

The Impact of Online Social Networks on Decision Support Systems

Francisco Antunes^{1,2} and João Paulo Costa^{2,3}

¹ Department of Management and Economics, Beira Interior University, Portugal

² INESCC – Computer and Systems Engineering Institute of Coimbra, Portugal

³ Faculty of Economics, Coimbra University, Portugal
francisco.antunes@ubi.pt; jpaulo@fe.uc.pt

Abstract. Previous research on this matter had already determined that many concepts are encompassed by both online social networking and decision support systems research. Due to the large number of concepts and using clustering techniques, we were able to determine four concept clusters, namely: the technical infrastructure, online communities, network analysis and knowledge management. Then, we intended to gain further knowledge on how those concepts influenced DSS related research and the contribution of each cluster to the support of the phases of decision-making process. We also wanted to perceive the interconnections among the concept clusters themselves, for which we used structural equation modeling techniques.

The obtained results evidence that not only online social networks are being used as a technical infrastructure to support the three decision making phases and to support knowledge management and online communities, but also that the other clusters only regard the intelligence phase of the decision process.

Keywords: Online social networks, decision support systems.

1 Introduction

Online social networking is almost omnipresent in nowadays personal life. However, recent research has revealed that such effect is rather potential than real, in business contexts [13]. The terms “Web 2.0” and “Web 3.0” have become such a commonplace that companies are now pasting it on as a marketing buzzword, with no real understanding of what it really means [15] or real application of its features, to fulfil its potential.

In order to understand the state-of-the-art of the decision support systems (DSS) research and its actual interconnection with online social networking (at both technical and social related levels), we underwent a large study using a thorough analysis of four major bibliographic resources: ISI WOK, SCOPUS, SCIRIUS and EBSCO. To the best of our knowledge, no other study of this extent has been performed until now.

This paper presents the most relevant concepts of the analyzed online social networking and decision support systems research literature, based on both manual and automatic text extraction procedures from the above-mentioned bibliographic resources. As the interconnections of online social networking and decision support

systems concepts were encompassed within the text, we resorted to network text analysis theory, as it assumes that language and knowledge can be modelled as networks of words and relations, encoding links among words to construct a network, analyzing the existence, frequencies, and covariance of terms or concepts. Baring in mind that, within the context of semantic network analysis, a concept is a single idea represented by one or more words [6], we used social network analysis tools to process and represent the obtained network of concepts.

The relevance of the obtained concepts was determined using the eigenvector centrality measure, as it determines the relative influence of concepts within the network. The underlying reason is that it assigns relative scores to all nodes in the network, based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes [4; 17]. Due to the large number of concepts we ran a modularity analysis of the concepts, in order to detect and study underlying concept clusters (communities), based on the Louvain method, which is specially fit for very large networks [3]. The process returned four clusters (technological infrastructure, network analysis, knowledge management and online communities), which represent the contact issues of both fields of research (please see [1] for further insights).

After determining the referred concept clusters we intended to gain further knowledge on how those concepts influenced DSS related research, specifically in terms of their actual support of the three traditional decision-making phases [18], namely the intelligence, design and choice phases. Not only had we wanted to determine the actual contribution of each cluster to the support of the phases of decision-making process, we also wanted to perceive the interconnections among the concept clusters themselves. To do so, we used structural equation modelling techniques. The results of this analysis are shown in section 4.

We present in the next section, the process of publication collecting, which was the basis for this study. Then, in section 3 we present the analysis with the social network techniques. Section 4 presents the exploratory study using structural network equation modelling techniques. The final section summarizes the main conclusions of this study, as well as some future research directions.

2 Data gathering

The time-span of published research chosen for this project is 2003–2010. The initial year is the one when the first publication explicitly dedicated to online social networks, according to [5], occurred.

We used four major bibliographic indexing resources: ISI Web of Knowledge (Science Citation Index Expanded; Social Sciences Citation Index; Arts & Humanities Citation Index; Conference Proceedings Citation Index – Science and Social Science & Humanities –; Current Chemical Reactions; Index Chemicus); SCOPUS (Life Sciences; Health Sciences); SCIRIUS; and EBSCO (Academic search complete). These bibliographic sources provided us the necessary grounds for searching information with scientific indexation across distinct scientific domains (as opposed to internet free content, such as commercial sites and blogs, whose validity is, most of the times, author dependent).

