Modeling Organizational Information System Architecture Using "Complex Networks" Concepts

José L.R. Sousa
Sistemas de Informação
IBMC
Porto, Portugal
jsousa@ibmc.up.pt

Ricardo J. Machado
Centro ALGORITMI
Universidade do Minho
Guimarães, Portugal
rmac@dsi.uminho.pt

J.F.F. Mendes
Group of Complex Systems & Random Networks
Aveiro, Portugal
jfmendes@ua.pt

***Abstract* — Organizations live in a world where interdependence, self-organization and emergence are factors for agility, adaptability and flexibility plunged into networks. Software-based information systems go into a service oriented architecture direction and the same goes to Infrastructures where services are become structures available in networks. Inspired into empirical studies of networked systems such as Internet, social networks, and biological networks, researchers have in recent years developed a variety of techniques and models to help us structurally understand or predict the behavior of these systems. Those findings are characterized by been supported on the "complex networks" concepts. On this PhD research we present the use of the concepts of complex networks from physics to develop organizational information system architectural models, as requirements modeling technique. The research is about the structure and function of networks and its use for modeling organizational information systems architectures by using a combination of empirical methods, analysis, and computer simulations.**

Keywords: software requirements, computer science related discipline: information management, management related discipline: information systems management

I. Introduction

Organizations live in a world where interdependence, self-organization and emergence are agility, adaptability and flexibility factors plunged into networks. It is a networked composed world in the design of collaborative-networked organizations. This networked structuration comes to the composition of complex systems, from cells, to society and enterprises (associations of individuals, technology and products). In those complex systems, characteristics of emergence, order and self-organization [1], develop a set of network interdependent actions not visible in the individual parts. This is what complex systems are about and networks concepts are becoming a common approach to describe and quantify these complex systems structures. A fundamental concern is to know the anatomy of theses structures in a relation that structure always affects function [2].

Organizational information system concentrates a set o elements that due to its role and nature are becoming similar to energy and raw materials, this is, are fundamental elements for organizations needs and successful existence. Developed in order to support the organizations processes, which are the center of organization efficiency and effectiveness, but without in-built value [3]. Organizational information system integrate commodities like computation power, whose current availability [4], is no longer a differentiation. In presence of this elements and in consequence, information systems are each more seen as a commodity [5] where organizational information systems are consumed by the idea of not being aligned with the enterprise information needs [7].

The realm of organizational information systems is the confluence between context, business and software-based technology and infrastructures. Organizational information systems and behavior are not dichotomous but inseparable [6] building a interdependent techno-social system. The use of complex network characteristics to model the information flow in organizational information systems is a tentative to better understand their techno-social nature and, thus, learn how to design better software-based organizational information systems architectures.

Present research is about the structure and function of complex networks and its use for modeling organizational information systems, studied using a combination of empirical methods, analysis, and computer simulations. The goal is to find the complex network structure for real organizational information systems architectures and produce models that can be analyzed and studied to better understand the information flow in the organizational techno-social system.

In this paper, a description of the state-of-the-art related with the subject of this research is presented in Section 2. Section 3 describes in detail the research objectives and the methodological approach. In Section 4, past work and preliminary results, already done in the context of this research, are briefly described. Section 5 presents future work and expected results, for the next 2 years of research. Finally, in Section 6 some conclusions are presented.

II. STATE-OF-THE-ART

Organizational information systems, are consumed by the idea of not being aligned with the enterprise information needs [7], whose existence and development is sustained by integrating a set of commodities. Organizational information systems should be planed, design developed and managed according to the organization strategy [8]. This approach makes information systems clearly responsive. There is a real need for information systems, with ability to support emergence, self-organization that are present and can support the organizations evolution [9]. Organizations are real world entities with dynamic elements, users, in space and time.

978-0-7695-4777-0/12 $26.00 © 2012 IEEE
DOI 10.1109/QUATIC.2012.40

Information systems must able to support the evolution of the interdependence between the techno and social systems. The construction of this techno-social system can be modeled by complex interactions [9] that develops co-evolution trough a flexible, adaptable and agile information systems architecture structure.

