requested the subtopic LO3.2 "Graph of the inverse". As a result the learning path including learning objects LO0, LO3 and LO3.2 visited by Saad (Fig. 2 - right most graph) constitute his personal experience in learning about inverse functions.

6 Conclusion

In this paper we introduced the Web of Learning that is a learning ecosystem which reuses and organizes web information as learning material. The approach retrieves hypermedia web resources based on semantic descriptions and packages them into learning objects that are mapped into a learning web allowing free user navigation. The Web of Learning sustains the generation of on-demand interactive and non-linear learning material delivered to learners who control the pace and duration of learning sessions. The implemented system is designed to generate on-demand learning material gathered and packaged for mobile learners. Different topics have been tested to match the user queries and web hypermedia resources. Semantic matching revealed that it is more accurate then classical information retrieval for retrieving web resources to match the user needs [5].

In the future, it is planned to test the system using more hypermedia web repositories. This will result in larger web coverage and the possibility to reach more interesting hypermedia web resources in order to satisfy a diversity of learning topics. Additionally, we would like to design a variety of learning object layouts for the graphical user interface to adapt the presentation of the learning material to the user's profile and real needs.

Acknowledgments. This work was supported by the Research Center of the College of Computer and Information Sciences, King Saud University. The author is grateful for this support.

References

1. Berners-Lee, T.: Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web. Harper Collins, New York (2000)
2. Hendler, J., Berners-Lee, T., Miller, E.: Integrating Applications on the Semantic Web. Journal of the Institute of Electrical Engineers of Japan 122(10), 676–680 (2002)
3. Manning, C.D., Raghavan, P., Schutze, H.: Introduction to Information Retrieval. Cambridge University Press, Cambridge (2008)
4. Han, J., Kamber, M., Pei, J.: Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers Inc., San Francisco (2011)
5. Berri, J.: Web Resources Annotation for the Web of Learning. International Journal of Advanced Computer Science and Applications (IJACSA) 4(11), 132–139 (2013), `http://dx.doi.org/10.14569/IJACSA.2013.041118`
6. Sampson, D., Spector, J.M., Aedo-Cuevas, I., Chen, N.S.: Special Issue on Current Advances in Learning Technologies. Educational Technology & Society Journal 15(4), 1436–4522 (October 2012) ISSN: 1436-4522
7. Seiz-Ortiz, R., Gimeno-Sanz, A., Macario de Siqueira, J.: APPRAISALWEB: an online platform for the pedagogical evaluation of Web-based Language Learning Resources. Procedia Social and Behavioral Sciences 15, 667–671 (2011)

8. Dunlap, J.C., Lowenthal, P.R.: Learning, unlearning, and relearning: Using Web 2.0 technologies to support the development of lifelong learning skills. In: Magoulas, G.D. (ed.) E-Infrastructures and Technologies for Lifelong Learning: Next Generation Environments. IGI Global, Hershey (2011), doi:10.4018/978-1-61520-983-5
9. Brusilovsky, P., Wade, V.P., Conlan, O.: From Learning Objects to Adaptive Content Services for E-Learning, pp. 243–262. IGI Global (2008)
10. Maxim, B., Akingbehin, K., Modesitt, K.: Just-in-Time Learning in Software Engineering. Journal of Computers in Mathematics and Science Teaching 18(3), 287–301 (1999)
11. Berri, J., Benlamri, R.: Context-Aware Mobile Search Engine, Next Generation Search Engines: Advanced Models for Information Retrieval, pp. 371–385. IGI Global (2012)
12. Atif, Y., Benlamri, R., Berri, J.: Learning Objects Based Framework for Self-Adaptive Learning, Education and Information Technologies, vol. 8(4), pp. 345–368. Kluwer Academic Publishers (2003)
13. Alzaabi, M., Berri, J., Zemerly, M.J.: Web-based Architecture for Mobile Learning. International Journal of Infornomics 3(1), 213–222 (2010)

Cross-Artefacts for the Purpose of Education

Dulce Mota[1,2,3], Luis Paulo Reis[4,5], and Carlos Vaz de Carvalho[2,3]

[1] FEUP - Faculty of Engineering, University of Porto, Porto, Portugal
[2] ISEP - School of Engineering, Polytechnic of Porto, Porto, Portugal
[3] GILT – Graphics, Interaction and Learning Technologies, Porto, Portugal
[4] EEUM/DSI – School of Engineering, University of Minho,
Information Systems Department, Guimarães, Portugal
[5] LIACC – Artificial Intelligence and Computer Science Lab, Porto, Portugal
{mdm,cmc}@isep.ipp.pt, lpreis@dsi.uminho.pt

