Improving Learning Performance in Industrial E-Learning Settings by Utilizing Social Networks

Seid Maglajlic, Dipl.-Ing.

DISSERTATION

zur Erlangung des akademischen Grades eines Doktors der technischen Wissenschaften
der Studienrichtung Informatik erreicht an der Technischen Universität Graz

Univ.-Doz. Dipl.-Ing. Dr.techn. Denis Helic

Institut für Wissensmanagement

Technische Universität Graz

Graz 2013

Deutsche Fassung:
Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008
Genehmigung des Senates am 1.12.2008

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am ....

(Unterschrift)

Englische Fassung:

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

....

date (signature)

ABSTRACT

Online social networks appear lately in various contexts, largely due to the popularity of the tools such as Facebook or LinkedIn. The traditional social networks and their relations to various human activities have been investigated for long time already, long before the Web 2.0 revolution. The methodology called Social Network Analysis has been developed for scientific investigation of such relations. In this work we applied the Social Network Analysis methods to investigate the role of online as well as traditional social networks in an industrial E-Learning setting.

For this thesis, we executed several E-Learning experiments in a pan-European organization with member companies that share cooperative business processes. In the next step, we applied Social Network Analysis to analyze the experimental results. One of the most important results was the fact that the learning performance in an industrial E-Learning setting was extremely weak whenever communication between trainees and tutors was also weak.

For that reason, in the second part of the thesis work, we provide a framework for the construction and use of collaborative online social networks within an E-Learning environment that aim to improve the learning outcome by improving the communication channels in the system. To that end, we apply further methodologies such as controllability theory and ontology engineering. We develop a novel algorithm for matching trainees and tutors in an E-Learning system to create groups that facilitate a better communication.

Finally, we evaluate the improvements of the developed framework by standard statistical tests such as sample t and chi-square tests. We are able to significantly improve learning performance in specific configurations of our framework.

Summarizing, this thesis contributes a novel matching algorithm for grouping of trainees and tutors in an E-Learning system, as well as an example of a practical integration of social networks and ontology engineering. This work is interesting for researchers since it provides an analysis of a specific situation in an industrial setting, which differs largely from the investigations on the same topic in academic settings. Moreover, this work is interesting for system designers in the industry since it provides the advanced methods of enhancement of E-Learning applications and tools.

For Bensara and Orhan

ACKNOWLEDGEMENTS

I would like to sincerely thank to my PhD supervisor, Denis Helic for his guidance and for helping me to put my "fuzzy" ideas in a form useful for the community.

This research has been carried out in cooperation between RailNetEurope (RNE) and TU Graz. I would like to thank to the students of TU Graz, Rene Winkler, Christoph Oberhofer, Christian Slamanig and Mario Ouchan who helped me to implement my ideas. They significantly contributed to the evaluation, installation, administration and development of add-ons of the E-Learning system of RNE.

I would like to thank to the Secretary General of RNE, Joachim Kroll, who supported and encouraged the cooperation with TU Graz and realized the benefit of the new methods for distance learning and knowledge management in industry.

Finally, I want to thank to my beloved family, my wife Bensara and my son Orhan who gave me the strength to carry on.

1. Motivation and Objectives

1.1. Background information

The author of this work works for the railway industry and was inspired by the implementation of an E-Learning system in this industrial environment. The railway industry is traditionally very resistant to change, therefore it was a considerable challenge to implement a new distance–learning, knowledge-sharing methodology in the chosen setting, as well as to observe and analyze the behavior of the community formed by the E-Learning participants. This setup significantly differs from the academic setup where students of a school or university are supposed to use a distance-learning methodology thus the issues related to knowledge advancement of E-Learning participants in an *industrial* setting are the focus of this work.

1.2. Introduction

E-Learning is a popular distance-learning method in diverse organizations, both academic and industrial ones. In academic organizations it is quite normal nowadays to use E-Learning as an extended mechanism of knowledge management and delivery in addition to the traditional educational methods. Setting up E-Learning methodologies in industry is also an ongoing trend but, of course, a little slower than in the education sector. However, the management of companies in the industry sector is increasingly aware of the importance of knowledge management and knowledge sharing between co-workers within their companies, especially after the boom enjoyed by the Web 2.0 and its growing utilization in corporations (Schachner and Tochtermann, 2008, Littlejohn et al., 2011).

During the literature review we found two guiding principles for the application of E-Learning in industrial organizations. The first is that these organizations aim to reduce workforce training costs compared to face-to-face training and the second that they aim to validate and improve the efficiency of knowledge transfer, again in comparison with traditional training methods (Lundvall, 2008, Sellen, 2002, Karrer, 2008). The obvious benefit brought by the first principle is that the companies reduce costs by lowering the travel expenses and being able to keep the workforce at the workplace even during the training.

However, the second principle is much more complicated to achieve than it might appear at first. Thus it is not surprising that efficiency has already been investigated in relation to E-Learning with regards to the following aspects: cognitive load theory (Paas, 2003, Clark et al., 2006, Plass, 2010), rapid dynamic assessment (Kaluga and Sweller, 2005), or adaptive and collaborative learning (Ruiz et al., 2006, Kumar, 2007, Weibelzahl, 2008), to mention just a few. In this work, we assume that the efficiency of knowledge transfer is reflected in the success rate of the learners, i.e. if knowledge transfer from the source to the trainees was efficient, the trainees will successfully reach their learning objectives.

Our aim is to investigate how to increase learning efficiency, i.e. how to improve the knowledge of the learners, in other words ensure their knowledge advancement. We took the research findings provided in (Scardamalia 2002, Trausan-Matu et al. 2012) as guidance; these indicate that knowledge advancement can best be achieved as a community rather than as an individual, therefore collaborative learning within an E-Learning setting is interesting to investigate. The investigation of the involvement of social networks in E-Learning also seemed to be a natural step, in order to analyze the communities and their collaboration activities within the E-Learning settings. Social networks in E-Learning have been a topic of research for years already, and some useful findings could be deployed in this work. Let us mention a few.

Implicit Social Networks

The implicit involvement of social networks in E-Learning setups has been seen as very intuitive (Haythornthwaite, 2005). The author of the referenced work provides a social network model of E-Learning participants for the purpose of examination, defines the roles and relations between the E-Learning participants (the actors in the social network), proposes measures using social network analysis (SNA) methods, and concludes by arguing for community building within E-Learning settings.

Social Network Modeling

Applying a social network model to the structuring of users in E-Learning systems such as the one made by (Chatti et al., 2007) clearly shows that taking social network methodologies into consideration can hardly be avoided in future E-Learning setups. The authors of the above-mentioned work claim that the social communication component (i.e. the paradigm ‘who do you know’) is becoming more important than sophisticated access methods to the learning material. The authors argue that previous attempts to relate Knowledge Management (KM) and Learning Management (LM) failed because they did not consider the social communication component as important.

Thirdly, the enhancement of collaborative learning by distance learners through the application of social network-based community-building according to behavioral rules is described in (Wang et al. 2007) as an example of the intensive utilization of the results of social network analysis for E-Learning.

Collaboration

The newest research attempts such as those shown in (Silva and Figuera 2012) support the idea of applying SNA (social network analysis) methodology to collaboration between trainees and tutors. The authors show that the social component of E-Learning can be closely investigated by visualizing the activities of the participants in a discussion forum in the chosen E-Learning setup.

However, in spite of these very useful findings, there was a lack of results about the impact of these collaborative social network-based methods, as applied in E-Learning settings, on trainee's learning success. The majority of the research work we examined contained advice about the structuring of the E-Learning community, but did not contain any indication about the knowledge advancement of the learners after the above-mentioned methods had been applied (Griol et al. 2012, Ullrich et al., 2010). **Although SNA is very important, only a few studies have investigated the impact of social networks on the learning success** (Cho et al., 2007, Vaquero and Cebrian, 2013)**.**

The work of (Hwang et al. 2012) sheds more light in this direction. Its experimental results show that the cognitive load of the learners can be reduced and their learning outcome can be improved if computerized collaborative concepts are implemented in the community of the learners. The experiments were made in an academic environment, with a sample of children aged 11 on average.

However, the question is how these findings relate to our sample in an industrial setting since it differs heavily from the academic one. Our samples are taken from a pan-European organization with member companies all over Europe that share a common business process for cooperation. The profile of our E-Learning participants largely differs from the profile of users in typical academic environments such as a school or a university: they are very likely to differ in age, as well as in educational, business, geographical, cultural and language background and skills.

To bridge the gap that we had detected, we started our own research and investigated the potential of E-Learning systems for the knowledge advancement of industrial learners. We chose an E-Learning system which initially suited the purpose and more or less reflected the standard of today's E-Learning systems: structuring of educational material, easy authoring of tests, testing reports, tutoring, virtual classrooms, discussion forums, chat, just to mention a few of their characteristics. For this purpose, several contemporary E-Learning systems were evaluated and the best-matching license-free (open source) system was chosen. The details about the system can be found in (Maglajlic and Helic, 2010). The users of the E-Learning system are employees of several companies geographically distributed across Europe. The courses in the E-Learning system contain lessons about new business processes that the workers in those companies are supposed to participate in, and about completely new workflow support tools that the companies within the organization have agreed to use, but the employees are not used to them.

The initial idea was to provide an approach to context-sensitive learning material delivery depending on the skills and competences of the E-Learning system users. However, during the observation of the knowledge advancement of the trainees – especially in relation to the efficiency of the knowledge transfer between tutors and trainees – we noted some problems that forced us to investigate the reasons more deeply and to come up with novel approaches to their solution.

