IRIT Report
IRIT/RR--2010-18--FR

Towards a Collaboration Network for the EURO Working Group on DSS

Fátima Dargam [a1], Rita Ribeiro [b], Pascale Zaraté [c]

[a]*SimTech Simulation Technology, Graz, Austria*
(F.Dargam@SimTechnology.com)

[b]*UNINOVA – CA3, Lisbon, Portugal*
(rar@uninova.pt)

[c]*Toulouse University, IRIT / INPT – ENSIACET – GI, France*
(zarate@irit.fr)

Abstract

The primary objective of this work is to develop a social-academic network for the EURO Working Group on Decision Support Systems (EWG-DSS), using as its defined population, the academic members of the group. Our motivation for carrying out this work comes from the importance and the need of analysing and representing the EWG-DSS various relationships that academically link its current 104 members. With this project, we will be able to evaluate the group's collaboration dynamics since its foundation in 1989, up to the present moment. As a by-product, we aim to encourage new research and promote further collaboration among the academic members of the group in common projects and joint-publications, since their current areas of research will be revisited and focused by this study. Another expected result of this work is to evaluate social informal relationships among the group members.

Keywords: social network, EURO working group EWG-DSS, decision support systems

[1] Fátima Dargam also collaborates as senior researcher at the Institute of Logic and Theory of Science – ILTC (www.iltc.br), in Rio de Janeiro, Brazil.

Introduction

Social networks are usually formed by actors (individuals organizations, teams), each having ties (connections) with others, which indicate their relationships [Newman 2001a, 2001b, 2003]. They are usually defined in terms of nodes (individuals/actors) and connections (ties), which indicate the relationships between the actors. Such networks can represent different real-life communities and play an important role in determining the way problems are solved, organizations are run, and the degree to which individuals succeed in achieving their goals [Newman 2001b, 2003], [Weimao and Simas 2007], [Rocha et al 2005].

The power of social network analysis stems from its difference from traditional social scientific studies [Newman 2003] [Freeman 2000]. Social network analysis produces an alternate view, where the attributes of individuals are less important than their relationships and ties with other actors within the network. They focus on how the structure of ties affects individuals (persons, organizations, states) and their relationships, rather than treating them as discrete units of analysis.

Social networks analysis faces two drawbacks [Newman 2003], most of the times due to a small number of actors (poor statistical accuracy and relevance) and intrinsic subjectivity (biases on the part of respondents to questionnaires or interviews). These phenomenon's would also occur on our study, since the EWG-DSS group has a relatively small number of members and the data collection was partially obtained via email, or links posted in the group website. To avoid the problems of poor statistical accuracy and intrinsic subjectivity, we use concepts and methods defined for a specific type of social networks, denoted as weighted collaboration/affiliation networks [Newman 2001 b].

Collaboration networks are affiliation networks in which actors collaborate in associations or clubs of some kind and the connections (ties) are established by common group memberships [Newman 2003]. For example, a network of scientists, in which any pair is connected if they co-authored one or more papers, is a good example of a Boolean collaboration network [Newman 2001a, 2001b]. When we consider the number of papers co-authored between two authors (two nodes) as the strength of the relation we have a weighted collaboration network [Newman 2001b].

In this work we start by using the generalized Jaccard similarity measure [Rocha et al 2005] to construct the social-academic fuzzy (weighted) networks [Simas 2008], from the initial Boolean matrices of: authors versus papers; and papers versus topics (keywords of interest for the DSS area). After constructing the weighted network we analyse it with the PAJEK Network framework [PAJEK] and also with the NWB Network Workbench framework [NWB]. Both frameworks are academic public-domain MS-Windows-based programs for network analysis and visualization.

Objectives and Expected Results

The main motivation for this work is to represent and analyse the various relationships that academically link the current 104 members of the EURO Working Group on DSS. For this purpose we developed an academic collaborative network to enable us to evaluate the group's collaboration dynamics since its foundation in 1989, up to the present moment. As a by-product, we hope to promote further collaboration among the academic-members of the group in common projects and joint-publications, since their current areas of research will be revisited and focused by this study. Another possible result of this work is to evaluate social informal relationships among the members of the group.

Methodology Overview

The methodology proposed for this work includes applying weighted graphs methods [Rocha et al 2005] [Newman 2001b] [Simas 2008] to develop a social-academic network, and then using existing software frameworks NWB Network Workbench [NWB] and PAJEK Network [PAJEK] to analyze and unfold existing relationships between members of the EWG-DSS.

Our network comprises the data gathered from the EWG-DSS members, from 1989 (group foundation) until 2008: a period of 19 years. To construct and analyze the social academic network, we follow five main steps:

1. **Acquisition Process:** Collecting the input data in a matrix format (MS-Excel files), which could relate authors and their papers, as well as the papers classified into topics;
2. **Extraction Process:** Creating the input files with nodes and labels to enable them to be manipulated by the network tools: PAJEK and NWB;
3. **Transformation Process:** Using Jaccard similarity measure [Simas 2008], we constructed a set of weighted networks by combining matrices including authors, publications and research topics;
4. **Weighted Network graphical analysis:** using PAJEK and NBW graphical tools we analyze the main characteristics of the EWG-DSS group;
5. **Weighted Network statistics:** using PAJEK and NBW statistical tools we discuss the main aspects of the EWG-DSS academic network.

Below we describe the proposed methodology steps, including the processes listed above for creating the weighted network and for analyzing the resulting social-academic network.

Step 1.
Acquisition of the academic production of the EWG-DSS members

Members involved

All the members of the EWG-DSS were requested by the coordination members of the group to submit to this project the relevant information, concerning their list of main publications since 1989, stating for each of them the authors names, title, journal/conference/..., editors, year of publication, as well as its main area(s) of research. Another important requested information for the network, was the indication of research cooperation with some other member(s) of the EWG-DSS group, if any. As a result, 70 members have replied to us with their corresponding data. In the sequel of this paper, they are going to be referred to as Authors. Table 1, to follow, shows the selected 70 authors of the EWG-DSS group, who were used for this case-study.

Ai	Author Name	Ai	Author Name	Ai	Author Name
A1	Arijit Bhattacharya	A25	Fréféric Adam	A49	Pascale Zaraté
A2	Adla Abdelkader	A26	Frieder Stolzenburg	A50	Peter Gelleri
A3	Albert A. Angehrn	A27	Frits Claassen	A51	Peter Keenan
A4	Alessio Ishizaka	A28	Ilya Ashikhmin	A52	Philip Powel
A5	Alexis Tsoukias	A29	Inès Saad	A53	Philippe Lenca
A6	Ana Respício	A30	J.Jassbi	A54	Pierre Kunsch
A7	Antonio Jimenez Martinez	A31	Jacques Calmet	A55	Rita Ribeiro
A8	Asis Kr. Chattopadhyay	A32	Jean Charles Pomerol	A56	Rudolf Vetschera
A9	Bertrand Mareschal	A33	Jean Pierre Brans	A57	Sanja Petrovic
A10	Bojan Srdjevic	A34	João Carlos Lourenço	A58	Suzanne Pinson
A11	Boris Delibasic	A35	Jochen Pfalzgraf	A59	Tawfik Jelassi
A12	Caludia Loebbecke	A36	Johannes Leitner	A60	Thanasis Spyridakos
A13	Camille Rosenthal Sabroux	A37	Jorge Freire de Sousa	A61	Yi Yang
A14	Carlos Antunes	A38	Jorge Pinho de Sousa	A62	Thomas Soboll
A15	Carlos Bana e Costa	A39	Jose Maria Moreno Jimenez	A63	BAZZANA Flavio
A16	Christer Carlsson	A40	Ladislav Lukas	A64	Guilan Kong
A17	Csaba Csaki	A41	Li Ching Ma	A65	Jason Papathanasiou
A18	Dirk Kenis	A42	Luís Cândido Dias	A66	Mikael Mihalevich
A19	Dobrila Petrovic	A43	Marko Bohannec	A67	Taghezout Noria
A20	Dorien De Tombe	A44	Michael Bruhn Barfod	A68	Warren Elliott Walker
A21	Eduardo Manuel Natividade Jesus	A45	Miklos Biros	A69	José Vicente segura Heras
A22	Fatima Dargam	A46	Natalio Krasnogor	A70	Antonio Rodrigues
A23	Franck Tetard	A47	Nguyen Dinh Pham		
A24	Frantisek Sudzina	A48	Olaf Herden		

Table 1 – *Authors Ai = Authors; for i = {1,…, 70}.*

Publications

A total of 1350 publications were taken into consideration for this case-study. The publications involved at least one of the 70 authors, members of the EWG-DSS, listed above. Only international publications that appear after the date that the authors became a member in the EWG-DSS were considered. Outside collaborators of the publications, not members of the EWG-DSS, were not included in the network. The list of 1350 publications was compiled into an Excel file by the team involved in this project, so that the publications could be then easily used in the network pre-formatting process, described in the sequel. For space reasons, we do not publish the publications list in this paper.

For representing the publications and their sources, we start by building a matrix where the lines contain the authors, members of the DSS group, and the columns comprise the set of papers written by the members, within the period of study. This task defines the Boolean matrix of relationships: R(A,P), among authors-members of the EWG-DSS group (Ai) and the papers published by them (Pj). Table 2 below depicts the template for this boolean matrix, where APij is either 1 or 0, depending on an author "*i*" having collaborated with a certain paper "*j*" or not, respectively. The number of lines "*n*"refers to the number of authors considered (70) and "*m*" is the number of papers collected from the selected members.