As we knew that using a “([DSS] AND [Web 2.0])” approach would lead us just to technical papers on DSS and internet technology, we decided to opt for a different approach, combining key concepts from both DSS [2] and online social networking [5; 14; 16].

DSS concepts:

- Decision support system (Personal Decision Support System/Group Decision Support System): A system developed to support (a) decision task(s);
- Group Support Systems: the use of a combination of communication and DSS technologies to facilitate the effective working of groups tangled with (a) decision task(s);
- Negotiation Support Systems: DSS where the primary focus of the group work is negotiation between opposing parties;
- Intelligent Decision Support Systems: the application of artificial intelligence techniques to decision support;
- Knowledge Management-Based DSS: systems that support decision making by aiding knowledge storage, retrieval, transfer and application by supporting individual and organizational memory and inter-group knowledge access;
- Data Warehousing: systems that provide large-scale data infrastructures for decision support;
- Enterprise Reporting and Analysis Systems: enterprise focused DSS including executive information systems, business intelligence, and more recently, corporate performance management systems. Business intelligence tools access and analyze data warehouse information using predefined reporting software, query tools, and analysis tools.

Online social networking concepts:

- Internet social networking: refers to the phenomenon of social networking on the Internet. As such, the concept subsumes all activities by Internet users with regard to extending or maintaining their social network.
- Social network sites: web-based services that allow individuals to construct a public or semi-public profile within a bounded system, articulate a list of other users with whom they share a connection, and view/traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site.
- Social software: wikis, micro-blogging and social bookmarking services are types of social software. In an enterprise context, feature-wise, social software is closely related to groupware, though social software originates from the public Internet and is heavily shaped by its users.
- Enterprise 2.0: it describes the adoption of social software in an enterprise context. Much as internet social network denotes the phenomenon and refers to the application of social network software as its main enabling technologies, Enterprise 2.0 refers to the phenomenon of a new participatory corporate culture (with regard to communication and information sharing), which is based on the application of various types of social software technologies. It describes a wider approach that advocates a new organization culture of participation, inclusion, and sharing, rather than simply adopting social software.

- Enterprise social networking: refers to the phenomenon of social networking in an enterprise context, whether using intranet social network or referring to the organizational usage of public social network sites.

After retrieving each data set from the selected data sources, the resulting database presented almost 1.000 records, which needed additional processing. This processing was achieved using SQL (structured query language) in order to perform actions like: the removal of duplicates; the repairing of fields' data types; the removal of abstract books, etc. After completing these steps, the dataset was reduced to 499 records, which needed 'human' processing.

The first 'human' step was to read all the abstracts in order to eliminate papers that were completely out of the scope of this study. Although our search keywords widely narrowed the search within each data source, their data extractors did not differentiate the body of the papers from their references section, for instance. The list of papers was reduced to 326. These ones underwent a thorough reading, in order to assess their contribution to our study. At the end of the process, only 89 papers were selected as an actual interconnection of the decision support systems and online social networking research.

3 Social Network Analysis

As the basis to ascertain the influence of online social networking on DSS are the above-mentioned bibliographic resources, it implied that the required information was text-based and that the interconnections of online social networking and DSS concepts were encompassed within the text. To extract and analyze such relationships, network text analysis theory stands on the assumption that language and knowledge can be modelled as networks of words and relations, encoding links among words to construct a network of linkages, analyzing the existence, frequencies, and covariance of terms or concepts. Knowing that within the context of semantic network analysis, a concept is a single idea represented by one or more words, concepts are equivalent to nodes in social network analysis (SNA) [6]. Consequently, SNA metrics are applicable to our study.

As complex socio-technical systems, online social networking and decision support are dynamic systems. Analyzing such complexity requires tools that go beyond traditional SNA and link analysis [7], namely through Dynamic Network Analysis (DNA), which combines the methods and techniques of SNA and link analysis with multi-agent simulation techniques. To that purpose we used the Automap CASOS (Computational Analysis of Social and Organizational Systems) toolkit to extract the relationship network among concepts of online social networking and decision support systems. AutoMap is a text mining tool that enables the extraction of network data from text, namely four types of information: content (concepts, frequencies and meta-data such as sentence length); semantic networks (concepts and relationships); meta-networks (ontologically coded concepts and relationships – named entities and links); and sentiment and node attributes (attributes of named entities).