1.3. Statement of the problem

We observed the learning outcome improvement of the trainees in the E-Learning system by comparing their test results before taking the E-Learning course, and after taking the (tutored) E-Learning course. We noted that trainees who had weak test results at the beginning were likely to have weak test results after the course if they didn't communicate with their tutors. More precisely, the lack of collaboration implied a weak learning outcome and was a sign of inefficiency in the knowledge transfer. One might claim that this is an obvious issue and that it is not worth investigating any further, but we disagree for the following reasons:

a) An analytical approach with a precise result evaluation that could show the statistical dependencies between communication intensity and learning outcome could rarely be found in existing research. An example can be found in (Chen et al., 2011), where the authors provide a very precise regression model which indicates that direct dialogue between student and teacher does influence the learning outcome. However, this research was carried out in a typical academic environment (a university), not in an industrial environment, and it was focused on face-to-face learning, not on E-Learning.

b) Teaching E-Learning participants is a new challenge compared to teaching students in an academic environment. Discussion between students and teacher in the (virtual) classroom is quite a natural thing, as well as the utilization of discussion results leading to improved learning. However, in an industrial setting, collaboration between E-Learning participants cannot be taken for granted since the participants may differ widely in their skills and competences, or in their working experiences. This might affect their personal attitude as regards asking the tutors for help. Hence, it is worth further analyzing whether their organizational, business sector-related, geographical or language-related background play a role in their learning process.

The problem statement is underpinned with the experimental results given in Paper 1 of Chapter 3.

1.4. Objective

According to the problem specification set out above, the main objective of this research is to discover methods that can enhance collaboration between trainees and tutors (actually, all E-Learning participants) within a specific industrial setting in order to promote their knowledge improvement.

In order to meet this objective, we consider that the following research questions have to be answered. Each of these questions has its own objective, but answering them helps to cover the general objective indicated above.

1.5. RESEARCH QUESTIONS

RQ1: What are the structure and properties of implicit social networks in E-Learning?

The objective of this research question is to analyze the E-Learning setting in depth and try to identify the implicit social networks in it. Answering this question helps us to analyze the eventual correlation between the social network role and/or position of the E-Learning participants and their knowledge advancement. We deal with this research question in Paper 1: 'How Do Social Networks Influence Learning Outcomes? A Case Study in an Industrial Setting'.

RQ2: How can we control and adapt social networks in the E-Learning setting to enhance collaboration?

The objective of this research question is to find ways to overcome the differences that constitute obstacles in the communication and collaboration of the E-Learning participants. Answering this question helps us to find out if there are any possibilities not only to use the potential implicit social networks but also to construct new social networks of trainees and tutors to enhance their information and knowledge interchange.

We deal with this research question in Papers 2 'Engineering Social Networks Using the Controllability Approach Applied to E-Learning', Paper 3 'Efficiency in E-Learning: Can Learning Outcomes Be Improved by Using Social Networks of Trainees and Tutors?' and Paper 4 'Implementation of a Framework for Collaborative Social Networks in E-Learning'..

RQ3: How adaptations to social networks within E-Learning settings improve collaboration and learning outcomes?

The objective of this research question is to investigate and compare the knowledge advancement of the trainees *with* and *without* the utilization of social networks for collaboration enhancement. Actually, by answering this research question, we also tackle the typical industrial, economic-efficiency question: has the workforce's knowledge improved by using this particular E-Learning system, i.e. is there a benefit of the particular utilization? We deal with this research question (and its 'industrial/economic' sub-question) in Paper 3 ‘Efficiency in E-Learning: Can Learning Outcomes Be Improved by Using Social Networks of Trainees and Tutors?’ and Paper 4 ‘Implementation of a Framework for Collaborative Social Networks in E-Learning’.

1.6. RESEARCH RESULTS

In order to answer the three research questions listed above, we needed to apply various methodologies, such as SNA (Wasserman and Faust, 2009, Knoke et al., 2008, Scott, 1988), the newest experiences gained from controllability theory (Sontag 1998, Åström 2008, Liu 2011, Kalman, 1963), as well as ontology engineering (Mika 2005, Pernas et al. 2012), and the methods of behavioral statistics, e.g. sample *t* tests, chi-square tests (Cohen, 1998, Myers, 1990, Mann, 1942) and Pearson-Product-Moment-Correlation-Coefficient (PPMCC), (Rogers, 1988). The work that was carried out with the help of these methodologies resulted in a framework for collaboration enhancement in E-Learning environments. The characteristics of this framework that make it original in comparison with previous research in this field are as follows.

NOVEL MATCHING ALGORITHM

It specifies a novel algorithm that matches tutors and trainees for the construction of tutored groups by taking the social network parameters of both trainees and tutors into account when calculating the matching criteria.

PRACTICAL INTEGRATION OF SOCIAL NETWORK ANALYSIS AND ONTOLOGY ENGINEERING

It provides a method for the joint utilization of learners' and learning ontology and SNA methods; in this way, the communication intensity between trainees and tutors within the E-Learning setting can be evaluated for the purpose of recommending tutors and fellow trainees to the 'weaker' trainees – the aim being to help them improve their learning outcome.

As a 'proof-of-concept', the framework was implemented as an add-on to the chosen E-Learning system in the chosen industrial environment. The detailed description of the framework and its results can be found in Paper 4 as well as in Chapter 4 ('Final Evaluation'), but we will describe the structure of the thesis in more detail in the next section.

1.7. ORGANIZATION OF THE THESIS

The rest of the thesis is organized as follows. Chapter 2 provides an overview of the scientific methods used in this research (as mentioned in the previous section) as well as a brief description of their theoretical background. Chapter 3 presents the cumulative work and contains the following papers:

Paper 1: Maglajlic, S., Helic, D. (2011), How do social networks influence learning outcomes? A case study in an industrial setting, Emerald Journal on Interactive Technology and Smart Education, Vol. 9 Iss: 2, pp.74-88.

Paper 2: Maglajlic, S. (2012), Engineering Social Networks Using the Controllability Approach Applied to E-Learning, Proceedings of 12th IEEE International Conference on Advanced Learning Technologies, pp. 276-280, DOI 10.1109/ICALT.2012.209.

Paper 3: Maglajlic, S. (2012), Efficiency in E-Learning: Can Learning Outcomes Be Improved by Using Social Networks of Trainees and Tutors?, Addleton Academic Publishers Journal on Economics, Management, and Financial Markets, Volume 7(4), 2012, , pp. 121-137, ISSN 1842-3191.

Paper 4: Maglajlic, S. (2013), Implementation of a Framework for Collaborative Social Networks in E-Learning, AACE International Journal on E-Learning (IJEL) Corporate, Government, Healthcare, & Higher Education, in review.

The graphical representation of Chapter 3 helps to better understand how the three research questions have been answered in the four papers in order to reach the final objective.

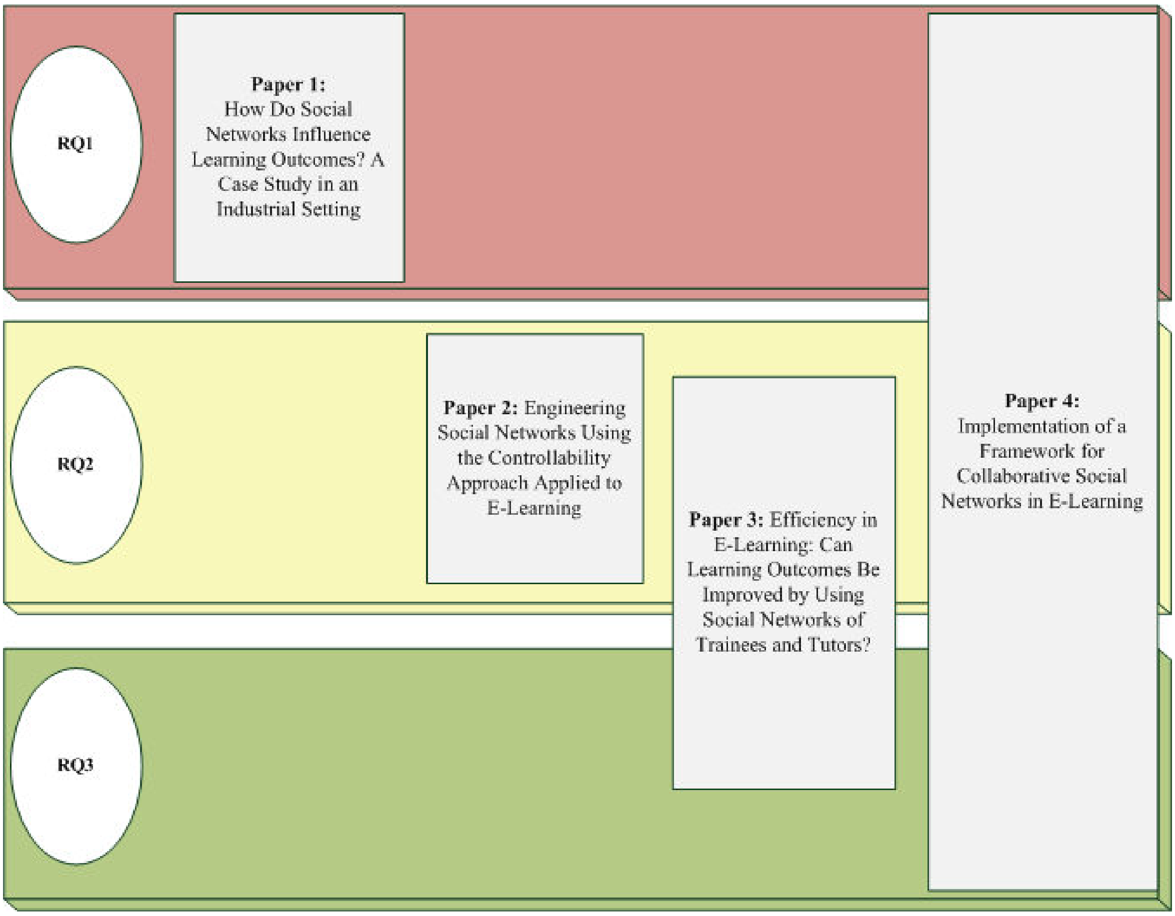

Figure 1: Coverage of research questions by papers in the thesis

Chapter 4 contains a final statistical evaluation of the methods applied in this research. We conclude the work with Chapter 5.

1.8. ALL PUBLICATIONS

During the research work, the following publications have been made:

Maglajlic, S., Helic, D. (2010), Integrating E-learning into work processes in industrial settings: a case study, IEEE Proceedings of the 9th International Conference on Information Technology based Higher Education and Training (ITHET), Cappadocia, Turkey, pp.151-157, ISBN 978-1-4244-4810-4.