Authors	**Publications**			
	P1	**P2**	**....**	**Pm**
A1	Ap*ij*			
A2				
....				
An				

Table 2: *Boolean Matrix R(A,P) relating authors and their publications.*

Step 2.
Representation of the relationships among the EWG-DSS members

To represent the relationships among the acquired data, we have considered the concept of collaboration/affiliation network already described, where the authors are connected via their publication collaboration. Another relationship considered in this study was related to the publications' areas of research. For implementing this type of relationship, it was urged that some topics of interests were identified, incorporating the main research areas pursued by the EWG-DSS members. This set of topics was defined by the team involved in this project, by taking into account the available input data of the publications.

Topics (Areas of Research)

A total of 34 areas of research, related to the main interest of the working group on decision support systems, were considered. Table 3 shows the topics that were used for classifying the papers published by the group members. For this first analysis, it was assigned one topic per publication only.

The methodological step of implementing the existing relationships of the network, concerns to create the Boolean matrix of relationships, denoted as S(P,T), among the publications of the members (Pi) and the pre-defined set of relevant topics for the EWG-DSS members (Tj). This task was done in the same way that we have created the relation R(A,P) for authors and their publications.

Table 4 depicts the template for this Boolean matrix, where PT*ij* is either 1 or 0, depending on the classification of paper Pi on topic Tj or not, respectively. The index "*n*" corresponds to the number of papers within the publications input data file *(n= 1350)*, and "*m*" corresponds to the number of relevant topics defined *(n= 34)*. To populate the matrix

S(P,T), we need to assign all papers to either 1 or more topics, until all cells PT*ij* of the matrix are filled with either 0 or 1.

#	Topic Code	Topic Abbreviation	Research Topic Description
1	T1	BM	Business Models
2	T3	CD	Collaboration Dynamics
3	T4	CDSS	Cooperative Decision Support Systems
4	T5	DA	Decision Analysis
5	T6	DAP	Decision Aiding Process
6	T7	DM	Data Mining
7	T8	DSS	Decision Support Systems
8	T9	E	Evaluation
9	T10	EB	E-Business
10	T11	ERP	Enterprise resource Planning
11	T12	ES	Expert Systems
12	T13	ET	Economic Theory
13	T14	FS	Fuzzy Sets
14	T15	GDN	Group Decision and Negotiation
15	T17	IR	Information Retrieval
16	T18	IS	Information Systems
17	T19	ITT	Information and Telecommunication Technology
18	T20	KM	Knowledge Management
19	T21	MAS	Multi-Agent Systems
20	T22	MCDA	Multiple Criteria Decision Aiding
21	T23	MLDM	Management Learning and Decision Making
22	T24	N	Network
23	T25	OR	Operations research
24	T26	PA	Preference analysis
25	T27	PE	Performance Evaluation
26	T28	PM	Preference Modelling
27	T29	PPS	Production Planning and Scheduling
28	T31	SCM	Supply Chain Management
29	T32	SD	Sustainable Development
30	T33	SN	Social Networks
31	T34	SS	Simulation Systems
32	T35	SSES	Systems Software Evaluation and Selection
33	T37	VC	Virtual Communities
34	T38	C	Context

Table 3: *Research Topics*

Papers	Topics			
	T1	T2	....	Tm
P1	PTij			
P2				
....				
Pn				

Table 4: *Boolean Matrix S(P,T) relating Publications and their Topics.*

Step 3.
Construction of the combined matrices including authors, publications and research topics & creation of the networks

After identifying the relevant key-information from the members, illustrated in Table 3, and determining the relationship matrices among authors x papers and papers x topics, as shown in Table 2 and Table 4, we can compose the initial matrices that are needed to construct the specified weighted networks. For accomplishing this step, we use a simple matrix multiplication process to combine the information of both Boolean matrices S(P,T) and R(A,P), via their input weighted networks as described in the sub-steps that follow.

Moreover, we create the set of weighted networks by using input graphs for the network simulation tools, by applying the extended Jaccard similarity measure, as suggested in [Rocha et al 2005]. This process will be implemented for both relations R(A,P), *Authors x Publications*, and S(P,T), *Publications x Topics*.

Authors x Publications Networks derived from the Matrix R(A,P)

In this step, we have used the input relational matrix **R(A,P),** implemented in an MS Excel document, expressing the relationships between *authors* and their *publications*, to obtain the networks **Author_AP** and P**ubl_AP.** Those networks respectively focus on the connections among Authors related by their publications; and among Publications related by their authors, into two respective output files. The graphic below illustrates the input matrix and its respective outputs.

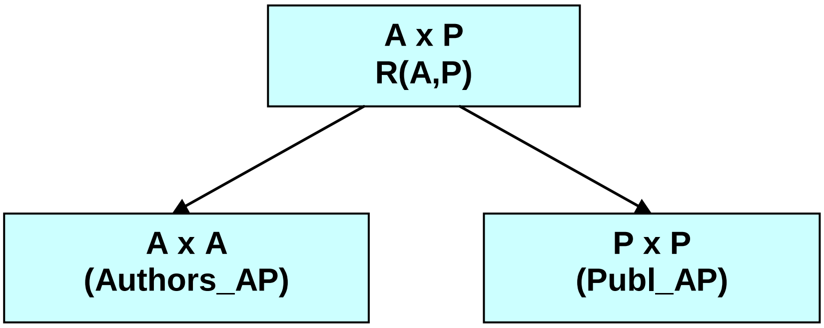

Figure 1: *Social networks for Authors and Publications. From the relation R(A,P), two types of networks, Authors_AP and Publ_AP, were obtained by applying similarity function.*

By building the networks **Author_AP** and P**ubl_AP**, we derive the existing relationships among the considered authors (A x A), and publications (P x P), relative to their ties from the original matrix of authors and publications (A x P) as illustrated in Figure 1.

Publications vs Topics Network derived from the Matrix S(P,T)

In the same way done with R(A,P), relating the *authors* and their *publications*, we now use the input matrix **S(P,T),** expressing the relationships between the *publications* and *research topics*, to obtain the networks **Publ_PT** and **Topic_PT.** The network **Publ_PT** shows all the connections among the given papers and their areas. Whereas, the network

Topic_PT shows the relationships among the considered research topics (T x T), as shown in the graphic of Figure 2 below.

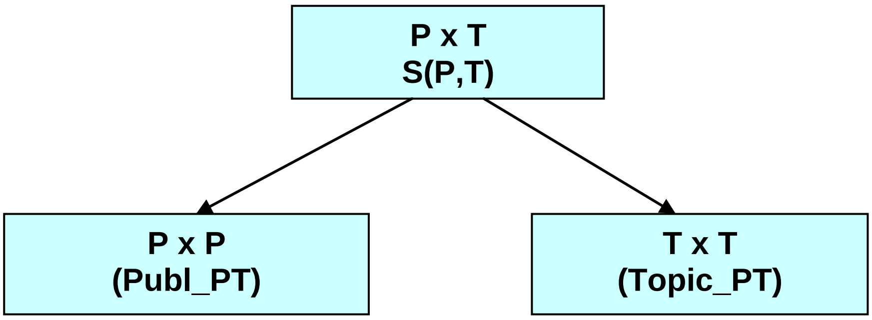

Figure 2: *Social networks for Publications and Topics. From the relation S(P,T), the networks Publ_PT and Topic_PT were obtained by applying similarity function*.

The results obtained for TxT (Topics x Topics) were not discriminative, hence no analysis was performed at this stage of our project. For this reason, we are assuming the network Topic_PT now as “not relevant”.

Authors vs Topics Networks derived from combining R(A,P) and S(P,T)

In order to obtain the existing relationships among authors regarding their areas of research, here denoted as “topics”, we combine, via matrix multiplication, the relations Authors x Publications (R(A,P)) with the Publications x Topics (S(P,T)), so that we are able to produce the matrix **U(A,T)** establishing the connections Authors x Topics. We then use this U(A,T) matrix to obtain the networks **Authors_AT and Topics_AT,** which show all the connections among the authors within their working topics, as well as the relationships among topics regarding the authors, respectively. The graphic in Figure 3 below shows the input matrix and its output files.

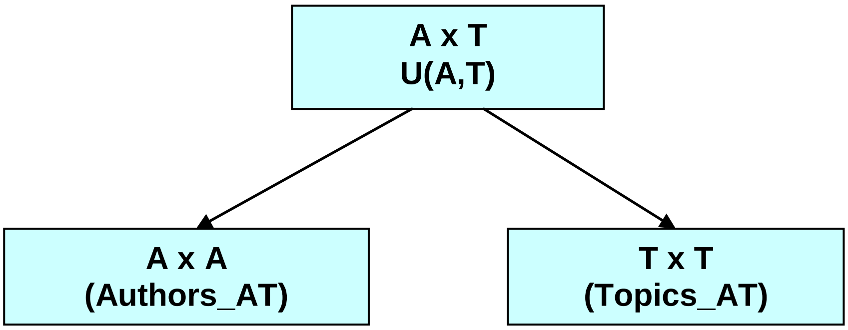

Figure 3: *Social networks for Authors and Topics. From the derived relation U(A,T), the networks Authors_AT and Topics_AT were obtained by applying the similarity function*.