We pre-processed the involved text files, manually removing the references and acknowledgments sections. We also created a personalized thesaurus from all the abstracts, in order to replace possibly confusing concepts with a more standard form (e.g. web 2.0, became "web_2d0"; decision support systems, became "dss"; business

intelligence, became `business_intelligence`; etc.) to be applied to all papers. The identification of Named-Entities was also performed using the corresponding Automap tool, in order to create a “delete-list” to remove the references to authors. In addition, all numbers were removed from the text, as well as special characters. Pronoun resolution was also performed using Automap.

To calculate the network metrics, as well as its visualization, we used the open source Gephi 0.8 (beta) software, which is easily integrated with Automap and DNA. Gephi has been used in extensive peer-reviewed scientific research (journals, conferences, workshops and thesis), providing a powerful, free-of-charge, intuitive and easy-to-use analysis tool.

Although several centrality measures can be used to identify key members playing important roles in a network (such as degree, betweenness, and closeness, see, for instance, [8], for further details), we chose Eigenvector centrality as the measure to determine the relative influence of concepts within the network. The underlying reason is that it assigns relative scores to all nodes in the network, based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes [4; 17]. If we rank nodes by Eigenvector centrality, we can determine the key nodes (concepts) in the network. In order to obtain an understandable network representation and due to the large number of involved concepts, we limited (filtered) the concepts to those bearing an Eigenvector centrality greater or equal to 0.15 (reminding that the values were encompassed within 0 and 1).

In spite of the fact that the concept network presents an interesting overview of the overlapping concepts of online social networking and DSS research, it is quite poor at revealing the thematic interconnection of the research field, which is, after all, our purpose. To obviate this problem, we ran a modularity analysis of the concepts, in order to detect and study underlying concept clusters (communities), based on the Louvain method implemented on Gephi 0.8 (beta), which is specially fit for very large networks [3].

3.1 Results

The results returned four concept clusters (again, please see [1] for further insights):

Technical infrastructure (TI) – It represents the concepts which stand for involved technical elements. The technical infrastructure encompasses research that elaborates, develops, proposes and analyzes social networking infrastructures, for distinct underlying purposes (data-gathering purposes, information extraction, taxonomy building, web-computing, consumer support, decision automation, etc.).

Online communities (OC) - This cluster focuses on people, users, teams, which points to community interaction. Instead of focusing on the network topology, it provides a focused view on the effects of online social networking among established online communities, bearing distinct decision purposes or options (academic, acquaintance, leisure, etc.). In addition, research is directed towards group dynamics (formation, cohesion, behavior, etc.) and its effects (actual or perceived) among specific online communities.

Network analysis (NA) – It encompasses a networked analysis of organization, companies and distributed structures. Although directly related to online communities, the main focus of this research lies on the description, community detection,

visualization of social networks, to provide interpretation and decision support according to the social network topology, by means of social network analysis measures (centrality, betweenness, closeness, degree, etc.).

Knowledge management (KM) - Finally, the last cluster represents knowledge management activities, especially around collaboration. The main focus of this theme is to address online social networking (and the so-called “wisdom of the crowds”), using the lens of knowledge management, namely its use (actual and perceived), usefulness and setbacks towards the objectives of knowledge creation, sharing, encoding, retrieval and representation.

4 Structural equation modeling

After finding the previous concept clusters, we wanted to find their interconnection and impact on DSS, namely on their actual contribution in supporting the phases of the decision-making process (as defined by [18]). To accomplish this task, we underwent an exploratory study using structural equation modelling (SEM). SEM is essentially a path analysis baring a structural model [9]. Path analysis is a method, developed by Sewall Wright [19], in which path models with multiple variables are analyzed, and where each variable is standardized. Path analysis is used to describe the directed dependencies among those variables and, therefore, suitable as a confirmation tool, trying to confirm certain pre-established hypotheses, usually causal, or as an exploratory tool, trying to find dependencies among variables. In this latter case, there is no previous model or hypothesis and, after finding significant dependencies among standardized variables, path models can be built. This was the followed strategy.