Abstract. The utility of computer-based teaching-learning systems is generally accepted but several relevant issues remain unsolved in the design of those systems, namely, how to adapt to a learner´s specific needs; how to plan corrective feedback; how to fit teaching-learning-assessment techniques to a specific educational context; how to choose the educational tools more appropriate to a teaching-learning-assessment method; how to choose a language to express a pedagogical model; how to adequate the teaching-learning-assessment activities deployment to a specific educational format (distance, face-to-face or blending learning). The aim of this paper is threefold: first, it surveys the most relevant computer-based teaching-learning systems since 1960. Second, it describes the learning design paradigm supported by specific modelling languages. Finally, it presents some reflections on educational material design, more specifically teaching-learning activities, that should be considered by teachers. Those considerations aim at bridging the gap between relevant theoretical aspects and the teachers' daily activities in the design of teaching-learning scenarios.

Keywords: computer-based teaching-learning systems, learning design scenarios, teaching-learning activities, educational tools, design tools.

1 Introduction

The study of instruction and learning suffered from the separation between learning, as the domain of psychologists and instruction, as the domain of educators [1]. But nowadays we are seeing a fruitful communication of principles and practices among teaching, learning and assessment domains.

The development in those domains along with innovation in information, computer technology and educational software design that took place in the second half of the last century reveals great changes in the way students learn and how they are taught too. The traditional educational settings are no longer the unique approach to learning. These patterns have emerged, to a large extent, because of the Internet and, particularly, the Web, whose potential has been widely studied [2]. Consequently,

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*, Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_47, © Springer International Publishing Switzerland 2014

planning learning scenarios is more and more a challenging task for teachers because many factors need to be combined and there is no one-size-fits-all formula for that. The educators' experience in teaching is certainly worthwhile but it is urgent to get them aware of the theoretical and scientific aspects in the instruction domain, such as the diversity of teaching-learning methods/techniques that can be put in practice with the help of recent developments in educational software design and digital technologies. Concerning these two issues, there are different architectures for different purposes, and the selection of those artefacts comprises also a very important topic for educators.

In this paper, we present in section 2 an overview of computer-based teaching-learning systems over six decades. In section 3, we describe the learning design paradigm supported by specific modelling languages. Some reflections on educational material design are presented in section 4. Finally, we conclude with the main ideas and contributions presented in this paper.

2 Educational Artefacts Overview

We divide teaching-learning artefacts in three threads: delivery, pedagogy and means. Concerning the delivery, we adopt the following approaches: face-to-face, distance education, online education and e-learning. In respect to the pedagogical thread, we include learning-teaching methods and techniques as well as strategies to learner's adaptation of contents and learning activities. The computer-based teaching-learning systems comprise the third thread which also includes software tools for designing teaching-learning-assessment activities. These issues are of real importance for educators to support the design of teaching-learning activities adapted to this knowledge age supported by digital technology.

2.1 The *Delivery* Thread

As mentioned earlier, the delivery thread is justified by the diversity of approaches to carry out a course, a module, a lesson or a simple teaching-learning activity. In addition to the traditional face-to-face approach other possibilities should be considered, namely, distance education, online education, and e-learning. In e-learning we include its subclasses, namely m-learning and b-learning.

Distance Education (DE) stretches back to 1880 and was characterized by the transmission of contents on paper. There was a complete separation between the teacher and the student without any type of interaction between them. From that time until today, several generations of DE can be mentioned, namely Correspondence, Broadcast radio & television, Open Universities, Teleconferencing and Internet/Web [3].

The e-learning and online education methodologies are also consequences of the developments in DE. The former is characterized by the use of the Internet; the course design can be a mixed of online and face-to-face classes, sustained in synchronous and asynchronous communications; and the interaction between learner and teacher is accomplished using all the kind of tools (chat, forum, email, videoconference, among others). It is worth to stress that e-learning is not e-publishing, as the focus should fall into the learning component rather than tools that support learning. The following definition wants to underline that idea: "e-Learning can cover a spectrum of activities from supporting learning, to blended learning (the combination of traditional and e-learning practices), to learning that is delivered entirely online. Whatever the technology, however, learning is the vital element." [4]. Nowadays, the e-learning domain covers scientific research in several aspects, namely accessibility, adaptability, interoperability and reusability. Concerning online education/learning, we resume this concept by presenting the following definition: "online learning describes education that occurs only through the Web, that is, it does not consist of any physical learning materials issued to students or actual face to face contact. Purely online learning is essentially the use of eLearning tools in a distance education mode using the Web as the sole medium for all student learning and contact." [5].