Step 4: Graphical Representation and analysis of the Network

By constructing the set of weighted networks for the EWG-DSS group, we were able to visualize and analyze them with PAJEK [PAJEK] and NWB [NWB] network frameworks. The PAJEK framework is more dedicated to large network analysis, whereas the NWB Network Workbench is a framework for pre-processing, modelling and analysing small networks. Both of them are MS-Windows-based programs designed for network analysis and visualization.

In this step, the output networks files "Authors_AP"; Publ_AP; Publ_PT; Authors_AT" and Topics_AT were used to produce graphical representations about the EWG-DSS social academic network.

The sequel of this section describes some relevant statistic and metric analysis, which were obtained via the adopted network frameworks: NWB Network Workbench and Pajek Network Tool. We start by publishing the graphical visualizations of all the network files. Remarks about the metrics of each network file follow next.

Network View of Authors relationships using Publications (Authors_AP Graphs)

The graphs illustrated in Figures 4 and 5 represent how the author nodes are connected among themselves, with relation to their publication-collaboration.

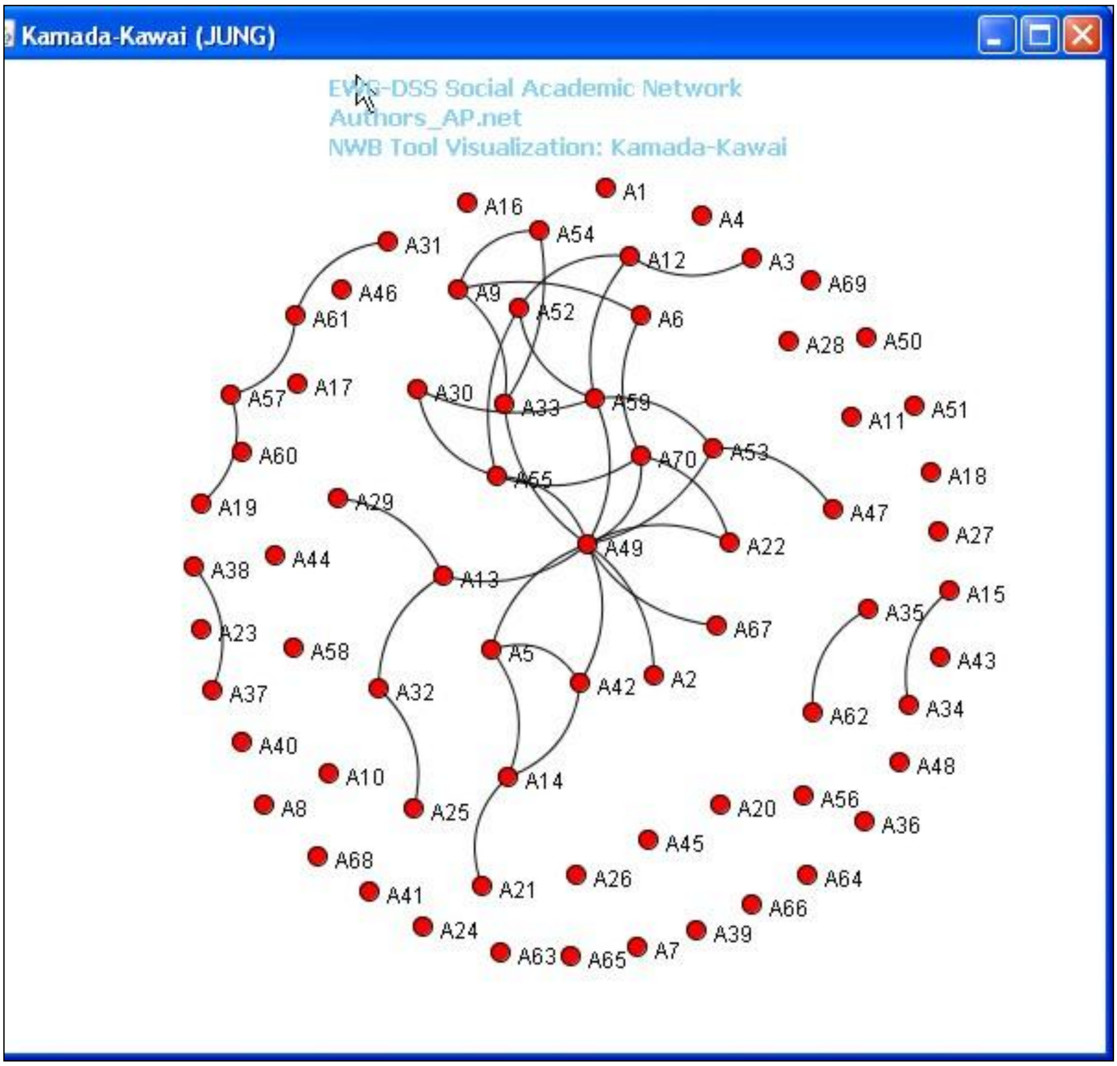

Figure 4: *Visualization of the Author_AP.net, related by the authors' collaboration in publications. Network graphically represented in NWB, via the Kamada Kawai algorithm.*

Figure 4 shows a network graphically represented in the NWB framework, using the Kamada Kawai visualization algorithm. The Kamada-Kawai Algorithm is a force directed layout algorithm, which considers the force between any two nodes. In this algorithm, the nodes are represented by steel rings and the edges are springs between them. The attractive force is analogous to the spring force and the repulsive force is analogous to the electrical force. The basic idea is to minimize the energy of the system by moving the nodes and changing the forces between them [NWB-KK].

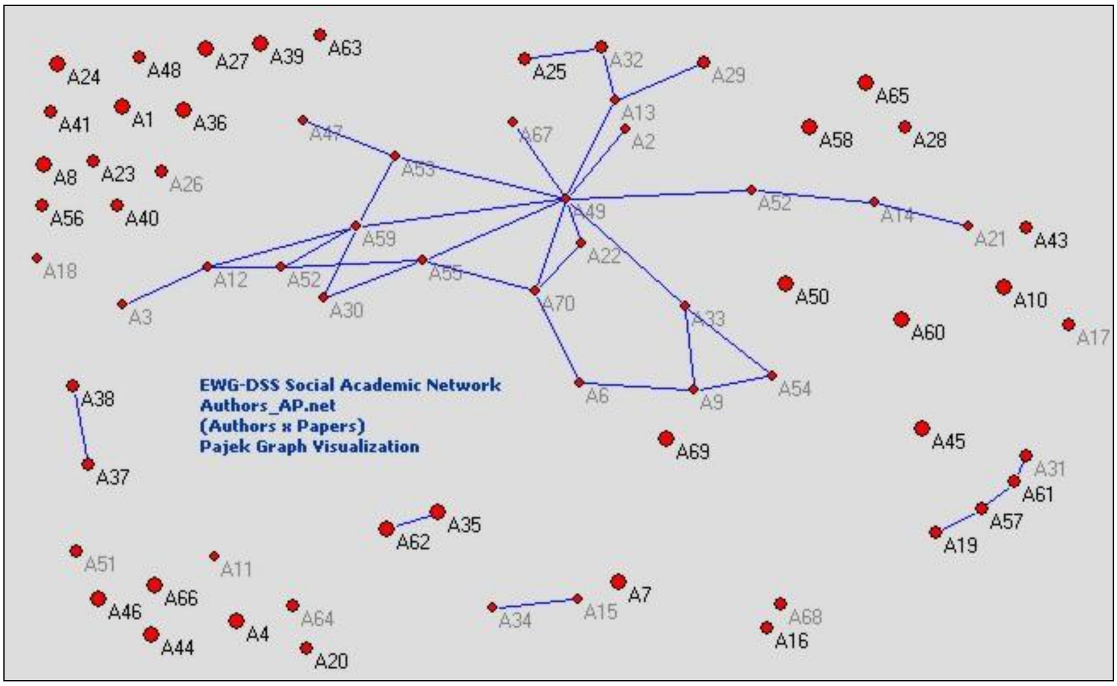

Figure 5: *Visualization of the Author_AP network, graphically represented in PAJEK.*
Ai = Authors; for i = {1,…, 70}

In Figure 5, the Pajek visualization of the Author_AP network shows larger nodes, which is only due to the 3D representation. The importance of both graphical representations in Figures 4 and 5 is to depict work relationships between members in terms of co-authored papers. For example, A33 (Jean-Pierre Brans) has 3 "direct-links", namely: A54 (Pierre Kunsch), A9 (Bertrand Mareschal) and A49 (Pascale Zaraté), but A49 is not a "direct-link" of A9 and A54. Another interesting example provided by this graph is that A57 (Sanja Petrovic) co-authored papers with A19 (Dobrila Petrovic) and A61 (Yi Yang). However, the later two did not publish papers together.

This visualizations of Figure 4 and of Figure 5, clearly show a central main cluster relationship of authors, which is directly linked from the node A49 to other 24 nodes. The big cluster linked by the author A49 (Pascale Zaraté), who is a founding member of the EWG-DSS group, as well as its coordinator for 10 years by now, is then explained by the fact that, through the years, this particular group member has continuously invested great effort in publishing special editions of journals related to the area of DSS, in direct or indirect cooperation with 24 other members of the group. Hence, the observed central cluster of the Authors_AP network can be well explained in this way. Furthermore, although the graphs do not provide this information, whenever the EWG-DSS group would take part in Conference Streams and Workshops, some publications were promoted by A49 (Pascale Zaraté), which resulted in her central role within the group.