A. Information systems

Defined by [10] as an "*organizational system that consists of technical, organizational and semiotic elements which are all re-organized and expanded during ISD (information systems development) to serve an organizational purpose*". According to [11], "*information systems is what emerges, from the usage and adaptation of the IT (information technology) and the formal and informal processes of all of its users*" .

Information systems must be constantly adapting to needs as the users change its use and the IT is updated or extended [11]. Information systems are also seen as communication system used to support a given human activity system, processes [12]. Information systems are a natural consequence of the need for humans to communicate and coordinate their activities [13][14]. An organizational information system, on this research proposal, is the complex network of interactions of software-based information systems (that are combinations of computation, communication, technology and processes) developed to fulfill enterprise goals and integrated in a certain context. This software-based information system is seen as SaaS (Software as a Service) anchored in a IaaS (Infrastructure as a Service) where context is the set of elements (Users) that interact with information technology. This complex networks approach to model information systems develops a architecture umbrella that is a tentative to understand the intrinsic nature of information flow and allows the design of better software-based information systems that expectably fully integrates emergence and self-organizing behaviors.

B. Commodities

Commodities lead to the concept that itself an information system does not add value. From technology [5], communications and computation [4, 15], information technology solutions are seen as a commodity. In this perspective information technology will be available at the enterprise ecosystem and can be easily plugged in and out, like energy. As an example, when a certain active entity integrates the ecosystem gets energy and everything else, including information technology. In this sense, service science is clearly developing this commodities deployment. So, this development of reality is demanding for a change that is represented in a particular part of any system, but emerges or results from a self-organization and interdependent evolution. It needs a view to develop information system architecture able to accommodate this new dynamics.

C. Complex adaptive systems (CAS)

The paradigm used to model and fundament organizations development has been changing. This change can be found in a fundamental core of articles and books that deal with this new organization dynamics [8, 18-20].

In this context, information systems architecture as a fundamental interdependent structure, develop trough an reductionist vision, will never have relevance due to its vision of parts, instead of the hole. Complexity theory and CAS can represent a response to the paradigm shift in order to address this new dynamics [21-23].

CAS integrates the concept of emergence from which adaptability and evolution arise as a result. For modeling this, complex networks are used to model self-organization, preferred attachment and fitness. Organization development is supported, by the ability to use information flow as a source for unique results, when facing change and competition along space and time [16, 17].

Complex adaptive systems are systems with great number of components, sometimes called agents that interact, adapt and learn. Many contemporary problems are under the complex adaptive system's theory [22-25]. In this systems emergence occurs near or in the limit of thermodynamic equilibrium. Such systems are common on the physical world and have "*emergent*" properties that result from interactions and are global or collective.

Emergence is founded in the existence of a global behavior that is new when related to all parts that compose the system, something not understood with the reductionist vision of the information system architecture.

Information flow then defines emergent patterns that can lead to adaptation, a familiar form in the biological process. Reorganizing genetically material, through which organism evolves to survive in environments that confront them. This process allows the modulation of non-linearity that comes from complex interactions [26, 27].

D. Complex networks

There are often cited examples of complexity, such as the Internet, WWW, immune system, ant colonies, economic markets or human social networks. Despite this fact, there is no central definition for complex systems, informally seen as a large network of relatively simple components with no central control, in which emergent complex behavior is exhibited [28]. This behavior is hard to define, and roughly, emergence refers to the fact of systems global behavior is not only complex but arises from the collective actions of simple components to which the notion of non-linearity is important: the whole is more than the sum of the parts [28].