2.2 The *Pedagogy* Thread

This thread is based on the teaching-learning methods and techniques supported by the developments in two main scientific domains: learning and instructional theories, in one hand, and human intelligence theories in other hand. It is generally accepted that the bridge between the scientific research in those domains and the educator´s practice is more and more essential. In relation to learning theories, we underline the importance and consequences of Behaviourism, Cognitivism and Constructivism learning theories. More recently, the Connectivism theory [6], and the principles of Collaborative and Cooperative learning are in great focus as well. It is also worthwhile to include the most recent developments in Human Intelligence Theories once contemporary developments in this field have showed great contributions to explain how learning occurs. The second half of the last century was very fruitful bringing innovative ideas that departed from the traditional monolithic view of the structure of intelligence. Among such theories, we underline the Multiple Intelligences Theory [7], the Emotional Intelligence Theory [8] and the Triarchic Theory of Intelligence [9]. Their authors bring to the stage innovative reflections and explanations on why some students have great success in traditional settings and others do not.

Finally, we introduce the learning styles concept (also referred as cognitive or intellectual styles). Learning styles can be described as stable individual differences in perceiving, organizing, processing, and remembering information or as the people's preferred ways to process information and carry out tasks [10][11]. That concept has been used in the design of educational systems in order to promote adaptation to learners' needs and preferences [12] [13] [14].

All those aspects offer several insights on the design of learning scenarios, and more specifically teaching-learning activities, carried out by teachers. Moreover, the scientific research in the design of educational software can be "animated" by those achievements.

2.3 The *Means* Thread

This thread aims at describing briefly the significant developments in computer-based teaching-learning systems (hereinafter referred to educational systems) over the second half of the last century until now. To accomplish this objective, we select adaptable or adaptive features as a common denominator among educational software systems. The former is a more static view whereas the latter promotes a more dynamic one, that is, the educational system aims at adapting to some characteristics of the user or context during the user-system interaction.

The research in educational systems has inspired many researchers since 1960. Initial systems, dubbed under Computer-Aided Instruction (CAI) were developed for teaching in many varied domains as logic, axiomatic mathematics and foreign languages. EXCHECK, LOGIC, INTEGRATION and BIP-I [15] are some examples of this category. Such systems inspired a new generation of tutoring systems, Intelligent Computer-Aided Instruction (ICAI) in the 70s, later called Intelligent Tutoring Systems (ITS) in the 80s [16]. These educational systems stress the importance of reproducing in some way the human behaviour in tutoring tasks. In general, those systems have three main purposes: representation of the knowledge to be learned, inference of the learner's state of knowledge, and planning of instructional steps to be followed by the learner [17]. To accomplish such challenging goal, scientific knowledge from Cognitive Science, Instructional Theory and Artificial Intelligence domains occupied a central role.

At the same time, the pedagogical actor concept was introduced in tutoring systems design, embodying an intelligent, reactive, instructable, adaptive and cognitive agent [18]. This educational systems design approach aimed at supporting learning that is in line with the constructivist views of learning. That agent could represent either a virtual teacher (tutor, mentor, mediator or other possible roles) or virtual students (learning companion) [19].

Other step ahead was achieved at the time when educational systems first incorporated face-to-face interaction with intelligent, animated agents capable of creating learning and training environments more interactive to the student [20] [21] [22]. The humanlike behaviour of 2D and 3D characters was capable of promoting high levels of motivation in the students. In the research literature, such systems are referred to as Interactive or Intelligent Learning Environments (ILEs).

Other relevant category of educational systems that appeared in the mid-90s is named Adaptation Hypermedia Systems (AHSs). Its roots emerged from Hypermedia and User Modelling. Brusilovsky defined such systems as "[...] all hypertext and hypermedia systems which reflect some features of the user in the user model and apply this model to adapt various visible aspects of the system to the user" [23]. AHSs can perform both content adaptation and navigation adaptation in order to personalize the teaching-learning environment as much as possible to user needs and preferences. In this sense, many adaptation methods and techniques have been developed to accomplished such objective [24][25]. The educators can use available authoring tools [26] for constructing of their own AH applications, customizing them for their purposes.