Another aspect that can be observed in both Figures 4 and 5 is that there are around 33 group members without any co-authored papers with other members, i.e. they do not yet show interaction with the other nodes of the network (the so-called “isolated-nodes”). Clearly, the EWG-DSS network helps us to identify this weak point in joint-collaboration within the group, so that this aspect can be encouraged to be improved.

Network View of Publications with relation to their Authors (Publ_AP Graphs)

The three figures to follow illustrate the graphs that represent how the publications’ nodes are connected among themselves, with relation to their authors. As done previously, these graphs were produced using the network simulation tools NWB and PAJEK.

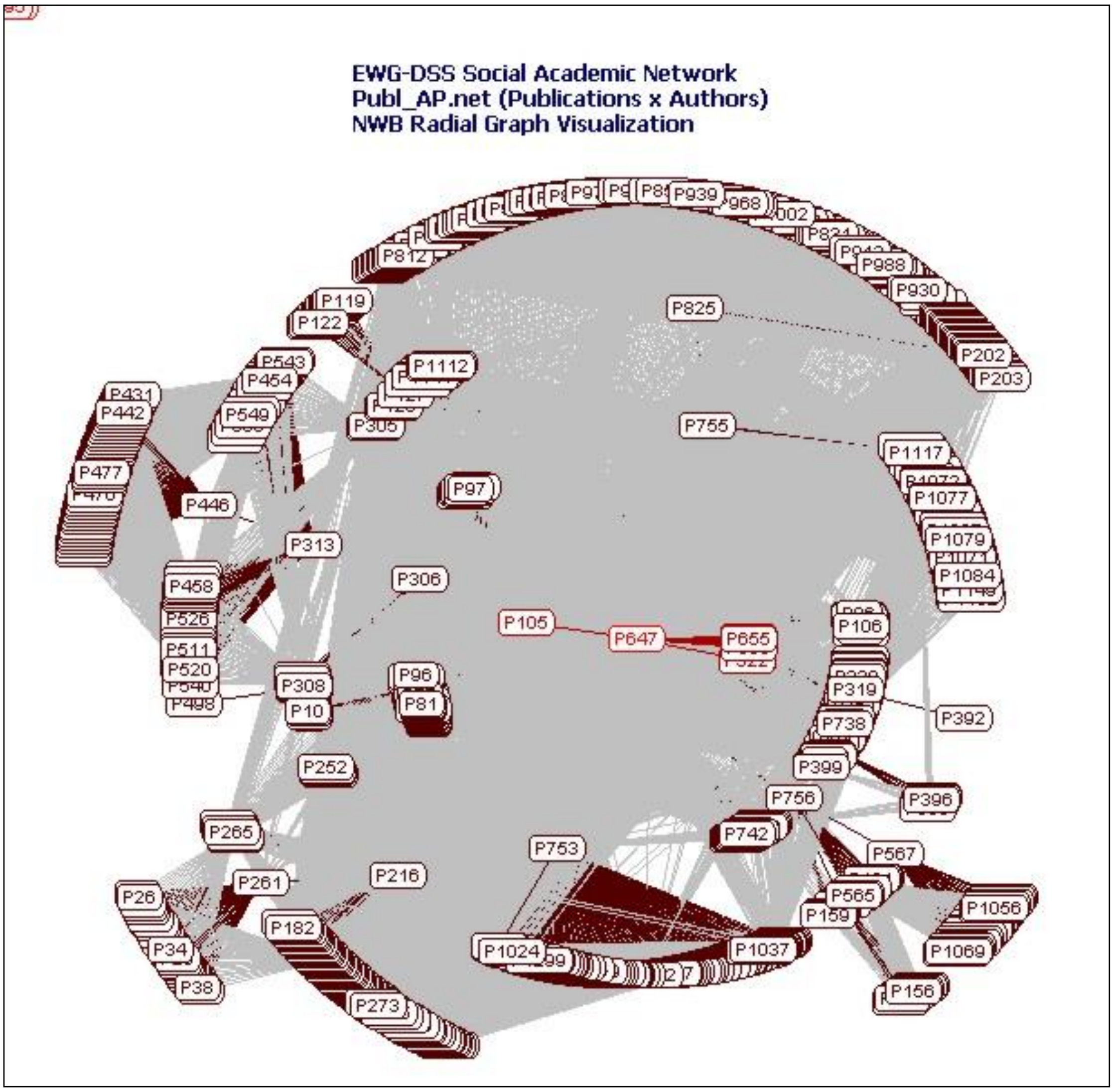

Figure 6: *NWB Radial Graph Visualization of the Publ_AP network, showing the collaboration of the joint- publications relative to their authors, where Pj = Publications of the authors; for j = {1,…, 1350}*

In Figure 6, we observe that the publications network Publ_AP is visualized via a radial tree layout, in which a single node is placed at the center of the graph (in the case, node P647) and all the other nodes are laid around it. The entire graph is like a tree rooted at the central node. The central node is referred to as the focus node and all the other nodes are arranged on concentric rings around it. Each node lies on the ring corresponding to its shortest network distance from the focus node. Any two nodes joined by an edge in the graph is referred to as neighbours. Immediate neighbours of the focus lie on the smallest inner ring, their neighbours lie on the second smallest ring, and so on.

The publication of node P647 (*"An Approach to Support Negotiation Processes with Imprecise, Information Multicriteria Additive Models", João N. Clímaco and Luis C. Dias, in Group Decision and Negotiation, 15(2):171 184, 2006*) was arbitrarily chosen to be the focus node of the radial graph illustrated in Figure 6. This publication node has four nodes as immediate neighbours and is certainly not the most connected publication, in terms of cooperation with other pieces of work, considering the sparse neighbour nodes that are shown in its closest rings. In further reports of this on-going project, we plan to select and identify specific publications to be analysed in relation to its joint-cooperation.

In Figure 7 below, we observe the same publications network Publ_AP, now visualized by the network framework PAJEK. In this graphical representation, although still difficult to be analysed, it is now possible to better visualize the various existing clusters of publications that are somehow related to each other via joint-cooperations of groups of authors, working within specific research areas.

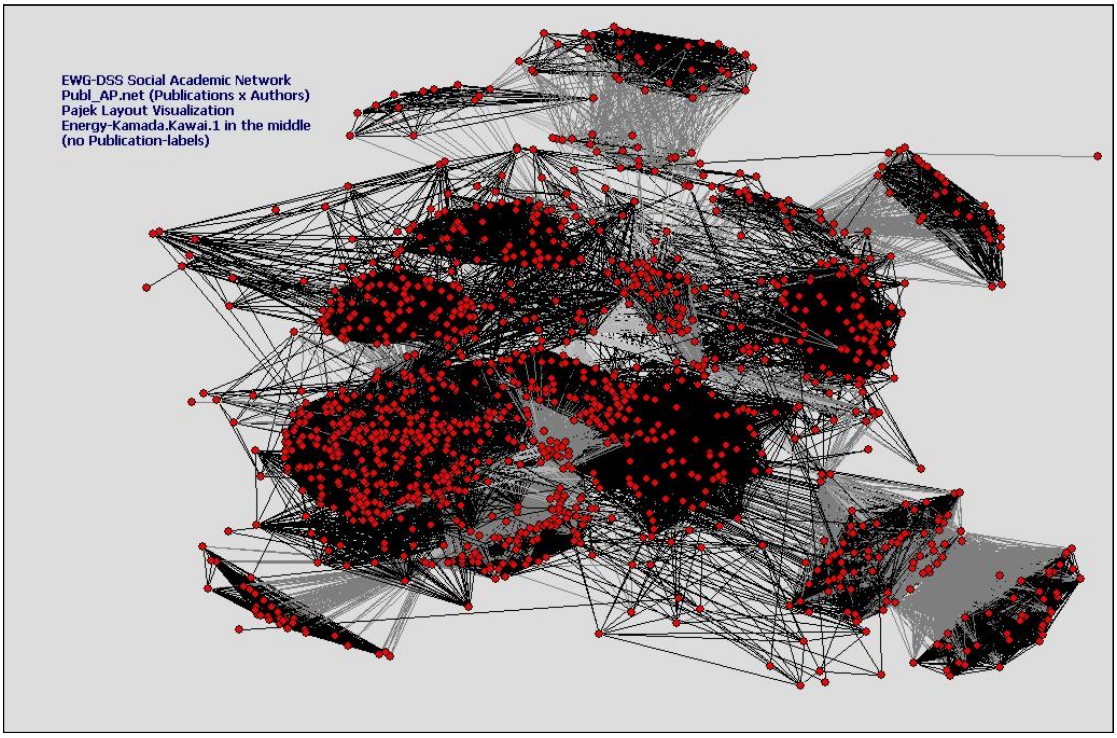

Figure 7: *Visualization of the Publ_AP network, graphically represented in PAJEK. Pj = Publications of the authors; for j = {1,…, 1350}*

The layouts used in both Figures 7 and 8, within the PAJEK network framework, are both based on *Kamada-Kawai* visualizations of the Publ_AP network. In Figure 7, the Energy Kamada-Kawai algorithm is applied with the parameter “Fix one in the middle”, whereas in Figure 8, the parameter used was “Separate Components”, as it can be clearly observed.

Even though in both graph representations, we have activated the option “no labels”, so that more of the shape of the graph could be printed without showing the 1350 publication labels, we have to admit that analyzing such a number of nodes in one only network becomes a hard task, whose main difficulty resides on the visualization of the proper characteristics for taking the relevant conclusions.