Maglajlic, S., Helic, D., Trattner, C. (2010), Social Networks and eLearning: New Model for Learning at Workplace, IEEE Conference on Information Technology Interfaces, ITI 2010, Cavtat / Dubrovnik, Croatia, pp. 373-378.

Trattner C., Helic D., Maglajlic S.: Enriching Tagging Systems with Google Query Tags, IEEE Conference on Information Technology Interfaces, ITI 2010, pp. 205-210, Cavtat, Croatia, June 2010.

Maglajlic, S. (2011), On the Importance of the Impact Analysis of Social Network Methods in Elearning, Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. Chesapeake, VA: AACE, pp. 749–752. (Award winning publication: "Outstanding Virtual Presentation Award" on AACE E-Learn 2011)

Maglajlic, S., Helic, D. (2011), How do social networks influence learning outcomes? A case study in an industrial setting, IADIS International Conference WWW/Internet 2011, Proceedings, pp. 203-213. ISBN 978-989-8533-01-2.

Maglajlic, S., Helic, D. (2011), How do social networks influence learning outcomes? A case study in an industrial setting, Emerald Journal on Interactive Technology and Smart Education (ITSE), Vol. 9 Iss: 2, pp.74–88.

Maglajlic, S. (2012), Engineering Social Networks Using the Controllability Approach Applied to E-Learning, Proceedings of 12th IEEE International Conference on Advanced Learning Technologies, pp. 276-280, DOI 10.1109/ICALT.2012.209.

Maglajlic, S., Gütl, C. (2012), Efficiency in E-Learning: Can Learning Outcomes Be Improved by Using Social Networks of Trainees and Tutors?, Proceedings of International Conference on Interactive Collaborative Learning (IEEE ICL), Villach, Austria, 2012; pp. 1-8., DOI 10.1109/ICL.2012.6402088.

Maglajlic, S. (2012), Efficiency in E-Learning: Can Learning Outcomes Be Improved by Using Social Networks of Trainees and Tutors? Addleton Academic Publishers Economics, Management, and Financial Markets, New York, Volume 7(4), 2012, pp. 121-137, ISSN 1842-3191

Maglajlic, S. (2012), Social Network Engineering and Ontology Engineering For E-Learning: How Do These Work Together? IADIS International Conference WWW/Internet 2012, Proceedings, in printing.

Maglajlic, S. (2013), Implementation of a Framework for Collaborative Social Networks in E-Learning, AACE International Journal on E-Learning (IJEL) Corporate, Government, Healthcare, & Higher Education, in review.

2. Theoretical Background

In this chapter we will explain the theoretical background of the scientific methods used in this piece of research, as well as present related research in the domain where these methods have been applied. The following methodologies have been used:

- Social Network Analysis (SNA)
- Graph Theory
- Control Theory
- Ontology
- Statistics.

A full description of these methods is clearly out of the scope of this thesis; however, basic definitions and an explanation of the interpretation of these methods will make it easier for the reader to understand the papers in the third chapter, and thus get the whole picture of the results achieved. Firstly we will describe the SNA and its relationship with graph theory, since these two approaches are very closely linked. Secondly, we will briefly describe the purpose of control theory, and note which specific relationship between control theory and graph theory is used in this work. Thirdly, we will shortly mention the principle of ontology and explain why this methodology has been used in this work. Finally, we will mention which statistical methods are used in this research and how they have been interpreted.

2.1. Social Network Analysis and Graph Theory

The concepts of social network analysis are used in this research by relying on the rules provided in (Wasserman and Faust, 2009). In reference to this source, we use the following definitions:

- *Actors* correspond to the basic unit of the social network. In our case these are the registered E-Learning participants.
- Relational *tie* – the link between actors, such as the evaluation of one person by another through friendship, transfer of material resources (buying/selling), transfer of non-material resources (communications; information sending/receiving). Thus, ties correspond to the affiliation of actors to various events or organizations or to direct communication between actors.
- *Group* – consists of a finite set of actors, which for conceptual, theoretical or empirical reasons, is treated as a finite set of individuals on which network measurements are made.

- *Relation* – a collection of ties of a specific kind among members of a group. For example, a set of friendships in a team, or a set of partnerships between companies in the industrial branch under study can be seen as relations between individuals or groups.
- *Social network* – the final set of actors and relation(s) detected within the network.

It is very convenient to represent the social networks with the help of graphs. For this purpose, mathematical graph theory is used. A graph may be defined as follows: *graph* G is an ordered couple consisting of the non-empty set of units called *nodes* (often referred to in the mathematical literature as *vertices* (sing. *vertex*), hence the notation *V(G)* for a set of nodes can be found in the literature) and a set of links between the nodes called *edges* (in the notation often referred as *E(G)*). The function that maps each edge to the non-ordered couple of (not necessary different) nodes, is called *incidence function* (Bondy and Murty, 2008). The 'relation' within the social network definition actually corresponds to the *incidence function* given by the definition of a graph. The 'relational ties' are interpreted as edges. 'Actors' and affiliation targets are (usually) represented by nodes. In the literature, diverse explanations about various types of social networks can be found, but we will concentrate on the following two.

Affiliation network

By definition (Wasserman and Faust 2009), an affiliation network is a type of social network that is shaped by the actors, organizations or events according to the actors' membership of an organization or registration for an event (affiliation). In our case, as already mentioned above, the actors are the E-Learning participants; the organization to which the participants belong, their business branch, geographical location and language preference can be considered as 'target of affiliation'.

Affiliation networks are usually modeled as bipartite graphs. A graph is called *bipartite* if its nodes can be separated into two disjoint sets *X* and *Y* (subsets of *V(G)*) and all the nodes from *X* have their edge ending in the *Y* and vice versa. Quite clearly, the nodes representing the actors can be put into the set *X* and the affiliation targets can be put into the set *Y*. The affiliation to the particular target is the edge between the actor and the target.

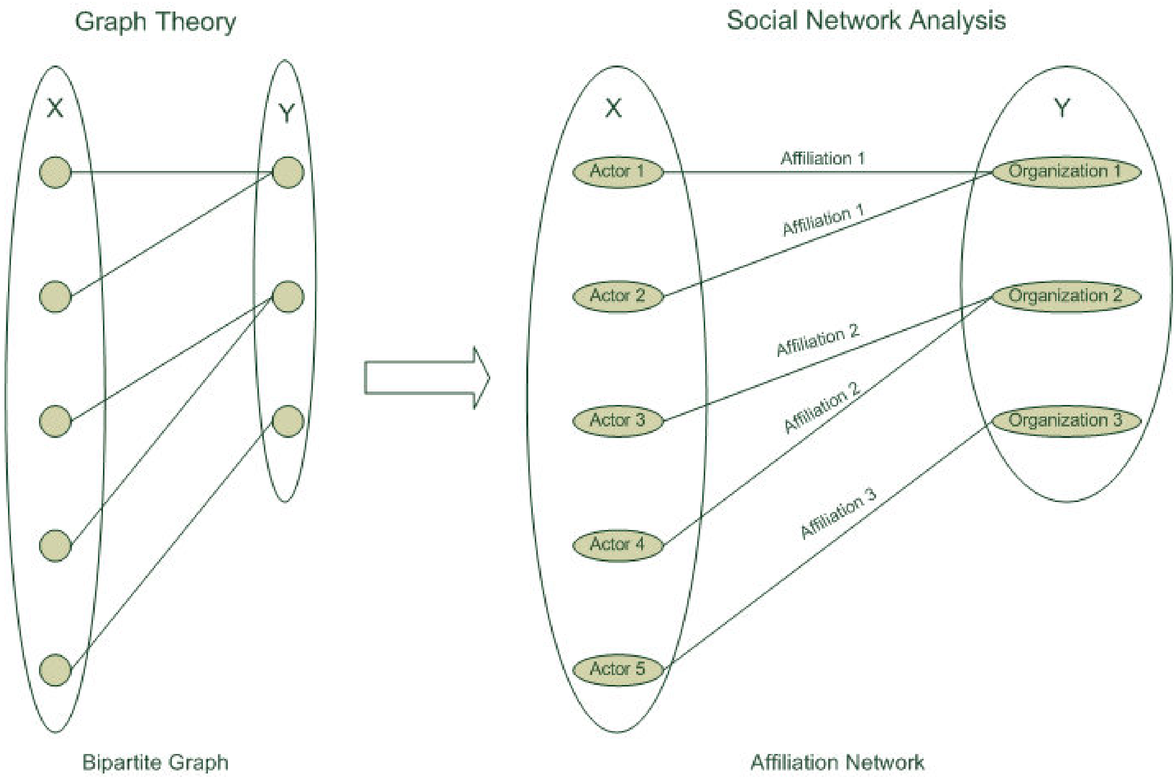

Figure 2: Bipartite graph and affiliation network: interpretation of graph theory in SNA

A social network with only one set of actors is called one-modal. Social networks where the set of basic units (actors and affiliation targets) can be divided into several (at least two) groups are called multi-modal. Affiliation networks are therefore multi-modal. Multimodal networks are represented with *n*-partite graphs, where *n* is the count of modalities. For simplicity reasons, we used two-modal affiliation networks, and present them as bipartite graphs. For example, affiliation networks according to organization and business branch will be treated separately.

COMMUNICATION NETWORK

Communication networks are modeled in a more straightforward manner than affiliation networks, and this task is intuitively easier, since we do not need to separate the sets of actors and affiliation targets – from the beginning, we try to form a network with actors of nearly the same type. The information flow between actors forms the communication network. In our case, E-Learning participants who exchange e-mails/messages or take part in a discussion forum are modeled in the graph as nodes connected by the edges, corresponding to the information flow. In graph theory, directed graphs are defined as graphs where the edges between the nodes are directed. The information flow direction defines the edge orientation in the graph. Such a graph is often called *digraph* in the literature (Bondy and Murty 2008; Wasserman and Faust 2009). The edges in a digraph are called *arcs*.