Networks are everywhere, from brain, to society passing by organizations. Using Karl Popper [29] approach, in higher degree, this pervasive presence of networks is a construction of human mind. Internet and WWW are the most impressive creation in the information system domain and probably the most moving creation of our civilization [30]. It is possible to imagine the past and the future without them; but for the e-generation it is not [8]. It is an element present in the daily-life and from which we know little, beginning on their complex organization structure or global topology. The understanding of the Internet and WWW inherent problems is not a topic for social sciences, computer and applied mathematics but rather of non-equilibrium statistical physics

[30-32]. For properties observation of the network data is the starting point [33, 34] and the same can be argued for organizational information systems. The study of how information flows and its support under software-based information systems interactions should be done with the same topological approach, which regards the big data that is stored.

Network's structure's study started in the mathematical study of graph theory [30]. In the beginning this theory seized to Poisson distributions, resulting from simple random graphs. Moreover, by definition, random graphs in graph theory, are graphs with Poisson distribution of connections [30]. By definition "*a network is simply a collection of nodes (vertices and links (edges) between nodes. The links can be directed or undirected, and weight or outweighed"* [28].

As first stage, all networks seemed random but, along the development analysis, some different and fundamental key characteristics were found. Networks are characterized by the way in which they are created, resulting into constructs such as, degree of distribution, average path length between node's pairs, clustering degree and communities [28, 31].Barabási and Albert (echoing the earlier work by Price and others) conclude that their simple "*growth with a preferential attachment*" mechanism, is what drives the evolution of real-world networks [28]. Network theory has been used to characterize a different set of systems. This use is making network theory a strong tool for using when emergence, self-organization, dynamic and co-evolution are characteristics to be analyzed in the systems.

E. Relevant information system architecture planning models

Due to the increasing in size implementations of organizational information systems, logical models (*or architecture*) for defining and controlling the interfaces and integration of all the system components were developed. Following, two of the most relevant are presented and a discussion is made about its bottlenecks in supporting organization techno-social systems information flow.

ZACHMAN Framework: It provides a view of the subjects and models needed for complete developing or/and documenting of organization architecture [35-37]. This framework provides a basic structure that supports the organization, access, integration, development, management and changing with a set of architectural representations of organization's information system. It uses a matrix structure of 30 cells and five perspectives of the overall architecture with six classifications of the various artifacts of the architecture as well as flow diagrams. For each cell of the matrix the documentation type is suggested, using ER technique for modeling the data description or using functional flow diagrams for modeling the process description [37]. In this framework, an organization has a whole range of diagrams and documents representing different aspects or viewpoints that can be developed. In the extended framework for information systems architecture there is a meta-model for cell data, and a classification of data, process or network is made. It has no specific associated methodologies and only a set of major principles and rules exist, working as a guide. Nothing is said about processes development for viewpoints or the associated order. It represents information technology and not information systems architecture interdependence in the organizational techno-social system.

TOGAF: The Open Group Architecture Framework (commonly known as TOGAF) is an industry standard architecture framework that may be used freely by any organization wishing to develop enterprise architecture descriptions for the use within that organization. It is defined as a comprehensive architecture framework and methodology, which enables the design, evaluation and implementation of the "right" architecture for an enterprise [38], supported by a set of well-defined tools. It is composed of three fundamental parts [39]: the ADM, the Enterprise continuum, and the Resource base. ADM (Architecture Development Method) forms de core structure for TOGAF, being able to detail procedural models in order to develop descriptions of enterprise architectures. It describes the different type of inputs and outputs but does not show guidelines; ER – entity relation - is used as a formalization model [39]. Design for development and not for exploitation of techno-social systems interdependent interactions in the road to co-evolution.

III. RESEARCH OBJECTIVES AND METHODOLOGICAL APPROACH

A. Research Goals

The present research is about the discovery of big data information flow structured under complex networks, and its use for modeling organization information systems architectures, using a combination of empirical methods, mathematical analysis, and computer simulations. These computer simulations will address the visual modeling of the network structure. From its review and results constructs of the complex network theory such as clustering, path-length, degree and communities, develop the ability to model emergence, self-organization, evolution and dynamic through space and time of the organization techno-social system

Regarding this, the main goals for PhD work are:

- Define an approach to support the adoption of the complex network meta-model to analyze models of already existent organizational information systems, based on information technology big data information flow.
- Adopt those models for monitoring the dynamic execution of existent organizational information technology systems.
- Based on the monitoring results, analyze the corresponding characteristics of the organizational information system architecture.