Finally, we also want to highlight other type of educational system architecture that had a significant growth in the last decade, the Collaborative Teaching-Learning Design Systems. These systems stress the learning design as a way to bridge the gap between theoretical aspects of learning theories with the practice in educational settings [27]. As described in the next section, that conception allows putting in practice different instructional methods thereby demonstrating a different kind of adaptation. The main participants are users (students and teachers), teaching-learning activities and environments variables (the educational context, tools, files, among others).

In summary, the presented educational systems have many common concerns but there are some design decisions that make them unique. The learner autonomy during his/her learning trajectory is one of the features that vary considerably from system to system. Other feature is the system capacity to reuse learning materials, including learning activities. Those systems have quite different approaches in this respect. One more feature that can be pointed out is related to the adaptation of the educational system to the user' preferences, needs or knowledge. Finally, the ability of the system to adapt to different learning scenarios. We claim that this educational systems diversity has its own space in the educational context and the question is how to use and combine them to get the students focused on learning and motivated too.

3 Design of Learning Scenarios

The preparation of teaching and learning activities is a common and consuming task for teachers whatever the degree of education those activities are intended for. Moreover, it could be very complicated, and sometimes not feasible, if teachers want to put in practice different teaching and learning methods/techniques. Collaborative techniques, for example, are normally very difficult to implement in traditional learning settings.

Two important contributions can minimize those issues. On one hand, the information and communication technology tools, and on the other hand, the formalization of learning design supported by proper languages and specifications. The former embraces the use of tools to support the design and delivery of learning [28]. The latter is twofold: diversifying teaching and learning methods in order to respond to diverse student population as well as to provide different learning experiences, and reusing of learning scenarios in different teaching-learning contexts. This is the learning design frame that we are going to introduce now.

3.1 Learning Design

Learning design can be introduced as follows: "it aims at providing teachers with a framework capable to bridge the gap between rich, descriptive models and technologies, and the everyday practice and understanding of teachers" [27]. It has the potential to go beyond the learning content creation itself and proceed into the "process". In other words, learning design helps to bring to the stage the learning issue while the technological aspects come after. In turn, the contributions of the new technologies to the learning design needs also to be underlined once usual teaching-learning methods and pedagogies that were previously taken for granted can now be reconsidered [28].

Learning design aims at contributing to the reflexion of the mentioned aspects, and consequently, helping educators in the preparation of learning activities that respond to different teaching and learning approaches. Another definition of learning design is presented in [29]: "as the application of learning design knowledge when developing a concrete unit of learning (UOL), e.g. a course, a lesson, a curriculum, a learning event".

Learning design is used to describe a learning experience supported by tasks to which students should be engaged to. For example, students may be formed into groups and required to discuss the relations between two given topics; or they may be asked to gather some information about a given theme and write a report afterwards. The meaning of learning design knowledge is transmitted by a series of prescriptive rules with the following format: "if situation, then method". The left-hand side of the rule is the learning situation which accommodates the situational factors. The term situational factors are justified by the assumption that one method may behave best in a particular situation whereas another method may work best in a different one. Learning outcomes and learning conditions are the two subclasses of situational factors. The former is related to the level of effectiveness, efficiency, attractiveness and accessibility of the learning design method. The latter is related to the characteristics of some elements, such as the learning objective (knowledge, skill, attitude, competence), the learners (pre-knowledge, motivation, situational circumstances), the setting (individual and/or group work, work at school and/or work and/or home) and the media (bandwidth, synchronous/asynchronous, linear/ interactive, media types) [29].

To develop reusable, interoperable and adaptive learning designs there is the need for formal languages understood by machines. The IEEE Learning Object Metadata (LOM) and the IMS LD specification [30], for example, provide two relevant examples of communication languages. Their use promotes the development of new technical architectures in addition to enable the move towards a service-oriented approach for the development of software and true interoperability purpose [31]. As a more complex concept, the IMS LD specification describes several components, namely metadata, roles, plays, acts, environment, role-parts, sequence of activities and conditions. This conception is comparable to a theatrical play where actors (students, teachers among other participants/roles) perform as expected by the script (that is, tasks to be performed during a learning activity episode).