In order to model the DSS concepts we considered the three main phases that can be found in a decision process and supported by a DSS [18] – the so-called traditional decision-making process: intelligence (DSS-I), design (DSS-D), and choice (DSS-C). Problem finding, analysis and definition occur during the intelligence phase, where divergence is supported through the generation of alternatives and, as the alternatives are evaluated by a group, the convergence process evolves. During the design phase, possible solutions to the problem are generated usually followed by the merging of related ideas and elimination of redundant or irrelevant ones, through a structuring process that might include the elicitation of criteria and their relative importance, as well as the indication of a value system. Choice involves divergent evaluation of the previous set of idea and convergent selection, possibly following an iterative process.

In summary, the variables within our SEM exploratory study were the four main concepts found with SNA techniques and the three DSS concepts: Technical Infrastructure (TI), Online Communities (OC), Network Analysis (NA), Knowledge Management (KM), Support to Intelligent Phase (DSS-I), Support to Design Phase (DSS-D) and Support to Choice Phase (DSS-C).

We built a questionnaire with 26 questions about the connections among the seven concepts related to the considered variables and we used a Likert scale to answer those questions. The typical question is resembles the following: “In this article there is a strong connection between online technical infrastructure and online communities”. The Likert scale had seven levels ranging from “Very strongly agree” (7) to “Very

strongly disagree” (1). The questionnaire was answered regarding the 89 papers that allowed the exploratory SEM study.

In order to conduct the exploratory SEM study we used the WarpPLS 2.0 tool [10-12]. This tool can be downloaded from (<http://www.scriptwarp.com/warppls/>). It is a powerful partial Least Squares based SEM tool that identifies linear and nonlinear relationships and estimates path coefficients. It also calculates p-values and model fit indices. We performed several tests considering lots of different models and we used the p-values and the ARS fit index in order to choose the significant models. The p-value is the probability of obtaining a test at least as significant as the one that was actually observed, assuming that the null hypothesis is true. We rejected the null hypothesis when the p-value was lesser than 0.01 and we considered the path to be statistically significant. The ARS is the average R-squared value and in the case in which the paths are significant, the best model is the one with the highest ARS. The paths coefficients are standardized weights of the regressions and are usually noted by the Greek letter β .

4.1 Results

The significant models and path coefficients are depicted in Figures 1, 2, 3 and 4, determined by the Technological Infrastructure (TI), Network Analysis (NA), Knowledge Management (KM) and Online Communities (OC), respectively.

When comparing these models, we can see that online social networks research has been focusing on TI, as there are more paths involved (Figure 1), and it is throughout this concept that the three phases of decision support are being achieved. Moreover, TI is also impacting the intelligence phase support through OC and KM. Finally, TI impacts the choice phase through KM.

Another major conclusion is that NA, KM and OC are basically being used, in the context of online software networking, to support the intelligence phase of decision processes (models of Figures 2, 3 and 4). It must also be noticed that NA is impacting both KM and OC as application purposes by themselves (Figure 2).

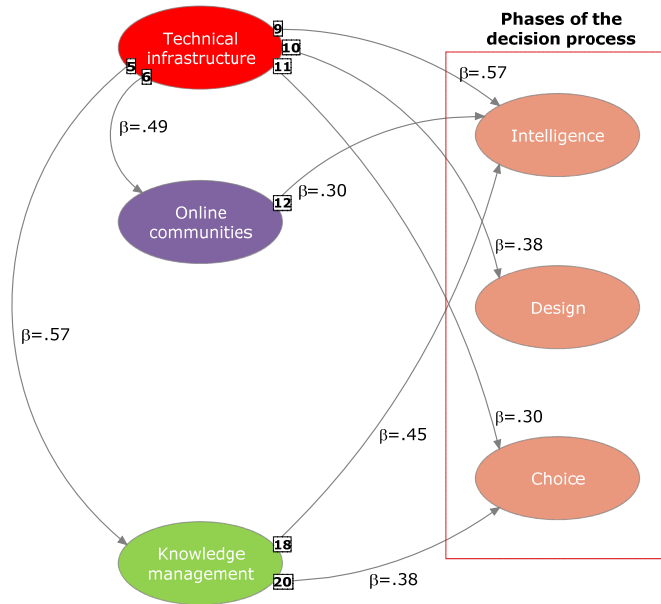


Fig. 1. Path model of the TI variable.