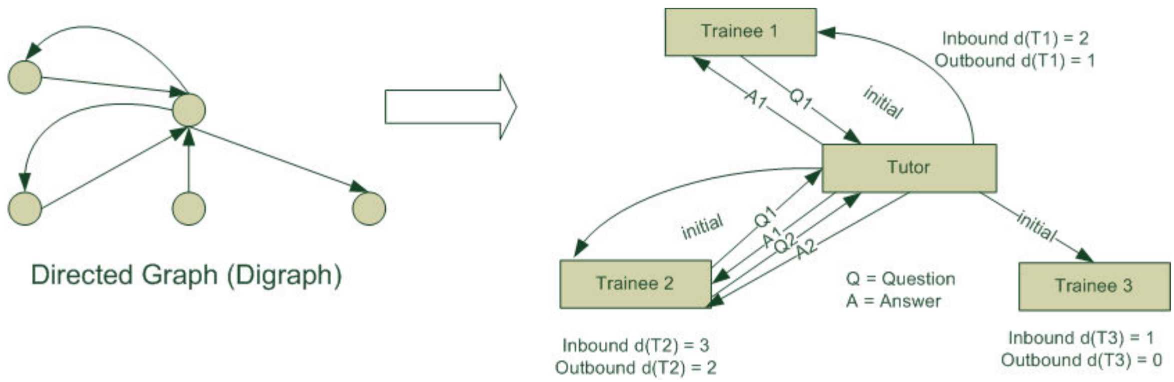

Figure 3: Nodes and arcs of a digraph: a natural way to present a communication network

In graph theory we count the number of edges *going from* or *coming to* the node. This count is called *node degree*. The notation is *d(v), v* denotes *vertex,* i.e. *node*. For digraphs, the number of ingoing arcs (edges pointing *to* the node, notation: $d^-(v)$) and outgoing arcs (the edges pointing *from* the node, notation: $d^+(v)$) is considered. These degrees provide us, for example, with information about the communication intensity of the trainees and tutors. The node degree is one of the most important quantifiers in SNA, and is / was therefore intensively used in this thesis. Additionally to node degree, some more terms are defined with the means of graph theory, which serve as the tools of SNA, but here we mention only the three that were used for this research:

- *Cliques*: the aim is to detect sub-networks inside a particular social network, e.g. concentrated communication between a group of nodes – in such a clique the nodes are more connected with each other than with other nodes of the network.
- *Degree Centrality*: the centrality of an actor in the social network is calculated by dividing the node degree by the total number of edges in the network.
- *Density*: the density of a particular social network is calculated as the total number of edges in the network divided by the maximum number of edges possible for this network.

Centrality and density are used in a similar way as node degree as important quantifiers of the social network.

The works provided by (Silva and Figueira, 2012; Haythornthwaite 2005) can serve as typical examples of utilization of the SNA methods in E-Learning. These methods have been used intensively in this thesis in Papers 1 and 3.

2.2. CONTROL THEORY AND GRAPH THEORY

Control theory is also referred to as 'feedback theory' and is used mostly in engineering for complex system investigation. It can be succinctly described as a methodology to transform a system from one state to another in a finite number of steps (Åström and Murray, 2008). According to (Sontag, 1998), the object of observation is treated as controllable if it can be manipulated so as to change from its initial state into the desired one.

The use of control theory described in (Liu et al., 2011) is especially interesting for this piece of research. In their paper, the authors focus on the controllability of complex networks. They seek to find an efficient method to make a network controllable according to the definition given above. After an analysis of the possibilities to use classic methods of control theory provided in (Kalman, 1963) to investigate the controllability of a complex communication network, the authors realized that this method was rather inefficient for large networks, since it requires the computation of 2^n-1 distinct combinations, for a network of n nodes. Clearly, such computing complexity (exponential!) should be avoided for large networks. Therefore, the authors investigated alternative methods and also applied graph theory here. They showed that control over the so-called *driver nodes* in the network leads to full network controllability.

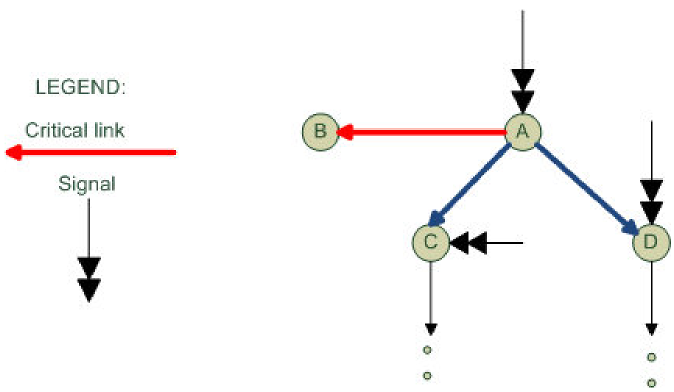

Figure 4: The nodes A, C and D are able to transmit the signal to other nodes in the network, but at node B, the signal terminates –node B is the driver node and the link from A to B is considered as *critical*.

If the communication network was represented as a digraph, the driver nodes are defined as the nodes of the network that are not able to transmit the signal coming as an input from the outside world into the communication network. By applying the graph theory method of finding the *maximal matching* in the arbitrary graph, the unmatched driver nodes can be determined more easily. According to (Murty and Bondy, 2008), the *matching* in a graph is defined as a set of pairwise nonadjacent edges. Two ends of each edge of a matching set are *matched* under the given matching set and each node incident with an edge from the matching set is said to be *covered* by the matching. Hence, a *maximal matching* covers as many nodes of a graph as possible. A *perfect matching* covers all nodes of the graph. Perfect matching is not possible for graphs with an odd number of nodes, because every matching involves an even number of nodes. The special case where matchings were originally investigated in graph theory is that of bipartite graphs (Bondy and Murty 2008).

Maximal matching is a matching that cannot be extended to a larger matching. For the algorithms that detect the matchings in the graphs, this is one of the basic preconditions for examining the matching in the arbitrary graphs. These algorithms execute in polynomial time. Therefore the authors in (Liu et al. 2011) propose to use these algorithms for the detection of driver nodes since they can be computed more efficiently than the classic method (exponential time) mentioned above.

We utilized this finding in Paper 2, proposing a method for the construction of bipartite graphs that guarantees an easy maximal matching computation – within the social network of E-Learning trainees and tutors – whose purpose is to locate the potentially isolated nodes in the network, i.e. to find those trainees who could possibly have a low communication level with their tutors. By finding such nodes in the network, thanks to the above-mentioned findings in control theory, the observed network is getting to be controllable.

2.3. Ontology

The term *ontology* originates from philosophy. It constitutes the study of being (existence, reality) and its categories, and the relationships between them (Smith, 2001). However, in computer science, the term ontology is defined in the following way (citation of the updated definition of ontology for computer science given by Gruber, 2009):

[start citation]

In the context of computer and information sciences, an ontology defines a set of representational primitives with which to model a domain of knowledge or discourse. The representational primitives are typically classes (or sets), attributes (or properties), and relationships (or relations among class members). The definitions of the representational primitives include information about their meaning and constraints on their logically consistent application. In the context of database systems, ontology can be viewed as a level of abstraction of data models, analogous to hierarchical and relational models, but intended for modeling knowledge about individuals, their attributes, and their relationships to other individuals. Ontologies are typically specified in languages that allow abstraction away from data structures and implementation strategies; in practice, the languages of ontologies are closer in expressive power to first-order logic than languages used to model databases. For this reason, ontologies are said to be at the 'semantic' level, whereas database schema are models of data at the 'logical' or 'physical' level. Due to their independence from lower level data models, ontologies are used for integrating heterogeneous databases, enabling interoperability among disparate systems, and specifying interfaces to independent, knowledge-based services.

[end citation]

Ontologies are used in numerous research areas, but we need only mention here the methods that influenced this work directly.

Semantic social networks

According to (Mika 2005), semantic social networks can be viewed as a tripartite model of ontologies. More precisely, the author in (Mika 2005) adds the social component to the traditional ontologies model: the social dimension expressed by actor is added to the model, which is made of concepts and instances. This produces the Actor-Concept-Instance model or tripartite model of ontologies. The first intended use of this model was the extraction of community-based ontologies from websites. However, due to its clarity, this analytical approach influenced several other researchers who used the term 'semantic social networks' to denote networks where relations between people (social networks) and relations between ontologies (ontology networks) are both present (Jung and Euzenat 2007).

Semantic social networks and knowledge sharing

In the work of (Jung and Euzenat 2007), the tripartite model of ontologies mentioned in the previous section is interpreted as a three-layered model consisting of the social, ontology and concept layers. However, in that particular work, the authors go a step further and define a metric to measure the distance between ontologies and use the distance function for a matching algorithm to find people in a social network who use the same, or similar (or, better expressed, close) ontologies and measure their affinity.

Ontology engineering in E-Learning

Ontology networks are defined in the work (d'Aquin et al. 2006) as a collection of ontologies related to each other through various relationships. In the work (Pernas et al. 2012) the term ontology engineering is viewed as a conceptualization of ontology networks. The authors propose that an ontology network for E-Learning should conceptualize the following domains:

- The student domain: information related to students' profiles and their preferences as well as behavior in an E-Learning environment
- The learning domain: learning objects and educational material
- The technological domain: devices used for E-Learning by students and their technological environment.

The authors explain how this conceptualization can be used to adapt the E-Learning setting to the context of the learner. The conceptualization of these three domains results in a context ontology network. The aim is to help learners obtain learning material that fits the learners' particular situation as much as possible, depending on the learners' already acquired knowledge and their different technological facilities, such as devices used for accessing the E-Learning system (smartphone / mobile device vs. computer with high-resolution screen and similar). For this purpose, the relations in the context ontology network are explored with the means of SWRL (Semantic Web Rule Language, (Horrocks et al. 2004)) in order to define the rules inside the context ontology language and assess the current values of instances existing in the network.

In this thesis, we combined the three methods described above. More precisely, we combined the semantic social network model (considered as a tripartite ontology model) with the situation-aware ontology model for E-Learning, applied to specific E-Learning settings. Our aim was to provide a framework for the recommendation of the trainees to each other, to enhance trainee-to-trainee collaboration, and to provide the possibility for trainees to ask each other for help, and to help each other, if their knowledge levels differed. This approach is described in detail in the Paper 4 in Chapter 3.