3.2 Learning Activities

In this research domain, the learning activity concept takes a central role. There is no single definition concerning this concept. We introduce the following definition as a good example to frame this issue: learning activity is "a specific interaction of learner(s) with other(s) using specific tools and resources, orientated towards specific outcomes" [28].

Four components are associated with a learning activity:

- Learner(s): This component combines identities (preferences, needs, motivations), competences (skills, knowledge, abilities) and roles (approaches and modes of participation).
- Learning environment: The focus is tools, resources and services.
- Learning outcomes: It comprises new knowledge, academic and social skills, and abilities.
- Other(s): Other people involved and the specific role they play in the interaction, e.g. support, mediate, challenge, tutor and guide.

The range of pedagogical approaches to the learning design process should be of a large spectrum using different perspectives on learning. The Associative, the Constructivist (individual and/or social focus) and Situated perspectives are considered a helpful support to create and sequence learning activities [32]. The learning outcomes to be achieved underpin the pedagogical decisions that educators should reflect carefully on.

A taxonomy of learning activities is introduced in [32] and forms an important basis to help teachers in the design of teaching-learning activities.

3.3 Design Tools Overview

Concerning the tools for supporting educators in learning activities design, we can witness a considerable work in this domain over the last two decades. Their design features vary considerably in some aspects, such us the degree of user's knowledge needed to deal with the metadata, the templates or wizards used to guide the teacher through a particular set of activities, the design process itself (textual or supported by graphic-based high-level facilities), and the output format chosen of the learning activity, which can restrict an openness to a broad student and educators' population. Some examples of those tools are: CopperAuthor [33], CoSMoS [34], MOT+LD [35], ASK-LDT [36], COLLAGE [37], Recourse [38], LAMS [39] and ACEM [40].

4 Reflections on the Design Process

The design of learning scenarios it is neither a linear nor an easy activity for teachers to accomplish. To turn that activity more efficient, we advocate a set of design principles that teachers should reflect on. Those principles could be divided into two main layers of issues, the educational context layer and the design context layer (see figure 1.). The educational context layer embraces three threads already explained in the last section along with two other components: teaching-learning activities and resources. The former is related to the sequence of tasks that students need to perform to get the teaching-learning activity completed. The latter comprises objects like files or services needed to support teaching and learning materials.

With regard to the design context the main idea is thinking about the tools that teachers should use to design the learning scenario. This decision has several consequences in the short and long terms. Some basic criteria to consider are the accessibility, the adaptability, the interoperability and the reusability of the tools and of the created learning scenarios. But also the capacity to enable the design of teaching-learning activities based on different pedagogies, including group work and collaborative learning, is relevant. To finalize this section, we want to underline the interoperability and the reusability features. It may be important that a teaching-learning activity could run in different players, and so the design tool should have the possibility to compile that activity into a proper language or specification. On its part, the reusable feature addresses a very recent issue as regards the searching facility of teaching-learning activities in repositories.

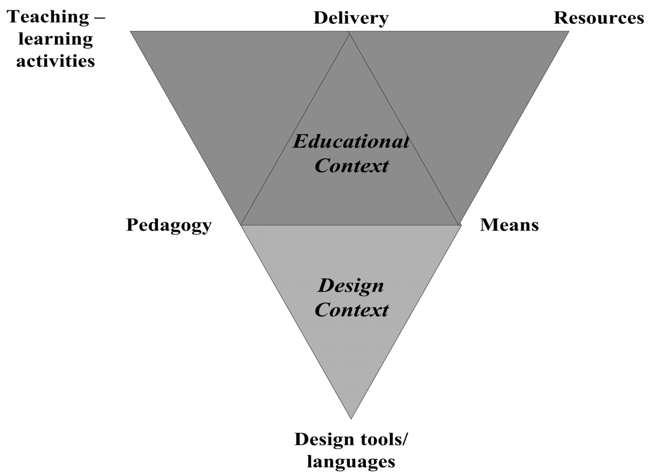


Fig. 1. Diagram describing the main principles in the design of teaching-learning scenarios

5 Conclusions

This paper has presented a survey of the most relevant computer-based teaching-learning systems since 1960 along with a description of the learning design paradigm supported by specific modelling languages. Concerning those systems, there are different architectures for different purposes, and the selection of those artefacts comprises also a very important topic for educators. In addition, we have described in more detail the learning design paradigm supported by specific modelling languages. Finally, we have also presented some reflections on educational material design and tools needed for that purpose. The choice of such tools by teachers has consequences in many aspects, namely accessibility, adaptability, interoperability, reusability and the implementation of different pedagogical models. All those considerations aim at bridging the gap between relevant theoretical aspects with teachers' daily activities in relation to the design of teaching-learning scenarios.