Nevertheless, when we observe the graphical representation in Figure 8, using the “Separate Components” visualization parameter of PAJEK, we can confirm that there is a main area of publication, which forms a cluster of the network Publ_AP significantly bigger than any of the others (cluster in the upper left part of the Figure). This cluster contains the publications related to the application area of “Decision Support Systems”, in which many authors cooperate in joint-publications.

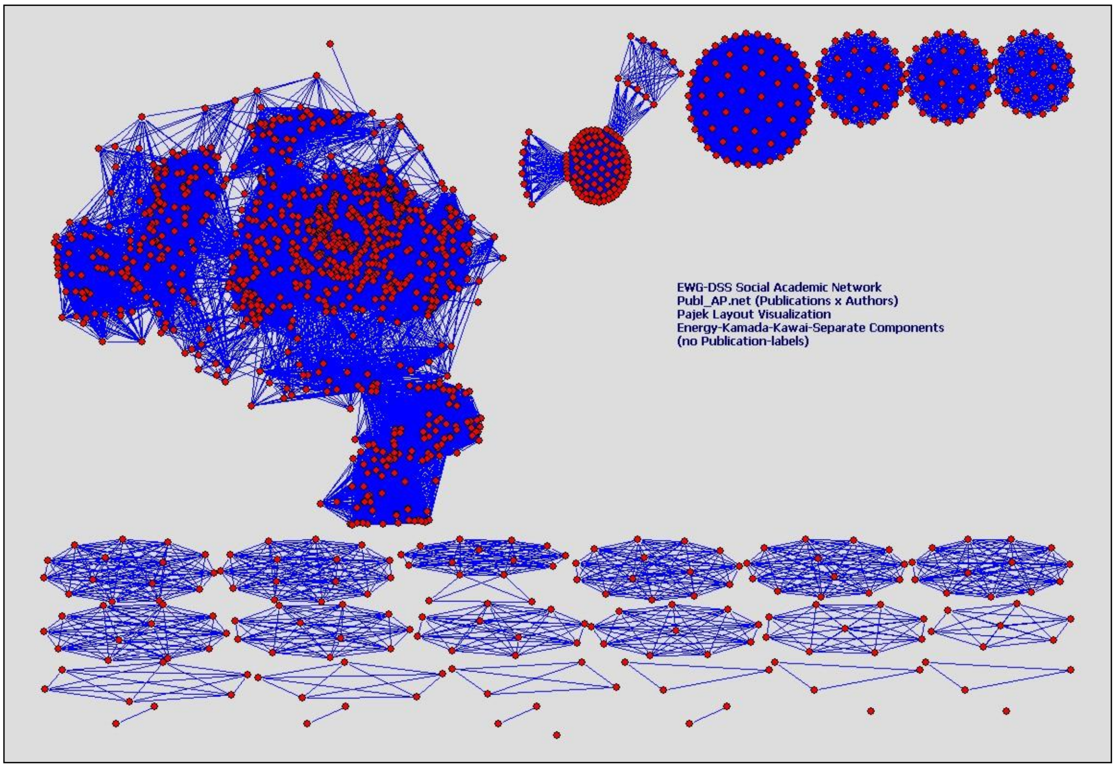

Figure 8: *Visualization of the Publ_AP represented in PAJEK, using the “ separate components” parameter.*

The representation in Figure 8 shows a total of 40 cooperation clusters among authors. Among them, we can also observe three cases of single nodes, revealing no interaction at all within this network, as well as 7 weak clusters containing less than 5 connections each, where 4 of those 7 are composed of only two nodes. Apart from the big DSS publications cluster, the circular representations of separate components of the Publ_AP network shown above refer to other less populated clusters of joint-publications. In further reports of this project, we plan to identify and analyse all the separate components of this network representation, in order to get a precise picture of the EWG-DSS performance in relation to its joint-cooperation.

One of our main objectives with this study is to carefully study the single cases, specially the cases of the weak clusters, in order to try to find options for them to better interact. This way, we aim at promoting better and new cooperation among the EWG-DSS group members.

Network View of Publications with relation to their Topics (Publ_PT Graphs)

In this session we illustrate the representation of how the publication nodes are connected among themselves, with relation to their main topics of research.

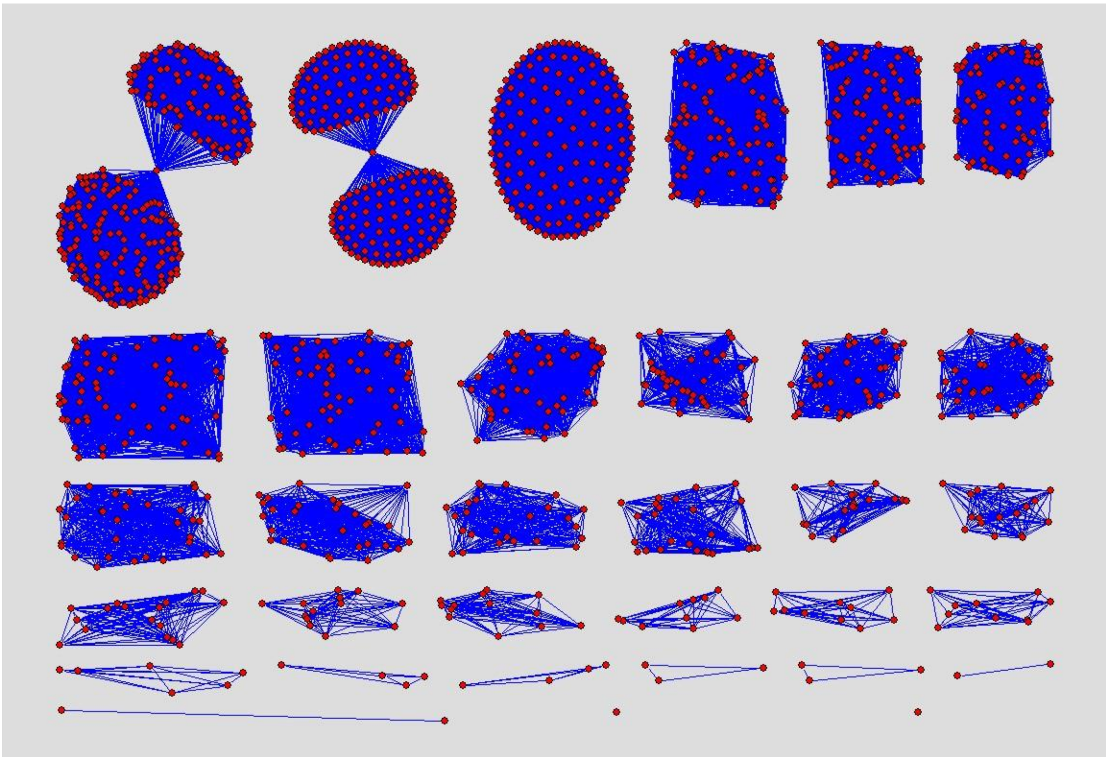

Figure 9: *Visualization of the Publ_PT represented in PAJEK, showing the publications connected to each other w.r.t. their research topi, using the parameter " separate components", where Pj = Publications of the authors; for j = {1,..., 1350} and Tn = Topics of the publications; for n = {1,..., 34}*

In Figure 9, we can identify the clusters of publications, relative to the topics listed in Table 3. In this visualization of the Publ_PT network, it is clearly seen that almost 25% of the topics, relative to 8 larger sub-nets, concentrate the great majority of papers published among the EWG-DSS group members.

The most used topics in publications, represented by the uppermost clusters of Figure 9, refer to the papers of the applications areas of "Decision Support Systems" (represented separately in a Column-Chart in Figure 10), "Operations Research", "Information and

Technology", "Expert Systems", "Knowledge Management" and "Multiple Criteria Decision Making", which seem to be the most exploited areas among the members of EWG-DSS, from the 34 topics listed in Table 3,. The other representations of separate components shown in Figure 9, refer to the smaller clusters of publications which relate to the other topics of interest within the EWG-DSS.