2.4. Statistics

The following methods, drawn from behavioral statistics, have been used in this piece of research:

- Checking dependencies between diverse quantifiers with Pearson Product Moment Correlation Coefficients (PPMCC)
- Null-hypothesis testing with *t* test and chi-square test.

The main source for the methods in behavioral statistics was the book of (Cohen, 1998).

Pearson-Product-Moment-Correlation-Coefficients (PPMCC)

Behavioral statistic scale according to (Cohen, 1998) is presented in the Table 1. The author argued that such a scale is more applicable to social sciences, i.e. that the interpretation depends on context and purposes. For example, the correlation of 0.9 can be interpreted as small dependency when investigating the results of the verification of a physical law and the accuracy of the results provided by some measurement instrument, but the same value can indicate a high dependency between two social factors.

Table 1: The scale of interpretation of PPMCC is provided by (Cohen 1998) and commonly used in social science experiments.

Correlation (dependency)	Negative PPMCC	Positive PPMCC
None	−0.09 to 0.0	0.0 to 0.09
Small	−0.3 to −0.1	0.1 to 0.3
Medium	−0.5 to −0.3	0.3 to 0.5
Strong	−1.0 to −0.5	0.5 to 1.0

In the experiments in this thesis (in all the papers containing experiments in the Chapter 3), this scale of interpretation of PPMCC results was used.

t Tests

t tests were used for sample statistic tests in order to evaluate the differences between the average learning outcomes of the trainees in different time periods, i.e. *before* and *after* applying the methods for collaboration between trainees and tutors. We applied this approach in Papers 3 and 4 as well as in the 'Final Evaluation' chapter.

Chi-Square tests

With chi-square tests we wanted to assess the homogeneity of the learning outcome results, i.e. to see the difference in the 'quality' of the learning outcomes in the different time periods. Therefore, we classified the learning outcomes of the trainees; then we checked with chi-square tests how significantly the levels of acquired knowledge (measured by the learning outcome categories) differed from each other before, during, and after implementation of the framework presented in this thesis. We applied this method in the 'Final Evaluation' chapter.

In both cases, *t* tests and chi-square tests, we used statistical inference in the following way: if the p-value was less than .05, the difference between the observed experimental result sets was treated as significant.

All the statistical results provided in this thesis were obtained by using the statistic calculation package R (R-Project, 2013).

3. Papers

This chapter contains the main papers that were published during the research work and present the results achieved step by step. The papers cover the research questions described in Chapter 1. Furthermore, they provide the following.

- Analysis of the impact of implicit and explicit social networks in the chosen E-Learning setting
- Specification of the framework for collaboration enhancement aiming to improve the learning outcomes of the trainees
- Two add-ons to the E-Learning system that were implemented as a consequence of the proposed framework
- Comparison and analysis of the learning outcomes of the trainees *before* and *after* the implementation of the framework.

For each paper, the main objective, applied methods and results are provided in a concise way in the sections below.

Note about the experiments and systems mentioned in the papers

All experiments mentioned in the papers were made within the organization where the author was employed during the whole research work. The experimental data is fully anonymized in order to protect the privacy of the members of the organization and confidentiality requirements of their companies. The ICT systems that were part of this research are owned by the organization where the author is employed; they were installed, and have been maintained and extended under the supervision and leadership of the author, in keeping with the contractual position of the author within the organization. The research has been carried out in cooperation with TU Graz. The students of TU Graz were supervised by the author and worked intensively on evaluation, installation, administration and development of E-Learning system of the organization the author works for.

3.1. Paper 1: How Do Social Networks Influence Learning Outcomes? A Case Study in an Industrial Setting

This paper was first published in the proceedings of the IADIS 2011 WWW Internet conference, and chosen as one of the best papers for publication in the ITSE (Interactive Technology and Smart Education) Journal.

The theoretical framework for the analysis of implicit social networks of trainees and tutors in industrial E-Learning settings, as well as the impact of such networks on the trainees' learning outcome, was investigated in this paper in detail. The findings in this case study were then used in further papers extensively. The paper indicated that the utilization of the implicit social networks of E-Learning participants was useful and should even be considered mandatory nowadays.

The methods used in the paper are SNA and PPMCC. In order to relate social network characteristics of the trainees with their learning outcomes, the SNA method involving "node degrees" was used for quantification. The theoretical framework for the analysis of the impact of the social network position and characteristics of the trainees on their learning outcomes was applied within an industrial experimental setup. The main findings of this research step were that the PPMCC calculation of the relation between social network parameters of the trainees and their learning outcomes showed that the common language relation of trainee with other trainees and tutors and the communication intensity – as measured by the ingoing and outgoing communication node degree of the trainees and their tutor – were positively correlated to the learning outcome.

The findings indicated to the author that further investigation of the potential of social networks of trainees and tutors being used to enhance their collaboration (in order to improve learning outcomes) was necessary.

Interactive Technology and Smart Education

Emerald Article: How do social networks influence learning outcomes? A case study in an industrial setting

Seid Maglajlic, Denis Helic

The current issue and full text archive of this journal is available at
www.emeraldinsight.com/1741-5659.htm

How do social networks influence learning outcomes? A case study in an industrial setting

Seid Maglajlic
RailNetEurope, Vienna, Austria, and
Denis Helic
Institute for Knowledge Management, Graz University of Technology, Graz, Austria

Abstract

Purpose – The purpose of this research is to shed light on the impact of implicit social networks to the learning outcome of e-learning participants in an industrial setting.

Design/methodology/approach – The paper presents a theoretical framework that allows the authors to measure correlation coefficients between the different affiliations that exist in an organization and the final learning outcome. The correlation between learning outcome and the communication intensity in the implicit social network of the e-learning participants is also observed. For the quantification of the communication intensity and affiliation network position of e-learning participants, the methods from the graph theory are applied.

Findings – The values of the correlation coefficients between communication intensity and learning outcome show the significance which motivates the authors for further research on engineering of the social networks in the e-learning environment.

Research limitations/implications – This case study is performed in an industrial setting.

Practical implications – The results of this case study influence the further development of the e-learning system that has been used in the experimental setup in this paper, especially the user management module. The algorithm for matching the trainees with tutors is in development.

Originality/value – The impact analysis of the influence of the social network position of the learner in e-learning environment by comparing the test results before taking the e-learning course and after taking the course (learning outcome) is provided by measurements of the correlations between the social network position and communication intensity of the learner with the learning outcome.

Keywords Implicit social networks, E-learning, Learning outcome, Social networks, Computer based learning, Learning methods

Paper type Case study

1. Introduction

The goal of this work is to shed more light on the dependencies between the factors underlying e-learning processes within an organization and the final outcomes of these processes.

In the research on methods for enhancement of collaboration between distance learners until now, some useful findings can be recognized: modelling of the users and groups in e-learning implicitly corresponds to social network modelling and can be utilized as such (Haythornthwaite, 2005), applying of behavioural laws to community-building when

The authors would like to thank RailNetEurope for the support of this case study.

Interactive Technology and Smart Education
Vol. 9 No. 2, 2012
pp. 74-88

1741-5659
DOI 10.1108/17415651211242224

organizing the social network units for the collaboration of distance learners is possible (Wang and Li, 2007), just to mention a few. However, during the investigation of related research we have recognized the lack of the impact analysis of the applied methods, i.e. we could not find sufficient indications on how social network based collaboration influenced the learning outcome of the distance learners.

Thus, in this paper we propose a theoretical framework that measures correlations between the social context, the communication activities and the improvements of knowledge of the workforce within a given organization.

In addition to this theoretical framework, the possibility of engineering the social network with a view to influencing (improving) the learning outcome is investigated. More precisely, the results of the application of the theoretical framework mentioned above to the industrial environment are considered useful for modelling the parameters to be used for the further construction of the social networks in the e-learning environment.

The rest of the paper is organized as follows: in the next section we present our theoretical framework. In Section 3 we discuss a case study carried out in an international railway industry setting and shortly present the dataset and experimental setup of our empirical analysis. Next, we discuss the case study's results and their possible implications. In Section 5 where the implications are discussed, we also provide a brief introduction to our most accurate research which includes utilization of some of the findings presented in this case study. Finally, we conclude and provide some directions for future work.

2. Theoretical framework

2.1 Social context: affiliation networks

In each organization there are numerous implicit social networks. People share interests, an internal group, competences, language, educational/professional background and so on. One of the goals of our theoretical framework is to represent, and then analyze such implicit social networks with respect to their influence on the final learning outcome of participants in an e-learning process. Implicit social networks are typically represented as two-mode or affiliation networks (Breiger, 1974). In affiliation networks we have two types of entities: actors and events (sometimes also called organizations (Wasserman and Faust, 2009)). These entities are related by links, i.e. by affiliations of the former to the latter. An explicit social network is then obtained by a one-mode projection of affiliation networks onto actors (Figure 1). In this process all actors sharing an affiliation become linked to each other by an undirected link in the one-mode network (Guillaume and Latapy, 2006).

Figure 1. One-mode projection of two-mode affiliation network

The basic idea here is to capture important affiliations in an organization and model them as affiliation networks. In the next step, different social networks are obtained by the one-mode projection of the affiliation networks. We can either select single affiliations, or any interesting combination of affiliations, to obtain social networks. In this way we can analyze the influence of single affiliations but also their combined effects on the learning outcome. For example, let us suppose that we are interested in the effects of spoken language in a geographically distributed organization. It is to expect that such a social network in today's global society will be almost fully connected as the majority of involved actors will speak a global language, e.g. English. However, if some of the actors do not speak English they might be disconnected from the network. If this disconnectedness correlates with a poor learning performance, we might broaden our analysis by looking for other affiliations that have positive correlations with the learning outcome. Let us suppose that we find that the business branch has such a positive correlation. Finally, we can engineer a change in the affiliation networks by suggesting to the affected actors to affiliate with yet another business branch in order to improve their chances of a better learning outcome.