References

1. Schunk, D.H.: Learning Theories: An Educational Perpective. Prentice Hall (2003)
2. Collis, B.: Tecnología de la información en la educación superior: paradigmas emergentes. In: Revista de Universidad y Sociadad del Conocimiento, vol. 2(2) (2006) ISSN 1698-580X
3. Moore, M.G., Kearsley, G.: Distance Education. A systems view of online learning, 3[a] edn. Wadsworth Publishing Co. Inc. (2011) ISBN-13: 978-1111520991
4. JISC: Effective Practice with e-Learning. A good practice guide in designing for learning (2004), `http://www.jisc.ac.uk/elearning_pedagogy.html`
5. Nichols, M.: A theory for eLearning. Educational Technology & Society 6(2) (2003) ISSN 1436-4522

6. Siemens, G.: A learning theory for the digital age (2004), `http://www.elearnspace.org/Articles/connectivism.htm`
7. Gardner, H.: Frames of mind: The theory of multiple intelligences. Basic Books, New York (1983)
8. Goleman, D.: Emotional Intelligence. Bantam Books, New York (1995)
9. Sternberg, R.: Beyond IQ: A triachic theory of human intelligence. Cambridge University Press, New York (1985)
10. Sternberg, R.J., Grigorenko, E.L.: Are cognitive styles still in style? American Psychologist 52, 700–712 (1997)
11. Sternberg, R.J., Grigorenko, E.L., Li-Fang, Z.: A Capsule History of Theory and Research on Styles. In: Sternberg, R. (ed.) Perspectives on Thinking, Learning and Cognitive Styles, LEA London (2001) ISBN: 0-8058-3431-1
12. Stash, N., De Bra, P.: Incorporating cognitive styles in AHA! (The adaptive hypermedia architecture). In: Proceedings of the IASTED International Conference WEB-BASED EDUCATION, pp. 378–383 (2004)
13. Martins, C., Faria, L., Carrapatoso, E.: An Adaptive Educational System For Higher Education. In: Proceedings of the 14th EUNIS 2008 International Conference of European University Information Systems, Denmark, June 24-27 (2008)
14. Mota, D., Carvalho, C.V.: Personalized Learning Activities based on the Student's Intelligence skills. In: Proceedings of the 9th IEEE International Conference on Advanced Learning Technologies, ICALT 2009, Riga, Latvia, July 14-18 (2009)
15. Wenger, E.: Artificial Intelligence and Tutoring Systems. Morgan Kaufmann, Los Altos (1997)
16. Sleeman, D.H., Brown, J.S. (eds.): Intelligent Tutoring Systems. Academic Press, London (1982)
17. Self, J.: The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. International Journal of Artificial Intelligence in Education 10, 350–364 (1999)
18. Frasson, C., Mengelle, T., Aïmuer, E.: Using Pedagogical Agents in a multi-strategic Itelligent Tutoring System. In: Workshop on Pedagogical agents in AI-ED 1997, World Conference on Artificial Intelligence and Education, Japan (August 1997)
19. Chan, T.-W.: Learning companion systems, social learning systems, and the global social learning club. Journal of Artificial Intelligence in Education 7(2), 125–159 (1996)
20. Johnson, W.L., Rickel, J.W., Lester, J.C.: Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. International Journal of Artificial Intelligence in Education 11, 47–78 (2000)
21. Martinho, C., Paiva, A., Gomes, M.R.: Emotions for a Motion: Rapid Development of Believable Pathematic Agents in Intelligent Virtual Environments. Applied Artificial Intelligence 14(1), 33–68 (2000)
22. Giraffa, L.M.M., Viccari, R.M., Self, J.A.: Multi-Agent Based Pedagogical Games. In: Goettl, B.P., Halff, H.M., Redfield, C.L., Shute, V.J. (eds.) ITS 1998. LNCS, vol. 1452, p. 607. Springer, Heidelberg (1998)
23. Brusilovsky, P.: Methods and techniques of adaptive hypermedia. International Journal of User Modeling and User-Adapted Interaction 6(2-3), 87–129 (1996)
24. Brusilovsky, P., Millán, E.: User Models for Adaptive Hypermedia and Adaptive Educational Systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) Adaptive Web 2007. LNCS, vol. 4321, pp. 3–53. Springer, Heidelberg (2007)
25. Brusilovsky, P.: Adaptive Hypermedia. User Modeling and User-Adapted Interaction 11(1-2), 87–110 (2001)