B. Research Approach

In order to validate this approach, is adopted a positivistic research method trough a quantitative data collection. Later on, an observing interpretation will be needed and a qualitative vision will be adopted, supported in the use of and abductive inference.. All the development phases will be based on action research theory [40]. The use of action research theory comes from its use on information systems research, when there is the need to determine the complex or diverse nature of information systems. The complexity support and diversity is clear in the modeling goals of our research, so the choice for action research is adequate. Traditional approach to the definition of the action research methodology is a spiral of interactions through time. At each cycle a more close set of interaction can be implemented [40]. For the present research proposal a more simplified version of action research will be used: the demonstration cases. This approach adopts a single cycle for data collection in one selected organization.

Information system architecture is a world of big data logs and traceable information flow interdependence in the techno-social organization system. There are many tools for collecting that information. In this research, standard tools will be used and it is expectable that a plethora of models will emerge from big data [41] evaluation using quantitative modeling within complex networks constructs. A qualitative approach will also be present in our efforts for the interpretation of the results obtained from the positivist construction of models using an abductive inference for its predictive fitness. Starting from the presented concepts in this paper, a more profound literature review will be done, exploring fundamental concepts from theoretical physics that will be used and also trough the evaluation of users behavior relating to technology and predictive fitness. This combination of research methods is a great enrichment for our research approach. The global research process will be centered in three main phases, described in Fig. 1.

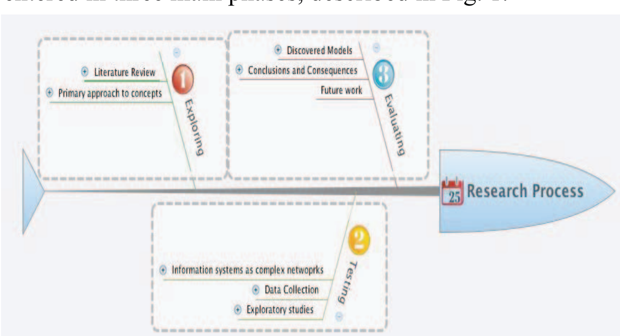

Figure 1 - Overview of the research process.

This research process starts with a broad question: *How already existent organizational information systems can be characterized adopting the "complex networks" concepts?*

Along our PhD work, three organizations will be involved for performing data collection and analysis supporting the construction of *"complex networks"* models. Two of these organizations are from two relevant industrial sectors in Portugal: automotive and textile. This phase will present modeling results of the application of complex network concepts to the organizational information systems using demonstration cases.

IV. PAST WORK AND PRELIMINARY RESULTS

This PhD work takes place within the Software Engineering and Management Group (SEMAG) from the ALGORITMI Research Centre at the University of Minho. SEMAG research group is devoted to study the development process of software-based information systems and related methodologies, focusing on both the engineering and management aspects. We will adopt a three-layer techno-social approach to regard into the commoditization process of enterprise information systems architecture in a foundation of service science (see Fig. 2): [42](1) the user level; (2) the SaaS (software as a service) level; (3) the IaaS (infrastructure as a service). We will manage a set of experiences (real big data collection and analysis) in order to produce the complex network models for each of the three levels.