Consequently, the influence of the position of an actor in the affiliation network on communication with other actors in the network is also worth noting. For example, it is interesting to compare whether the organizational affiliation has more impact on the communication between actors than the business branch affiliation. Analytical results (correlation coefficient calculation) might express some unexpected values that can be used for further engineering of the affiliation networks. Actually, we intend to enhance the information flow between the actors in the network for a common purpose: the achievement of a satisfactory learning outcome.

2.2 Communication network

The communication network is modeled in a more straightforward manner than affiliation networks. This task is also intuitively easier, since we do not need to separate the sets of actors from the affiliations targets. We define actors as network nodes and information flow between actors as links between the nodes in the network. Thus, whenever there is some kind of information flow occurring between actors we connect those two actors by a link. This network is a directed network – links point from the actor who initiated the information flow to the actor who received that information. The structure of the communication network also tells us how efficiently the information has flown between the actors involved. The basic idea here is, on the one side, to investigate the effects of communication between actors on their learning outcomes – by measuring correlation between, e.g. communication intensity and the learning outcome. On the other hand, by applying the same method of measuring the correlations, one can analyze the dependencies between the affiliation networks and communication and, accordingly, communication intensity in the future steps of our research. As we expect communication intensity to positively correlate with the learning outcome, the goal of the future analysis would be to identify prerequisites for the improvement of communication between affected actors.

2.3 Learning outcomes

Typically, during an e-learning course, the participants will be tested twice: before they start the course and after they have completed it. Comparison of the results provides a simple indication of the learning progress. In our study, we quantified the test results by the ratio l, which is the ratio of the points achieved after and before the course. Finally, we can measure the correlation of this ratio with other figures obtained from the social networks and the communication network. Our aim here is to identify promising affiliations and promising communication practices that have a positive effect on the learning outcome. By considering the learning outcome in our analysis, we achieve a major milestone in our research: the analysis of the impact of the applied methods (Maglajlic, 2011).

2.4 Correlation coefficients

The usage of Pearson product-moment correlation (or more simply, the Pearson correlation) will help us to detect the linear dependency between the characteristics of the trainees as members of the implicit social network and the learning outcome. In order to be able to calculate the correlations, the characteristics of actors (i.e. the nodes of the network) have to be quantified. Therefore, the degrees of the nodes in the networks (affiliation network and communication network) on one side, and the learning outcome ratio on the other side are used. Hence, in order to estimate correlations, we apply the Pearson correlation between the node's in-degree and the node's out-degree from the communication network (or combined node's degree in the undirected social networks) and the learning outcome ratio l. Moreover, the Pearson correlation between, e.g. the node's degree in a social network obtained by different affiliations and the learning outcome is interesting for us. Let us denote node's in-degree and node's out-degree of the communication network with d^- and d^+, respectively. Similarly, we denote the node's degree of a particular social network by $d(a)$ where a is the given affiliation used for the creation of the social network. Then, the Pearson correlation coefficient between, for example, $d(a)$ and l is defined as below.

The Pearson correlation between $d(a)$ and l:

$$\mathrm{corr}(d(a), l) = \frac{\sum_{i=1}^{n}(d(a)_i - \overline{d(a)})(l_i - \bar{l})}{\left[\sum_{i=1}^{n}(d(a)_i - \overline{(d(a|)})^2 \sum_{i=1}^{n}(l_i - \bar{l})^2\right]^{1/2}}$$

Now, substituting different affiliations such as spoken language or business branch for a will provide us with an indication on how a particular affiliation and learning outcome ratio depend on each other. To calculate the correlation between the out-degree in the communication network and the learning outcome ratio, we will use the below equation.

The Pearson correlation between d^+ and l:

$$\mathrm{corr}(d^+, l) = \frac{\sum_{i=1}^{n}\left(d_i^+ - \overline{d^+}\right)(l_i - \bar{l})}{\left[\sum_{i=1}^{n}\left(d_i^+ - \overline{d^+}\right)^2 \sum_{i=1}^{n}(l_i - \bar{l})^2\right]^{1/2}}$$

The result of the Pearson correlation coefficient calculation, exactly for this case shown in equation (1), is further discussed in Section 4.4, since it shows the significance (existing of the linear dependence) which helps us for further research.

Finally, the Pearson coefficient calculation can also be provided for the affiliations and out-degree in the communication network. The corresponding equation is shown below.

The Pearson correlation between $d(a)$ and d^+:

$$\mathrm{corr}(d(a),\ d^+) = \frac{\sum_{i=1}^{n}(d(a)_i - \overline{d(a)})\left(d_i^+ - \overline{d^+}\right)}{\left[\sum_{i=1}^{n}(d(a)_i - \overline{d(a)})^2 \sum_{i=1}^{n}\left(d_i^+ - \overline{d^+}\right)^2\right]^{1/2}}$$

In the same way as described above, different affiliations will have to be substituted for a during the concrete calculation. Furthermore, the out-degree can be replaced with in-degree in the formula – in that case the correlation coefficient for affiliations and in-degree of the communication network will be calculated.

2.5 Summary

To sum up, our theoretical framework can be described as a four-step process:

(1) Identify affiliation networks that are of interest. Represent the affiliation networks as two-mode networks. Create the social networks as one-mode projection of affiliation networks using diverse combinations of affiliations.

(2) Obtain the communication network of actors using the information flow as criteria for linking the actors.

(3) Calculate the learning outcome ratio l for each participant in an e-learning course.

(4) Calculate correlation coefficients for all interesting combinations of node's degree from diverse social networks, in-degree and out-degree from the communication network and the learning outcome ratio. Compare the correlation coefficients with each other in order to recognize the relation strength of the particular social network aspect of the e-learning participants (actors in the network). The analysis sheds light on the impact of different affiliations and actors' communication activity on their final score in the e-learning course. The analysis results provide an input for engineering of the underlying networks, the final goal being to improve the learning outcome.

3. Experimental setup

In this section, we introduce our case study, which was carried out in an industrial setting. We also describe how we applied the above proposed theoretical framework to a practical situation. We chose an international organization that coordinates the work of the railway infrastructure companies in Europe. The organization has 38 member companies in 27 different European countries. The customers of the railway infrastructure companies include more than 100 train operators/train operating companies. The members of this organization are meant to work together, as an association with a common business purpose, in order to provide better quality of service for their customers. To this end, the organization provides several IT tools which are used across the member companies for the coordination of some of their business processes and to facilitate cooperation.

One of these IT tools has a special focus on train path management and supports the international path booking process. There are currently 767 users registered for this tool (their number is growing, with an average two new registered users per week) belonging to 150 companies (organization members and their customers). Thus, the users belong basically to two different business branches: the railway infrastructure and train operating companies. For the purpose of our study, the infrastructure branch can be divided further into infrastructure manager activities and path allocation bodies. The train operating companies can be roughly divided into freight train operators and passenger train operators.

Due to the fact that the tool supports a rather complex international business coordination process, users require some training in order to be able to use it. However, it would be fully unpractical to reach all the users with classic face-to-face training sessions as their number is growing considerably. Moreover, the users are spread across a vast geographical area. Therefore, an e-learning system has been implemented and deployed.

An e-learning course containing information on how to use the tool is offered to all users. This e-learning course contains seven lessons; these describe very precisely the steps required to be able to make good use of the tool within the business process. Tests are provided, containing 30 questions; the maximal score is 100 points. All users of the tool are automatically registered for the e-learning course. The users are advised to take an examination before and after taking the course. Thanks to this approach, the improvement ratio can easily be detected. In addition, users from the member companies are candidates to become certified trainers for the tool. The common agreement is that an applicant who scores more than 75 percent at the examination automatically becomes a trainer. Certified trainers then become tutors in the e-learning course. As a consequence, the e-learning users are divided into tutors and participants. The participants are assigned to tutors.

Our experiment was conducted with 26 e-learning users from 16 countries. Those who took part in our experiment speak 17 different languages and belong to 19 different companies. Among them, there were eight tutors (certified trainers) and 18 e-learning course trainees.

All of the trainees applied for the certification. Hence, the challenging aim was to improve their knowledge significantly, so that they could achieve at least 75 percent of the total score. The trainees had two weeks of time to take the e-learning course after the first test and at the final session they were required to take the examination. The starting knowledge, measured by the score at the first test, varied from 61 to 85 percent. Similarly, the improvement also varied from 1.06 (almost no improvement) to 1.58 (great improvement). Two trainees did not reach the 75 percent mark and an additional four were under 80 percent, with a low improvement ratio. Investigating the cause for the weak improvement of these candidates inspired us to carry out this research, since the social network measurements that were proposed in the previous chapter show that the position of the participant as an actor in the affiliation network and his/her characteristics in the corresponding communication network do correlate with the learning outcome. As we will see later in more detail, the most significant parameter is the degree of the outgoing communication of the trainee.

3.1 Affiliation networks

We applied the following affiliations to construct the social networks:

- *Language.* As a natural precondition for communication between actors in a social network, the language affiliation influences the information flow as well as the acknowledgment of information received.
- *Country.* When actors share a geographical location, this intuitively simplifies communication between them, due to (in most cases) the common language, and greater opportunities for face-to-face communication.

- *Business branch.* Furthermore, the industry branch should not be underestimated – in some industry branches, participants communicate across national and organizational borders very intensively due to the business process in which they participate. Historically, the railway industry has shown this behaviour since the beginning of international railway traffic – the border-handling business processes have always been agreed and specified by the particular branch experts in order to avoid the risk of traffic accidents.
- *Organizational unit.* Actors belonging to the same organization in most cases use business-driven communication, especially if they participate in a common business process.

3.2 Communication network

The communication between tutor and trainee is especially interesting. E-mail messages between tutors and trainees as well as messages posted in the forum by tutors and trainees are counted as communication attempts, i.e. information flow. If a trainee is assigned to a particular tutor, one-way communication from tutor to trainee is guaranteed. Tutors are required to initiate contact with their trainees, by sending them an introductory text and an invitation to collaborate. Communication from tutor to other trainees out of the virtual classroom is also possible, i.e. it is not restricted or forbidden. Any trainee can, theoretically, contact any tutor in order to get the required information. Communication in this direction (opposite to the initial flow from tutor to trainee) is valuable for further analysis. For example, there can be some questions from the trainee to the tutor in that case, an answer by the tutor increases knowledge acquisition of the trainee. Also, in some cases the trainee may send an acknowledgment of the explanation of the tutor, which might also be considered as a positive sign of knowledge acquisition.