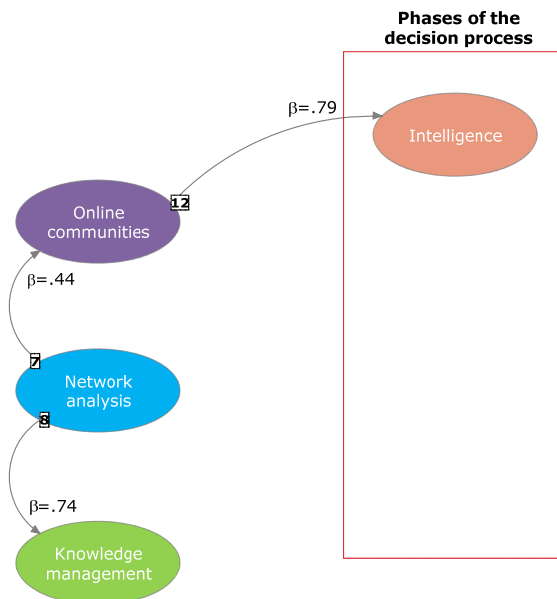


Fig. 2. Path model of the NA variable.

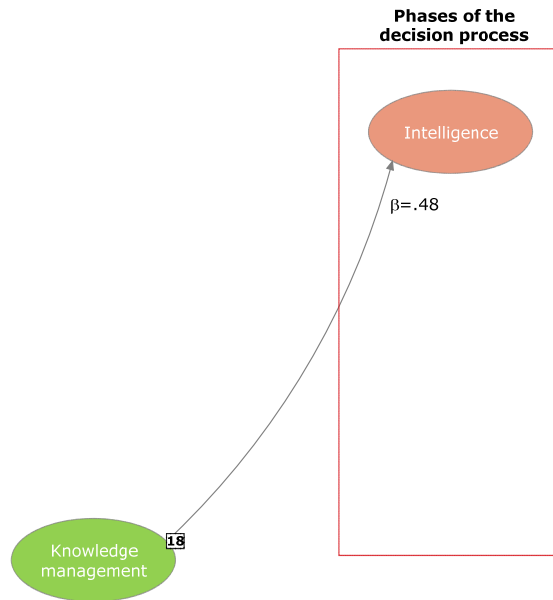


Fig. 3. Path model of the KM variable.

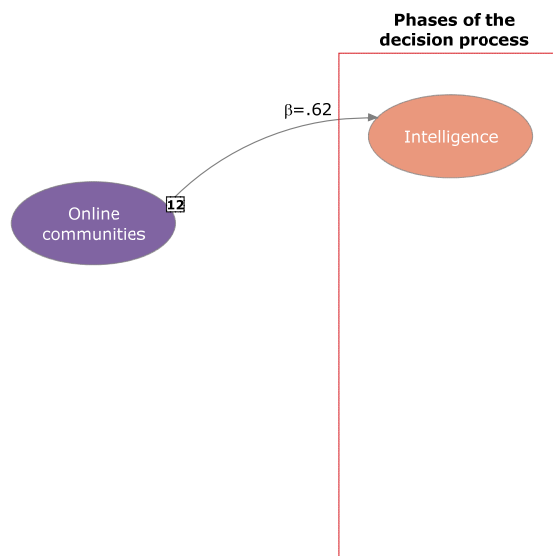


Fig. 4. Path model of the OC variable.

5 Conclusions and future research questions

We underwent an exhaustive exploration of the so called “indexed publications” of research on applying online social networks to support decision making.

We applied social network analysis techniques and software in order to find the key concepts of the ongoing research and we found four concept clusters: technical infrastructure, representing the involved technical elements; online communities, representing people, users, teams and their interaction; network analysis, encompassing the analysis of organizations and distributed structures; and knowledge management, representing the objectives of knowledge creation, sharing, encoding, retrieval and representation.