A first data collection experiment was put in place at IBMC addressing the SaaS level- the level were processes are translated into software based packages - using a small link with a volume 40 to 50 connected users and *Allot NetEnforcer Series AC-500* equipment (with the cooperation of Palo Alto Networks[1]); a commodity solution that can collect data flows. The data was collected during a three-week period and then exported to CSV files through the use of *Netxplorer*, to be processed by external software for complex networks evaluation..This resulted in the construction of the adjacency list of the networks [33]. The software used for complex networks structure discovery and constructs evaluation was *Gephi.*

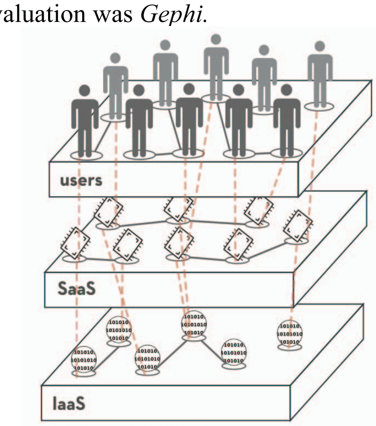

Figure 2 – 3-layer information systems modeling architecture.

The experienced was conduced with the guarantee of total confidentiality, since the data flow is only related within the IP address and additional information is needed to correlate. The architecture of data-collecting infrastructure is presented at Fig. 3. It was performed in a five-flours building in a LAN with a clear DMZ and core switching linking all the distributed connections. The data collection allowed us to discover initial structures of the information flow that are the core element of the "complex network" model.

[1] www.paloaltonetworks.com

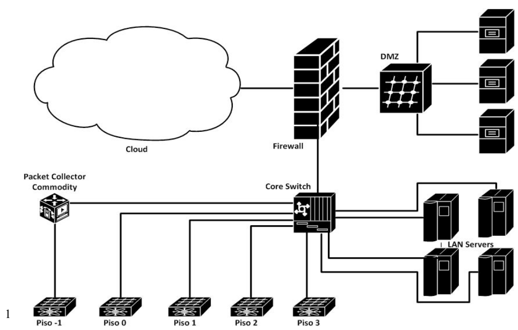

Figure 3 - Data collecting ecosystem.

Fig. 4 presents the "complex network" model obtained and Table I shows numerical characteristics of the model. Clustering coefficient defines the complex networks structure as having small-world properties. We are now studying these values, its relation to the SaaS for which we have collected data, and what they can mean in that context.

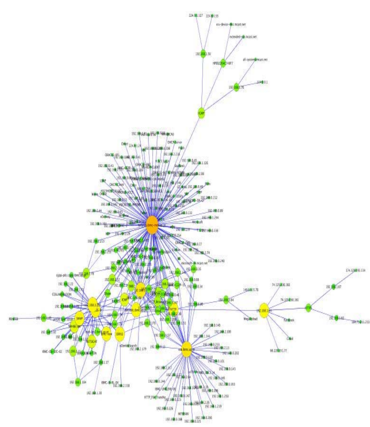

Figure 4 **- Complex network structure at the SaaS level.**

TABLE I. VALUES OF COMPLEX NETWORK AT THE SAAS LEVEL

Element	Value
Average degree *(k)*	2.869
Average shortest path *(l)*	3.074
Clustering *(C)*	0.004

Another experience was conducted within the ISOFIN project to assess the characteristics of an information system architecture that is being designed. The ISOFIN project [43] aims to provide a set of functionalities based on the cloud paradigm as defined by NIST [44] and enacting the coordination of independent services relying on private clouds in a coordinating public-cloud application (the ISOFIN Platform). The resulting ISOFIN platform will allow the semantic and application interoperability between enrolled financial institutions (Banks, Insurance Companies and others). In the presented real industrial case, the process-level 4SRS method [45, 46] is used to create the necessary context to elicit the requirements for designing an architecture capable to be implemented in the three typical cloud-layers: IaaS, PaaS (platform as a service), and SaaS.

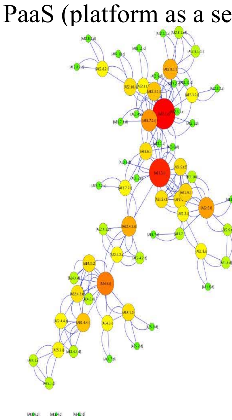

Figure 7 **- Complex network of process-level logical architecture of ISOFIN in a circular layout with expansion of packages and relevant edges**

Fig. 4 and Table I present the results from the IBMC experience; they allowed discovering the existence of "small-world" network properties in the existing information system. Fig. 5 and Table II present the results from the ISOFIN experience; they allowed discovering the existence of communities that are not equal to the packages defined in the designed architecture for the ISOFIN information system. A study of the construct communities should also be addressed in the future work of this PhD.