3.3 Learning outcomes

The test results are shown in Table I (the data are anonymized, trainees are denoted by P < number > and tutors by T < number >).

The column "before" carries the information about the score (number of points) of the trainee before taking the course. In the same way, the column "after" carries the score (number of points) of the trainee after taking the e-learning course. The learning outcome is measured by the values from column l (the improvement ratio), which is calculated as the ratio between the scores after and before taking the e-learning course.

3.4 Correlation coefficients

In our study we applied the following correlation coefficient calculations for analysis of the dependency (i.e. the link) between social network degrees and learning outcome:

- Language affiliation and the learning outcome ratio: $corr(d(lang); l)$.
- Country affiliation and the learning outcome ratio: $corr(d(co); l)$.
- Business branch affiliation and the learning outcome ratio: $corr(d(bb); l)$.
- Organizational unit and the learning outcome ratio: $corr(d(ou); l)$.

We also measure the following correlations between the in-degree and out-degree in the communication network and the learning outcome ratio:

Name	Before	After	l	Tutored by
P1	85	100	1.176	T1
P2	63	100	1.587	T1
P3	63	67	1.063	T1
P4	61	65	1.066	T1
P5	70	85	1.214	T3
P6	72	85	1.181	T5
P7	81	90	1.111	T2
P8	85	100	1.176	T3
P9	82	95	1.159	T7
P10	70	81	1.157	T4
P11	65	76	1.169	T7
P12	70	76	1.086	T3
P13	72	85	1.181	T2
P14	74	77	1.041	T3
P15	65	91	1.4	T8
P16	66	79	1.197	T6
P17	73	90	1.233	T5
P18	80	100	1.25	T8

Table I.
Test results

- Communication network in-degree and the learning outcome ratio: $corr(d^-; l)$.
- Communication network out-degree and the learning outcome ratio: $corr(d^+; l)$.

Finally, we investigate the correlations between the affiliations and in-degree and out-degree of the nodes in the communication network:

- Language affiliation and both in-degree and out-degree: $corr(d(lang); d^-)$ and $corr(d(lang); d^+)$.
- Country affiliation and both in-degree and out-degree: $corr(d(co); d^-)$ and $corr(d(co); d^+)$.
- Business branch affiliation and both in-degree and out-degree: $corr(d(bb); d^-)$ and $corr(d(bb); d^+)$.
- Organizational unit and both in-degree and out-degree: $corr(d(ou); d^-)$ and $corr(d(ou); d^+)$.

4. Results

4.1 Degrees of affiliation

In Table II, the language, country, organizational unit, and business branch affiliations are shown, along with the improvement ratio.

The numbers in column $d(lang)$ contain the number of actors from the social network that speak the same languages as the trainee. In the terms of graph theory, d is denoting the degree of a node: the number of edges (links) going out from or pointing to a node (Bondy and Murty, 2008). To clarify it in more detail for our case, we can say that this is the number of nodes linked to the particular node as the result of the one-mode projection described in Section 2.1, i.e. by representing the language affiliation of the nodes in the network as the links between them. The most significant values are shown for P3 and P4, who have a low $d(lang)$ value and the lowest

Table II.
Affiliation degrees and learning outcome

Name	d(lang)	d(c)	d(ou)	d(bb)	l
T1	23	1	1	20	
P1	16	2	2	20	1.176
P2	16	2	2	20	1.587
P3	1	2	2	20	1.063
P4	1	2	2	3	1.066
P5	22	1	1	3	1.214
P6	21	2	2	20	1.181
T2	23	3	3	20	
P7	1	3	3	20	1.111
T3	23	2	1	20	
P8	15	1	1	20	1.176
P9	18	1	1	1	1.159
P10	1	2	2	20	1.157
T4	6	2	2	20	
P11	1	1	1	20	1.169
P12	5	2	2	20	1.086
P13	3	3	3	20	1.181
P14	16	2	2	20	1.041
T5	25	2	2	20	
T6	23	2	1	20	
P15	15	1	1	20	1.4
T7	25	1	1	20	
P16	16	2	1	1	1.197
P17	18	1	1	3	1.233
T8	22	2	1	20	
P18	22	2	2	20	1.25

improvement ratio. Other values, $d(c)$, $d(ou)$, and $d(bb)$ compared to the improvement ratio, are not showing such an obvious result which could lead us to some intuitive conclusion, i.e. they do not show significant dependencies that would have obvious implications. In other words, we cannot conclude that the trainees that belong to the same organizational unit or country as several others are achieving better results on the testing in our experiment.

4.2 In-degree and out-degree

Table III shows the information about in-degree, out-degree and improvement ratio of all actors in the experimental network (trainees denoted by P < number > and tutors denoted by T < number >). In-degree (ingoing communication intensity for a node in the communication network, i.e. the number of the directed links from other nodes to the particular one) is denoted by d^- and out-degree (the number of outgoing links from the node in the communication network) is denoted by d^+. The right-most column carries the information about the improvement ratio l of the particular trainee. Clearly, the improvement ratio is skipped for the tutors, since they have not been tested in this experiment.

Similarly as concluded for the $d(lang)$ and the participants P3 and P4 in Table II, the improvement ratio is the lowest as their out-degree is 0. More precisely, the trainees with the lowest communication intensity have achieved the weakest results on the tests after taking the e-learning course in our experiment. The correlation calculation results

Name	d^-	d^+	l
T1	6	10	
P1	4	3	1.176
P2	4	3	1.587
P3	1	0	1.063
P4	1	0	1.066
P5	3	2	1.214
P6	1	1	1.181
T2	4	6	
P7	4	3	1.111
T3	5	9	
P8	4	3	1.176
P9	4	3	1.159
P10	3	2	1.157
T4	2	3	
P11	1	1	1.169
P12	1	0	1.086
P13	2	1	1.181
P14	1	0	1.041
T5	4	5	
T6	2	3	
P15	4	3	1.4
T7	4	5	
P16	3	2	1.197
P17	4	3	1.233
T8	6	8	
P18	4	3	1.25

Table III.
In-degree and out-degree

which are presented in Section 4.4 containing the final comparison lead us to this conclusion.

4.3 Affiliation and communication

In Table IV, we cross-check the correlations between degree of affiliation and communication intensity of the actors. More precisely, we investigate whether a linear dependency exists between the position of the actor in the affiliation social network and her/his communication during the course.

We see in Table IV (in the rows 1 and 2) that the language affiliation possesses the highest correlation coefficient with communication in-degree and out-degree. These correlation coefficients will be compared with the coefficients calculated with learning outcome in Section 4.4.1.

1	$corr(d(lang); d^-)$	0.31359
2	$corr(d(lang); d^+)$	0.33806
3	$corr(d(ou); d^-)$	−0.15255
4	$corr(d(ou); d^+)$	−0.18847
5	$corr(d(co); d^-)$	−0.12495
6	$corr(d(co); d^+)$	−0.16084
7	$corr(d(bb); d^-)$	−0.14022
8	$corr(d(bb); d^+)$	−0.09647

Table IV.
Degree of affiliation and communication in-degree and out-degree

4.4 Final comparison

The final results of the experiment, including the correlation calculations with the learning outcome, are shown in Table V.

On the basis of the results shown in Table V, we may conclude the following:

(1) In total, 16 trainees out of 18 have successfully passed the certification test. These 16 participants are, therefore, candidates to become future tutors in the e-learning system. They are also the key persons in their companies to spread knowledge on how to use the tool efficiently to enhance the common business process.

(2) The average improvement figure, l, signifies a general knowledge improvement of the group of almost 20 percent.

(3) The correlation between language affiliation degree of the trainees and their knowledge improvement testifies to the significant level of linear dependency between these two parameters. Hence, the language affiliation of the participants does have an influence on the results. Referring to Cohen (1988), the correlation coefficient value given in (3) can be seen as a medium correlation.

(4 and 5) The correlation coefficient between organization and country affiliation on the one side, and the improvement ratio on the other side, does not lead us to conclude that organizational or geographical affiliation influences the results much.

(6) Since the correlation coefficient between branch affiliation and knowledge improvement is near 0, we can say that these two parameters are rather independent, but still positively correlated.

(7 and 8) The most significant parameters showing the highest linear dependency level with knowledge improvement are the ingoing and outgoing degree of the actors in the communication network. According to Cohen (1988), as this correlation has the value greater than 0.5, it can be, therefore, seen as a strong correlation.