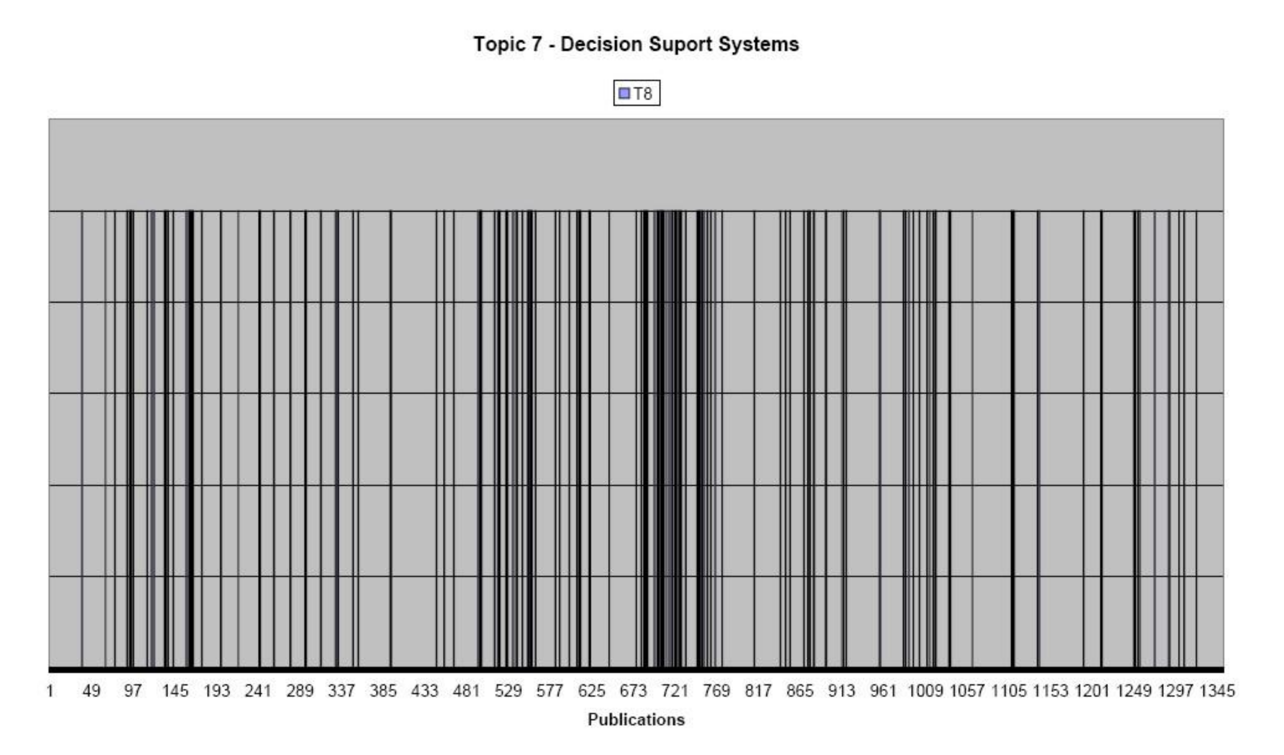

Figure 10: *Column-Chart representation of the Publications within the Topic T8 of DSS*

Figure 10 shows the amount of publications related to the topic T8 of "Decision Support Systems", within a Column-Chart. A full report of this project, containing the the specific analysis and the representation of all the publications relative to their main research topics is planned for publication soon.

Network Views of Authors with relation to their Topics of Research (Authors_AT and Topics_AT Graphs)

In the representation of Figure 11, a *Radial Graph* visualization of the network Authors_AT produced using the NWB framework, we can visualize the 70 authors considered and the way they are interconnected to each other with relation to their main topics of research, taking two arbitrarily authors as central nodes represented by A9 (Bertrand Mareschal) and A65 (Jason Papathanasiou). It is relevant to notice that the darker connections, represented in the foreground, express the stronger connections among the authors and the nodes in focus.

Using the NWB tool and the radial graph visualization option, it is very easy to dynamically change the nodes to be focused and in this way, analyze all the individual cases among all the authors, to get the exact picture of their cooperation relative to the many different topics within their areas of interest around Decision Making.

Figure 11 illustrates just one of the possible cases. Another possible representation of a single case analysis are shown in Figures 12 and 13, where the authors A70 (Antonio Rodrigues) and A49 (Pascale Zaraté) are placed respectively in focus as the central nodes.

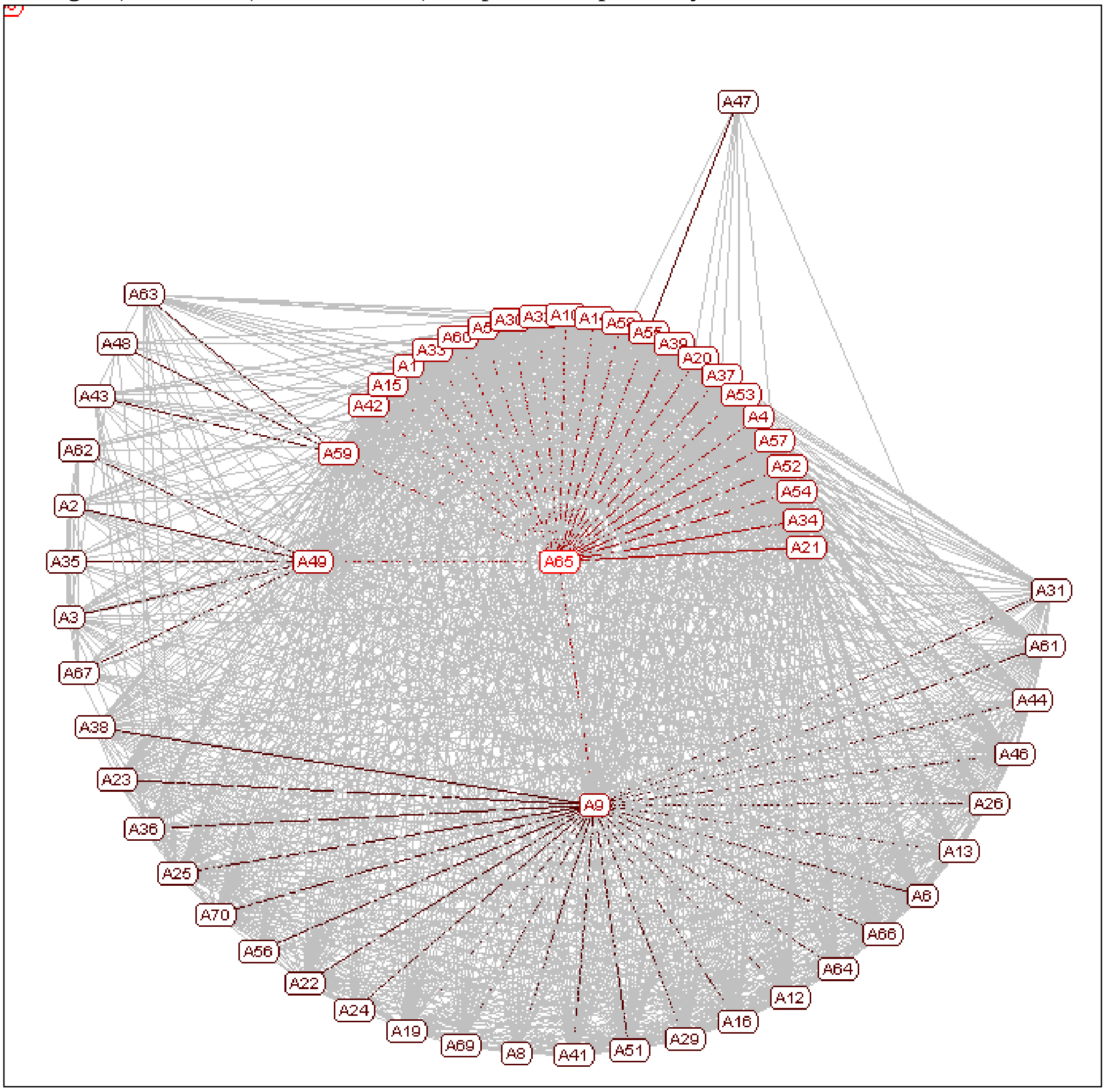

Figure 11: *NWB Radial Graph Visualization of the Authors_AT network, showing the collaboration among the authors, with relation to their common research topics.*

In Figure 12, it is easily identified that the author represented by the node A70 (Antonio Rodrigues) has direct links, with relation to areas of interests in publication topics, to 45 authors; indirect links to 13 other authors (who themselves are directly linked to some of the 45 authors of A70's inner circle); has even a more indirect link to 1 author (A62); and no links at all to 10 members of the EWG-DSS group. Those 10 isolated nodes represent members of group who have published either on their own within specific areas of research, or with other collaborators who do not belong to the EWG-DSS.

If we now analyze the case illustrated in Figure 12 statistically, taking into account our scenario of 70 authors, we can say that, throughout the period of time being considered, the EWG-DSS member A70 has cooperated strongly with 64% of the authors; in an indirect way with 18% of them; weakly with 1.4% of the authors; and had no joint-work at all with

14% of the considered group of authors of the group. In average, we conclude that member A70 has had a positive profile of interaction among the group.