26. De Bra, P., Calvi, L.: AHA! An open Adaptive Hypermedia Architecture. The New Review of Hypermedia and Multimedia 4, 115–139 (1998)
27. Cameron, L.: How learning design can illuminate teaching practice. In: The Future of Learning Design Conference (2010), http://ro.uow.edu.au/fld/09/Program/3
28. Beetham, H.: An approach to learning activity design. Rethinking Pedagogy for a Digital Age. In: Beethan, H., Sharpe, R. (eds.) Designing and Delivering e-Learning. Routledge, Taylor & Francis Group (2007) ISBN 978-0-415-40874-5
29. Koper, R.: An Introduction to Learning Design. In: Koper, R., Tattersall, C. (eds.) A Handbook on Modelling and Delivering Networked Education and Training, pp. 3–20. Springer, Heidelberg (2005) ISBN: 3-540-22814-4
30. IMS Global Learning Consortium, http://www.imsglobal.org/learningdesign
31. Conole, G.: Tools and resources to guide practice. Rethinking Pedagogy for a Digital Age. In: Beethan, H., Sharpe, R. (eds.) Designing and Delivering e-Learning. Routledge, Taylor & Francis Group (2007) ISBN 978-0-415-40874-5
32. Mayes, T., Freitas, S.: Learning and e-learning. In: Beetham, H., Sharpe, R. (eds.) The Role of Theory, Rethinking Pedagogy for a Digital Age- Designing and Delivering e-Learning, pp. 13–25. Routledge, Taylor and Francis Group (2007)
33. OUNL. CopperAuthor project website (2010), http://www.copperauthor.org/
34. Miao, Y., Hoeksema, K., Hoppe, H., Harrer, A.: CSCL scripts: Modelling features and potential use. In: Koschmann, T., Suthers, D., Chan, T.W. (eds.) Proceedings of Computer Supported Collaborative Learning 2005: The Next 10 Years!, pp. 423–432. Lawrence Erlbaum, Mahwah (2005)
35. Paquette, G., Teja, I., Léonard, M., Lundgren-Cayrol, K., Marino, O.: An instructional engineering method and tool for the design of Units of Learning. In: Learning Design: A Handbook on Modelling and Delivering Networked Education and Training, pp. 161–184. Springer, Heidelberg (2005) ISBN: 3-540-22814-4
36. Karampiperis, P., Sampson, D.: Designing learning services for open learning systems utilizing Learning Design. In: Uskov, V. (ed.) Proceedings of the 4th IASTED International Conference on Web-based Education, pp. 279–284. ACTA Press, Anaheim (2005)
37. Hernández-Leo, D., Villasclaras-Fernández, E.D., Asensio-Pérez, J.I., Dimitiadis, Y., Jorrín-Abellán, I.M., Ruiz-Requies, I., Rubia-Avi, B.: COLLAGE: A collaborative Learning Design editor based on patterns. Educational Technological & Society 9(1), 58–71 (2006)
38. RECOURSE TenCompetence Project, http://tencompetence-project.bolton.ac.uk/ldauthor
39. LAMS Foundation, Learning activity management system, http://www.lamsfoundation.org
40. Mota, D., Reis, L.P., Carvalho, C.V.: Fostering Collaborative Work between Educators in Higher Education. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics, SMC 2011, Anchorage, Alaska, EUA, October 9-12, pp. 1286–1291 (2011)

Erratum: Transformation of Coloured Petri Nets to UML 2 Diagrams

Ayman Yassin and Hoda Hassan

Computer Science, Arab Academy for Science and Technology, Cairo, Egypt
Ayman.Yassin@Gmail.com, huda.hassan@aast.edu

Á. Rocha et al. (eds.), *New Perspectives in Information Systems and Technologies, Volume 2*,
Advances in Intelligent Systems and Computing 276,
DOI: 10.1007/978-3-319-05948-8_13, © Springer International Publishing Switzerland 2014

DOI 10.1007/978-3-319-05948-8_48

In the original version, the first author wants to change his email id. It should read as:

Ayman.yassin81@gmail.com

The original online version for this chapter can be found at
http://dx.doi.org/10.1007/978-3-319-05948-8_13

Author Index