We then wanted to find the impact of these concepts on the three main phases of the traditional decision making processes: intelligence, design and choice. We used SEM techniques to find path models among the concepts and found that online social networks are being used as technological infrastructures to support the three decision making phases and also to support knowledge management and online communities. Another major finding is that, besides the technological infrastructure, the other concepts are only being used to support the intelligence phase of a decision process.

The conclusions of this study are limited by the inherent subjectivity of the study, because several key options were taken solely based on the authors opinions.

In spite of the revealing results on the subject, a new unanswered question seems to arise: considering that online social networks concepts and applications are relatively new concepts, is that fair to study them under traditional decision making paradigms (like the three phase model)?

To answer this question, we will try to understand in future work if other decision making paradigms are being put forward by online social networks.

Acknowledgements: This work has been partially supported by the Portuguese Foundation for Science and Technology under project grant PEst-OE/EEI/UI308/2014.

6 References

1. Antunes, F and Costa, J P, Disentangling Online Social Networking and Decision Support Systems Research Using Social Network Analysis. In: Proceedings of the EWG-DSS Liverpool-2012 Workshop on Decision Support Systems & Operations Management Trends and Solutions in Industries, Liverpool, UK, April 12th-13th, (2012)
2. Arnott, D and Pervan, G, Eight key issues for the decision support systems discipline. *Decision Support Systems*, 44:657–672, (2008)
3. Blondel, V D, Guillaume, J-L, Lambiotte, R and Lefebvre, E, Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, P10008(2008)
4. Bonacich, P, Some unique properties of eigenvector centrality. *Social Networks*, 29(4):555–564, (2007)
5. Boyd, D M and Ellison, N B, Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210-230, (2008)
6. Carley, K M, Network Text Analysis: The Network Position of Concepts. In: Roberts C W (ed) *Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts*

- and Transcripts, Laurence Erlbaum Associates, Inc., Mahwah, New Jersey, pp 79-100, (1997)
7. Carley, K M, Diesner, J, Reminga, J and Tsvetovat, M, Toward an interoperable dynamic network analysis toolkit. *Decision Support Systems*, 43:1324-1347, (2007)
 8. Freeman, L C, Centrality in Social Networks - Conceptual Clarification. *Social Networks*, 1(3):215-239, (1978-1979)
 9. Kline, R B, Principles and practice of structural equation modeling. The Guilford Press, New York, NY, (1998)
 10. Kock, N, Using WarpPLS in E-collaboration Studies: An Overview of Five Main Analysis Steps. *International Journal of e-Collaboration*, 6(4):1-11, (2010)
 11. Kock, N, Using WarpPLS in e-Collaboration Studies: Descriptive Statistics, Settings and Key Analysis Results. *International Journal of e-Collaboration*, 7(2):1-18, (2011)
 12. Kock, N, Using WarpPLS in e-Collaboration Studies: Mediating Effects, Control and Second Order Variables, and Algorithm Choices. *International Journal of e-Collaboration*, 7(3):1-13, (2011)
 13. Lager, M, No one's Social (yet). *CRM Magazine*, 13(6):29-33, (2009)
 14. McAfee, A P, Enterprise 2.0: the dawn of emergent collaboration. *MIT Sloan Management Review*, 47(3):20-29, (2006)
 15. O'Reilly, T, What is Web 2.0? - Design patterns and business models for the next generation of software. <http://oreilly.com/web2/archive/what-is-web-20.html>, Cited 2011-07-06, (2005)
 16. Richter, D; Riemer, K and Brocke, J v, Internet Social Networking: Research State of the Art and Implications for Enterprise 2.0. *Business & Information Systems Engineering*, 2:89-101, (2011)
 17. Ruhnau, B, Eigenvector-centrality—a node-centrality? *Social Networks*, 22(4):357–365, (2000)
 18. Simon, H A, The New Science of Management Decision. Prentice Hall, Englewoods Cliffs, NJ, (1977)
 19. Wright, S, The method of path coefficients. *Annals of Mathematical Statistics*, 5(3):161–215, (1934)