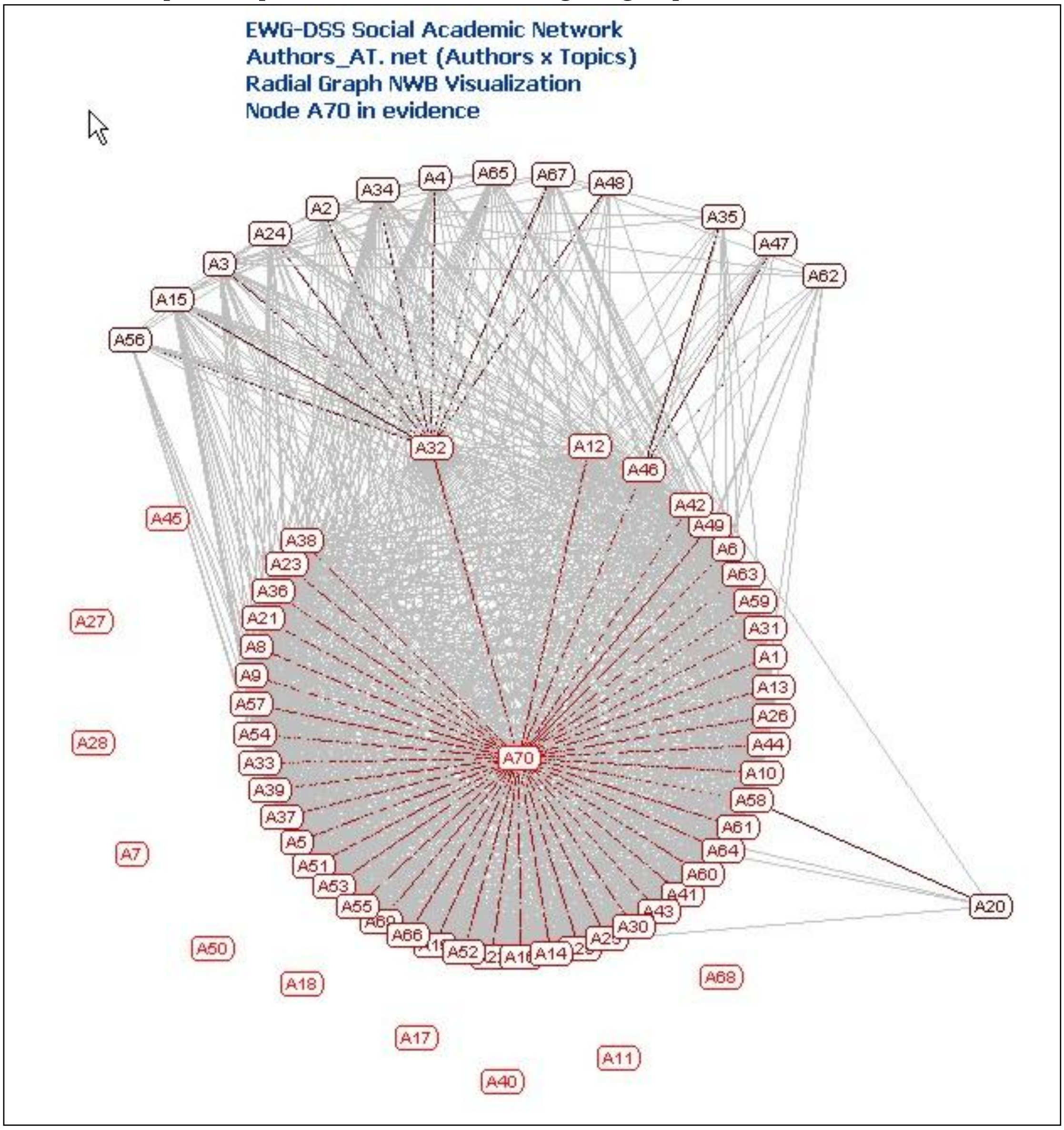

Figure 12: *NWB Radial Graph Visualization of the Authors_AT network, showing the author A70 as its focus node.*

The next Figure 13 shows that the author A49 (Pascale Zaraté) has direct links to 51 authors, with relation to areas of interests in publication topics, indirect links to 8 other authors; and as already shown for author A70, A49 has no links at all to those 10 isolated nodes that represent the members of group who do not yet interact within the group. Statistically analyzing the case shown in Figure 13, taking into account our scenario of 70 authors, we can say that, during the time being considered, A49 (Pascale Zaraté) has cooperated strongly with 73% of the authors and indirectly with 11% of them. A49 has no joint-work at all with 14% of the considered group of authors of the group. In conclusion, member A49 presents a very positive profile of interaction among the EWG-DSS group.

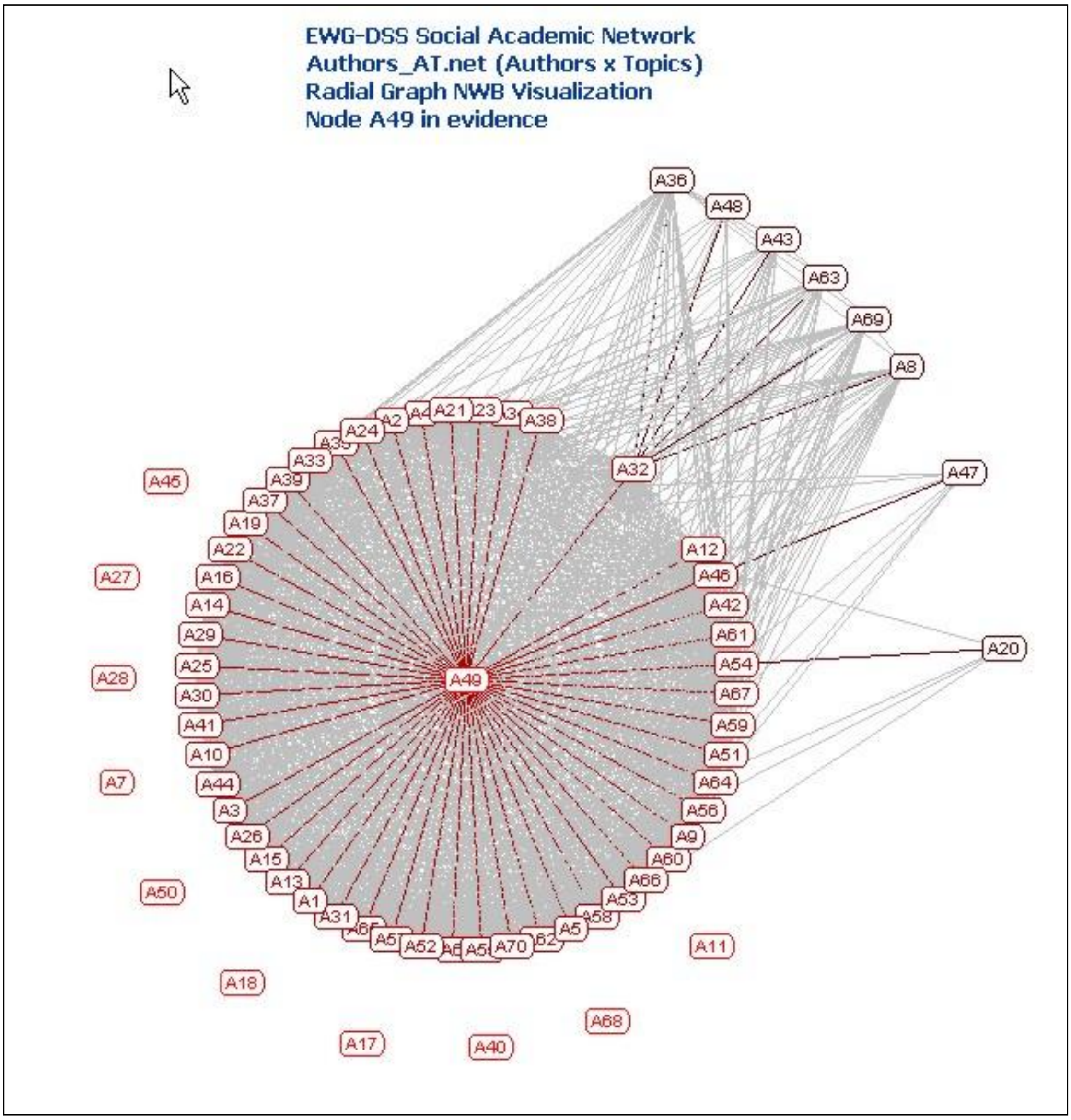

Figure 13: *NWB Radial Graph Visualization of the Authors_AT network, showing the author A49 as its focus node.*

Step 5: Network Statistical Analysis

In this section, we will focus on the metrics of some of the network representations shown previously in this paper. These metrics help us to build up a consistent basis for analyzing the network graphs that were generated via the input data available.

It is well known that the basic characteristics of a social network are described by the following parameters:

- Number of nodes;
- Number of isolated nodes;
- List of nodes attributes;
- Number of edges (connections);
- List of edges attributes;
- Density of the graph;
- Type of the graph (directed / not directed);
- Type of the connections of the graph (weakly connected or not);
- Number of weakly connected components (nodes); and
- Number of nodes in the largest connected components.

In the sequel of this section, we will make use of most of the parameters above to try to infer some relevant conclusions from our represented networks, considering at this stage the NWB representations and its Network Analysis Toolkit (NAT).

NWB Network Analysis Toolkit NAT

Using the Network Workbench Tool NWB, we could analyze the resulting academic network via its Network Analysis Toolkit: NAT.

NAT performs basic analysis on a represented network. It calculates the number of weak component clusters, strong component clusters, self-loops, parallel edges, whether the network appears to be directed or undirected, and the attributes present on both nodes and edges. It also calculates the number of nodes, the number of edges, and the density of a network.

By combining a lot of common analysis into a single algorithm, the NAT toolkit allows a user to get a good overview of the network and quickly discover any errors that may be present in the data [NWB-NAT].

For illustrating the NWB-NAT analysis carried out in this project, we have selected two obtained results, given by Figures 14 and Figure 15, where we publish the analysis of Authors_AP and of the Publ_AP network files, which mainly represent the EWG-DSS Academic Network.

AUTHORS_AP
Graph and Network Analysis

File: Authors_AP.net

This graph claims to be undirected.

Nodes: 70
Isolated nodes: 35
Node attributes present: label

Edges: 40
No self loops were discovered.
No parallel edges were discovered.
Edge attributes:
Did not detect any nonnumeric attributes
Numeric attributes:

	min	max	mean
weight	0.00328	1	0.08246

This network seems to be valued.

Average degree: 1.1428571428571428
This graph is not weakly connected.
There are 40 weakly connected components. (35 isolates)
The largest connected component consists of 25 nodes.
Did not calculate strong connectedness because this graph was not directed.