TABLE II. VALUES OF COMPLEX NETWORK AT THE SAAS LEVEL

Element	Value
Average degree (k)	2.735
Average shortest path (l)	4.150
Clustering (C)	0.253

V. FUTURE WORK AND EXPECTED RESULTS

For the next two academic years (2012/13 and 2013/2014), we will address the other two levels (users and IaaS). With the lessons learned from these three kinds of experiments we expect to define the main elements for the complex adaptive information systems architecture.

VI. CONCLUSIONS

Organizational information systems are today faced with a need for management of information flow through space and time in order to support organization information needs. Inspired on the work been done on the definition of WWW and social relations, this PhD proposal is presenting a new approach to the modeling of organizational information systems architectures, by using the "complex networks" concepts. It adopts the physics concepts of complex networks and proposes a research agenda for modeling organizational information systems architectures as a first step to the engineering of information systems. Although those concepts have already been used in modeling the WWW, power grids or air traffic systems, they have never been tested in the information systems domain. The collaboration of an enterprise like Palo Alto Networks, making available the last version of a context firewall in order to be used for data collection, makes a big impact on what can be collected and on what models that can be produced.

A special attention was paid to the choice of the selected organizations for data collection, trying to address relevant domains in order to give the vision for the fitness of the "complex networks" concept in different context and behaviors. This approach addresses the use of physics concepts that once more shows that information systems research benefit from the knowledge any area domain. Its different way of seeing things can be a central research for the leverage of organizational information systems architecture to the center of co-evolution of organization socio-technical systems trough the ability to exploit architecture for agility, flexibility and adaptability

REFERENCES

[1] Barabási, "The Architecture of Complexity," *Control Systems Magazine, IEEE,* vol. 27, pp. 33-42, 2007.

[2] S. H. Strogatz, "Exploring complex networks," *Nature,* 2001.

[3] T. H. Davenport, "The coming commoditization of processes," *Harvard Business Review,* 2005.

[4]N. Carr, "The end of corporate computing," *MIT Sloan Management Review,* vol. 46, pp. 67-73, 2005.

[5] N. Carr, "IT doesn't matter," *IEEE Engineering Management Review,* 2004.

[6] A. R. Hevner, S. T. March, and J. Park, "Design science in information systems research", *MIS Quarterly,* 2004.

[7] Y. E. Chan, "Why haven't we mastered alignment? The importance of the informal organization ," *MIS Quarterly Executive,* 2002.

[8] A. W. Don Tapscott, *Wikinomics - how mass collaboration changes everything*. New York: Portfolio, 2006.

[9] Y. E. Chan and R. Sabherwal, "Antecedents and outcomes of strategic IS alignment: An empirical investigation," *IEEE Transactions ,* 2006.

[10] K. Lyytinen and M. Newman, "Punctuated equilibrium, process models and information system development and change: towards a socio-technical process analysis," *Sprouts,,* 2006.

[11] R. J. Paul, "Challenges to information systems: time to change," *European Journal of Information Systems,* 2007.

[12] P. Beynon-Davies, "Informatics and the Inca," *International Journal of Information Management,* 2007.

[13] P. Beynon-Davies, "Neolithic informatics: The nature of information," *International Journal of Information Management,* 2009.

[14] S. Alter, "Defining information systems as work systems: implications for the IS field," *European Journal of Information Systems,* 2008.

[15] M. P. Papazoglou, "Service-oriented computing: Concepts, characteristics and directions," *Fourth International Conference on Web Information Systems Engineering (WISE'03),* 2003.

[16] E. Bonabeau and C. Meyer, "Swarm intelligence: A whole new way to think about business," *Harvard Business Review,* 2001.