Density (disregarding weights): 0.01656
Additional Densities by Numeric Attribute
densities (weighted against standard max)
weight: 0.00137
densities (weighted against observed max)
weight: 0.00137

Figure 14: *NWB-NAT Network Analysis of Authors_AP (Authors x Publications).*

Figure 14 presents the results of the NWB-NAT analysis for the network Authors_AP, which consider the interconnections of its 70 author-nodes with respect to joint-publications. This analysis presents us with a scenario of the EWG-DSS academic interaction since 1989. The network includes 40 weakly connected components / nodes, where 35 isolated nodes can also be found among them. This result has allowed us to identify the authors / members of the group, who are still passive in mutual collaboration with other members. Moreover, on the basis of this somehow weak analysis, we have profited with the retrieval of the precise information to attack the interaction problem of the EWG-DSS group and try to improve the level of connectivity of this network for the future.

PUBL_AP
Graph and Network Analysis

File: Publ_AP.net

This graph claims to be undirected.

Nodes: 1350
Isolated nodes: 3
Node attributes present: label

Edges: 57475
No self loops were discovered.
No parallel edges were discovered.
Edge attributes:
Did not detect any nonnumeric attributes
Numeric attributes:

	min	max	mean
weight	0.2	1	0.85285

This network seems to be valued.

Average degree: 85.14814814814774
This graph is not weakly connected.
There are 31 weakly connected components. (3 isolates)
The largest connected component consists of 903 nodes.
Did not calculate strong connectedness because this graph was not directed.

Density (disregarding weights): 0.06312
Additional Densities by Numeric Attribute
densities (weighted against standard max)
weight: 0.05383
densities (weighted against observed max)
weight: 0.05383

Figure 15: *NWB-NAT Network Analysis of Publ_AP (Publications x Authors).*

Figures 15 presents the results of the NWB-NAT analysis for the Publ_AP network, which consider the interconnections of its 1350 publication-nodes with respect the involved authors. The scenario presented with the analysis of Publ_AP is somehow more positive than the one of Authors_AP. Again it involves the academic interaction of the EWG-DSS since 1989. From a total of 1350 nodes, the network includes only 31 weakly connected components, where a minor quantity of 3 isolated nodes were found among them. This positive result has shown us that even the 40 weakly connected authors of the previous analysis are producing a reasonable rate of research publication. However, they are not yet in collaboration with other members of the group. Once more, on the basis of this analysis, we have now some relevant and accurate information at hand to better handle and encourage joint-collaboration within the EWG-DSS group.

Also in relation to the publications interconnections of Publ_AP, some other relevant statistical parameters were derived from the network files, generating the following results:

- Diameter: 7
- Average path length: 3.1941558
- The average clustering coefficient is 0.9684615657261947
- Network Betweenness Centralization = 0.16360

The ***Diameter*** parameter represents the maximum distance between the publications found in the network. The value “Diameter: 7” for the Publ_AP network means that it is a well connected network, since there are a lot of publications and none of them is located at a distance greater than 7.

The ***Average path length*** parameter classifies the type of network. In our case the aproximate value of 3.2 shows that the Publ_AP network is viewed as a ***small-world***, due to its number of nodes Pj = Publications of the authors; for j = {1,…, 1350}, even though for us 1350 publications were shown to be difficult to visually process.

The ***Average clustering coefficent*** parameter provides a measurement for transitivity, i.e. when it is high, like in our case: 0.968, it indicates that communities exist within the network. This characteristic was also easily observed in the visualization of the Publ_AP network, illustrated in Figure 8 considering the option of “separate components”.

The ***Network Betweenness Centralization*** also called the ***Centrality*** paramter correlates the degree of the nodes of the network. For example, a node's betweenness centrality indicates the number of shortest paths going through it. This notion tend to be more important in scale-free networks, where the disproportionately large number of connections carried by hubs make them lie in the path between many other nodes. In the case of the Publ_AP network, the centrality parameter was shown to be reasonably good with the value of 0.16360.

All the NWB-NAT analysis with relation to the complete set of generated network files that fully represent the EWG-DSS Academic Network, can be obtained in the final report of this project, to be published soon as a technical report of IRIT.

Concluding Remarks

In the current project, we have worked on the development of a network structure for the EURO Working Group on DSS, using 70 of its current 104 group members as its defined population of authors, together with their relationships of research collaborations and main topics of research, within the areas around Decision Making.

By identifying the collaboration relationship that exists among the EWG-DSS members, we were able with this project to show: 1. how the members of the group relate to each other, in terms of topics of research; 2. what are the most relevant topics of research within in the group as a whole; and 3. what are the relevant statistical data/knowledge concerning the amount of publications that circulate within the group, how the authors are tending to cooperate with each other and also how many external collaborators (co-authors not in EWG-DSS) exist and indirectly participate in the EWG-DSS.

It was not a central objective of this paper to identify each one of the network clusters and analyze its population density. But it is certainly the aim of this project as a whole to perform a detailed analysis of each represented cluster of publications, authors and topics in a constructive way. By doing this we hope to be able to spot and try to eliminate the isolated cases in the EWG-DSS network, whenever possible, by showing to the members involved, other possible options available for cooperation within correlated research areas.

Based on the study conducted by Granovetter in the earlier developments of social networks [Granovetter, 1973], we believe that there is great potential for the weak connections of our EWG-DSS network still to be able to develop into strong ones. In Granovetter's study, he has explained that information was far more likely to be "diffused" through weaker ties, than through already strong connections. We shall, however, be concerned with the absent ties, i.e. connections that are beyond the concept of weak ties. Although those connections are of no relevance to the network as a whole, they should be encouraged to become "weak ties", in order to interact and gain importance within the network.

In the particular case of the EWG-DSS network, absent connections are represented within the network files, by for example authors who exist in the network, are publishing actively within their areas of research, also with collaborators outside the scope of the group, nevertheless do not interact with other members of the EWG-DSS group, via joint collaborative research work. This issue was little exploited in this paper, but it is part of the project's main objectives and will be worked out in the future.

On the way to accomplish the main aim of our EWG-DSS social-academic network analysis, we have already covered most of the planned work. However, for performing the full analysis of the network, as well as for optimizing its relationships adding also missing input data, our project still has some steps to cover, which we plan to fulfil within the next six months.

We are certain that with this work, we have established a powerful tool to identify the EWG-DSS group interaction on one hand, and also to provide a useful feedback for further collaboration in joint-research on the other. We hope and expect that this tool gives us enough grounds for extinguishing the existing "absent connections" among the members of the EWG-DSS network, by enabling us to identify them on one hand and encourage them to take part in new research and academic cooperation within the group, on the other.

As coordinators of the EURO Working Group on DSS, we also encourage the other existing Working Groups of EURO to follow our example and try to identify the cooperation of their groups in order to encourage more participation and research production in joint-work.

Acknowledgements

The authors are grateful to all the members of the EWG-DSS working group, for their important participation in this project. This work would not be possible to be carried out without the cooperation of all the EWG-DSS members, who have replied to us.

Many thanks also for the support of Rahma Bouaziz and Tiago Simas in the implementation of the methodological approach and analysis of this work. Their work was of vital importance for starting our implementation and for achieving the analysis stage that we have, so far.

References

[Granovetter, 1973] Granovetter, M. S. “The Strength of Weak Ties”. American Journal of Psychlogy, 78(6), pp. 1360-1380, 1973

[Newman 2003] Newman, M. "The Structure and Function of Complex Networks". SIAM Review Vol 45, No 2,pp 167–256, 2003.

[Newman 2001a] Newman, M. “Scientific collaboration networks. I. Network construction and fundamental results”. Physical Review E, Vol. 64, 016131, 2001.

[Newman 2001b] Newman, M. “Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality”. Physical Review E, Vol. 64, 016132, 2001.

[Miyamoto 1990] Miyamoto, S., “Fuzzy Sets in Information Retrieval and Clustering Analysis”. 1990:Kluwer Academic Publishers.

[Weimao, Simas 2007] Weimao Ke, Tiago Simas, “Mapping a Local Web Domain”. 11th International Conference on Information Visualisation, IV 2007, 2-6 July 2007, Zürich, Switzerland. IEEE Computer Society 2007: 228-231.

[Simas 2008] Simas, T., “How to Build Weighted Graphs from Bi-partite Graphs”. Technical Report, UNINOVA, 2008.

[Freeman 2000] L. Freeman (2000) Visualizing Social Networks. *Journal of Social Structure*(JoSS) Vol. 1, No. 1.

[Rocha et al. 2005] Rocha, L.M., T. Simas, A. Rechtsteiner, M. DiGiacomo, R. Luce. MyLibrary@LANL: Proximity and Semi-metric Networks for a Collaborative and Recommender Web Service". In: Proc. 2005 IEEE/WIC/ ACM International Conference on Web Intelligence (WI'05), IIEEE Press, pp. 565-571, 2005.

[Verspoor et al. 2005] Verspoor, K., J. Cohn, C. Joslyn, S. Mniszewski, A. Rechtsteiner, L.M. Rocha, T. Simas. "Protein Annotation as Term Categorization in the Gene Ontology using Word Proximity Networks". BMC Bioinformatics, 6(Suppl 1):S20. doi:10.1186/1471-2105-6-S1-S20, 2005.

[PAJEK] PAJEK Website: http://vlado.fmf.uni-lj.si/pub/networks/pajek/

[NWB] NWB Website: http://nwb.slis.indiana.edu/