[17] A. Desai, "Adaptive complex enterprises, *Communications of the ACM,* 2005.

[18] G. Hamel, "O Futuro da Gestão," 2008.

[19] I. Nonaka and G. v. Krogh, "Perspective---Tacit Knowledge and Knowledge Conversion: Controversy and Advancement in Organizational Knowledge Creation Theory," *Organization science*, 2009.

[20] J. Surowiecki, *A sabedoria das Multidões Como Inteligencia colectiva transforma a economia e a sociedade*, 1ª ed.: Lua de Papel, 2007.

[21] M. Iansiti and R. Levien, "The keystone advantage," *harvardbusiness.org,* 2004.

[22] M. Schneider and M. Somers, "Organizations as complex adaptive systems: Implications of Complexity Theory for leadership research," *The Leadership Quarterly,* pp. 351-365, 2006.

[23] J. Sutherland and W. J. van den Heuvel, "Enterprise application integration and complex adaptive systems," *portal.acm.org,* 2002.

[24] K. J. Dooley, "A Complex Adaptive Systems Model of Organization Change," *Nonlinear Dynamics, Psychology, and Life Sciences*, 1997.

[25] K. J. Dooley, T. L. Johnson, and D. H. Bush, "TQM, chaos and complexity," *Human Systems Management,* 1995.

[26] J. Holland, "Studying Complex Adaptive Systems," *Journal of Systems Science and Complexity*, 2006.

[27] J. H. Holland, "Adaptation in natural and artificial systems," *mitpress.mit.edu,* 1992.

[28] M. Mitchell, *Complexity - A guided Tour*: Oxford University Press, 2009.

[29] K. Popper, "Three worlds," *bengin.net,* 1979.

[30] S. N. Dorogovtsev and J. F. F. Mendes, "Evolution of networks," *Advances in Physics,* 2002.

[31] S. N. Dorogovtsev and J. F. F. Mendes, "Evolution of Networks - From biological nets to the internet and www," *Oxford University Press,* 2010.

[32] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, "Critical phenomena in complex networks," *Reviews of Modern Physics,* 2008.

[33] M. Newman, *Networks: An Introduction*: Oxford University Press, Inc., 2010.

[34] M. E. J. Newman, "The structure and function of complex networks," *SIAM review,* 2003.

[35] J. A. Zachman, "The Zachman Framework," *, Institute for Framework Advancement. Available ,,* 2007.

[36] J. Zachman, "Enterprise architecture and legacy systems, getting beyond the "legacy"," 2004.

[37] J. A. Zachman, "A framework for information systems architecture," *IBM Systems Journal,* 1999.

[38] A. Gerber, P. Kotze, and V. d. Merwe,A, "Towards the formalisation of the TOGAF Content Metamodel using ontologies," *ICEIS,* 2010.

[39] S. Leist and G. Zellner, "Evaluation of current architecture frameworks," ACM symposium on Applied computing, Dijon, France, 2006.

[40] M. R. De Villiers, "Three approaches as pillars for interpretive Information Systems research: development research, action research and grounded theory," *SAICSIT,* 2005.

[41] M. C. Gonzalez and A.-L. Barabasi, "Complex networks: From data to models," *Nat Phys*, 2007.

[42] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," *NIST*, 2009.

[43] Research Project: http://www.i2s.pt/i2ssite/Projectos/isofin.asp

[44] NIST, www.nist.gov/itl/cloud/upload/cloud-def-v15.pdf

[45] Nuno Ferreira, Nuno Santos, Ricardo J. Machado, Dragan Gasevic. Derivation of Process-Oriented Logical Architectures: An Elicitation Approach for Cloud Design. PROFES, 2012.

[46] Ricardo J. Machado, João M. Fernandes. Heterogeneous Information Systems Integration: Organizations and Methodologies. Markku Oivo, Seija Komi-Sirviö (Eds.), Product Focused Software Process Improvement, LNCS Springer-Verlag.