4.4.1 Comparison of affiliation-communication correlations and effect size. In Table VI we provide the coefficient of determination calculation: if we denote the result of correlation coefficient calculation with r and call it effect size (Cohen, 1988), the coefficient of determination (Steel and Torrie, 1960) is calculated as the square of r, i.e. r^2. Let us briefly check the results shown in Table VI and stress an interesting finding: the coefficient of determination r^2 calculated from the correlation coefficient for the business branch and communication out-degree (0.00931) is close to the value of the coefficient of determination of the correlation between degree of business branch affiliation and the learning outcome (0.00708). We compare these two values, because the analysis in the previous section has shown that out-degree of the communication network has the

1	Successful examination (>75 percent score)	16/18
2	Average improvement ratio (l)	1.19149
3	$corr(d(lang); l)$	0.40947
4	$corr(d(ou); l)$	− 0.20354
5	$corr(d(co); l)$	− 0.2043
6	$corr(d(bb); l)$	0.08412
7	$corr(d^-; l)$	0.55585
8	$corr(d^+; l)$	0.59494

Table V.
Final comparison

Correlation	r	r^2
$corr(d(lang); l)$	0.40947	0.16767
$corr(d(ou); l)$	−0.20354	0.04143
$corr(d(co); l)$	−0.2043	0.04174
$corr(d(bb); l)$	0.08412	0.00708
$corr(d^-; l)$	0.55585	0.30897
$corr(d^+; l)$	0.59494	0.35395
$corr(d(lang); d^-)$	0.31359	0.09834
$corr(d(lang); d^+)$	0.33806	0.11428
$corr(d(ou); d^-)$	−0.15255	0.02327
$corr(d(ou); d^+)$	−0.18847	0.03552
$corr(d(co); d^-)$	−0.12495	0.01561
$corr(d(co); d^+)$	−0.16084	0.02587
$corr(d(bb); d^-)$	−0.14022	0.01966
$corr(d(bb); d^+)$	−0.09647	0.00931

Table VI. Effect size coefficient of determination

strong (actually, the strongest among other observed parameters) correlation with the learning outcome. On the other hand, the coefficient of determination of the correlation coefficient for the business branch and communication in-degree (the second strong correlation) shows a significant slope compared to the value of the coefficient of determination of the correlation coefficient between degree of business branch affiliation and the learning outcome (from 0.00708 to 0.01966). Thus, the stability of the correlation coefficient for this particular affiliation on one side and significant increase of the coefficient of determination on the other side will help us in further research with choosing the parameters for a further social network engineering designed to improve the learning outcome of the trainees. We will further discuss the implication of this phenomenon in Section 5.

Other affiliation-communication correlations (those between country/organization affiliation and communication intensity) do not show the significance that would lead us to such an interesting finding as it was the case with the business branch affiliation. Their effect sizes differ from those of the correlation coefficients between affiliations and learning outcome more than the effect size values of business branch correlation with out-degree and business branch correlation with learning outcome. The absolute values of the correlation coefficients of country and organization affiliation and in-degree and out-degree are lower than the absolute values of the correlation coefficients of country and organization affiliation and the learning outcome. Therefore, the affiliation-communication correlations for country and organization affiliation are less interesting for further discussion than the correlation coefficients related to business branch affiliation.

5. Implications

According to the results of our research, the next stage for the improvement of knowledge about e-learning could be defining actions designed to improve communication in the social network. The problem that was detected can be depicted as follows: the actors in the social network with a low degree of communication intensity, as measured by the outgoing node degree, have scored weaker results on the tests compared with the trainees with a higher node and affiliation degree. The early prediction of potentially poorly connected trainees would be useful from the analytical point of view. As we have already mentioned in Section 2.2, the measurement of correlation between the affiliation degree ($d(a)$) and the communication intensity (d^- and d^+) will help us to see which affiliation has a significant correlation with the communication intensity. Hence we have provided the correlation coefficient calculation between affiliations and communication intensity in Section 4.3 and we have stressed the finding derived from this observation: the correlation coefficient of one of the affiliations remains stable both as regards learning outcome and communication out-degree (the parameter which is strongly correlated with learning outcome). Technically speaking, it could be managed by applying a checking procedure of the trainees' connectivity at an early stage, i.e. during the creation of the trainee's profile in the e-learning system and the assignment of the trainee to a particular tutor and/or virtual classroom, the affiliation with the strong correlation to communication intensity from the previous experiments can be used as assignment criterion. As we have seen in Section 4.3 and mentioned in Section 4.4.1, the correlations that are interesting for such a checking procedure are language and business branch affiliation: language affiliation because it has the highest correlation coefficient; business branch affiliation because it has a stable behaviour compared to its correlation coefficient value, as calculated with the learning outcome. Hence, such a checking procedure encourages us to make further investigations. The preconditions for its implementation are:

- the e-learning users model has to be implemented as a social network model;
- the social network model has to be built according to the affiliation network rules;
- the social network model has to provide the possibility for an initial setup of the communication network between participants and tutors; and
- an algorithm has to be designed that considers the trainees' affiliations and purposes, according to their parameters in the affiliation network, their position in the communication network (assignment to the tutor/virtual classroom).

In our latest research (Maglajlic, 2012) we have proposed such an algorithm for connecting trainees and tutors upon the matching criteria derived from the social network positions of both trainee and tutor. For that purpose, with the strong relation to the results of the research presented in this paper, we have proposed the following matching categories:

- *Language*. Since the language affiliation has the strongest correlation among all observed affiliations with both learning outcome and communication intensity, as shown in Sections 4.4 and 4.3, we have weighted this category with the highest value.
- *Business branch*. Due to its stable correlation coefficient value for both communication out-degree and learning outcome, as well as the significant slope of the effect size (coefficient of determination) of the correlation coefficient between this affiliation and the communication in-degree we have weighted this category with the higher value than two remaining categories: organization and country. The increase of the effect size of correlation coefficient between business branch and communication in-degree indicated us that this particular affiliation increases communication intentions between trainees and tutors, which is very positive if we consider the results of the research given in this paper.
- *Organization and country*. These two categories are used as a good alternative to the business branch matching.

The categories described above are aimed to be used by the matching algorithm which is applied when a new trainee joins the e-learning environment. The algorithm tries to find the tutor for the trainee that suits the best according to the matching criteria defined by the weighting of the categories given above. The resulting network trainees and tutors is the consequence of the proposed social network engineering approach.

6. Conclusion

Thanks to the implications of the study results described above, the problem which will be the focus of our further research has been detected. We propose the following research agenda:

- Establish an algorithm to detect candidates that can, potentially, experience less knowledge improvement than other trainees, according to their parameters and their position in the social network, and to propose a placement in a "better" position in the network. We have already started the implementation towards this aim in the work (Maglajlic, 2012).
- Investigate the possibility of integrating such an algorithm into the user management module of e-learning software platforms.

The analysis of algorithms that may be applied in such a case leads us to the further investigation of the correlations between the parameters in the social network model – including the affiliation network's and communication network's key indicators. The aim of this analysis would be to find the indicators that help increase communication intensity. After such indicators have been detected, the algorithm may take them into consideration, as we have described in the previous section, when proposing the participant's position in the social network model of the e-learning system users' management. The results of the proposed research could help to improve the knowledge of e-learning participants in an industry setting, the implications of this being: a less expensive learning process in industry settings, bringing a clear benefit to the industry.

References

Bondy, J.A. and Murty, U.S.R. (2008), *Graph Theory*, Springer, New York, NY.

Breiger, R.L. (1974), "The duality of persons and groups", *Social Forces*, Vol. 53 No. 2, pp. 181-90.

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, L. Erlbaum Associates, Hillsdale, NJ.

Guillaume, J.-L. and Latapy, M. (2006), "Bipartite graphs as models of complex networks", *Physica A: Statistical and Theoretical Physics*, Vol. 371 No. 2, pp. 795-813.

Haythornthwaite, C. (2005), "Social network methods and measures for examining e-learning", *Social Networks*, April, Citeex:10.1.1.135.6993.

Maglajlic, S. (2011), "On the importance of the impact analysis of social network methods in e-learning", *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, AACE, Chesapeake, VA, pp. 749-52.

Maglajlic, S. (2012), "Engineering social networks using controllability approach applied to e-learning", working paper (accepted for the conference, paper-id: 269) for International Conference for Advanced Learning Technologies (IEEE), Rome, Italy.

Steel, R.G.D. and Torrie, J.H. (1960), *Principles and Procedures of Statistics*, McGraw-Hill, New York, NY, pp. 187-287.

Wang, Z. and Li, L. (2007), "Enable collaborative learning: an improved e-learning social network exploiting approach", *Proceedings of the 6th WSEAS International Conference on Applied Computer Science, Hangzhou, China*.

Wasserman, S. and Faust, K. (2009), *Social Network Analysis: Methods and Applications*, Cambridge University Press, New York, NY.

About the authors

Seid Maglajlic is Chief Technical Officer at RailNetEurope (Vienna, Austria), the joint organization of European Railway Infrastructure Managers. He has an MSc in Mathematics from the University of Zagreb and he is working on a PhD thesis at Graz University of Technology. He currently investigates innovative e-learning methods combined with social network analysis in industry. Seid Maglajlic is the corresponding author and can be contacted at: seid.maglajlic@rne.eu

Denis Helic is an Assistant Professor at the Institute for Knowledge Management (KMI) at Graz University of Technology, in Austria. He has an MSc in Computer Technics from the University of Zagreb and a PhD in Computer Science from Graz University of Technology. His research interests include social computation, online social network analysis, network science, multimedia and hypermedia information systems, and the web.

3.2. Paper 2: Engineering Social Networks Using the Controllability Approach Applied to E-Learning

This paper was presented at the IEEE ICALT 2012 conference in Rome, July 2012, and published in the conference proceedings. It describes a social network engineering concept for the construction of tutored groups of trainees by taking the concordance between tutor and trainees into account. It explains the matching algorithm that was implemented as an add-on to the E-Learning system and helps potentially weak trainees to be placed in the tutored group where they can be tutored more easily, since their social network background is considered to be similar to that of the tutor. This add-on is the first module that was deployed in the chosen E-Learning system according to the framework set by this research. The second add-on is described in detail in the Paper 4.

Engineering Social Networks Using the Controllability Approach Applied to E-Learning

Seid Maglajlic
RailNetEurope
Vienna, Austria
seid.maglajlic@rne.eu

Abstract— **In our previous research we focused on the analysis and further utilization of social networks for E-Learning in industrial settings. The existence of implicit social networks in E-Learning setups has been investigated [4], [5], demonstrated [9], [14] and used [12], [13]. We started investigating the influence of the social network position of a trainee / trainer (tutor) in E-Learning in [9], [12]. In the research provided in [12], some light was shed on the dependencies between the factors underlying E-Learning processes within an organization and the final outcomes of these processes. A theoretical framework for measuring correlations between the social context, communication and knowledge improvements of the trainees in an industrial environment was proposed. The analysis in [12] shows that enhancement of the communication between tutors and trainees within the social network is recommended in order to improve the learning outcome. For this purpose, we investigate how experiences gained elsewhere with applied control theory could be used in our industrial setting.**

Key words: E-Learning; social network controllability

I. Introduction

The construction of the communication network for the purpose of improving cooperation between tutors and trainees has to be optimized. The difficulty is that the implicit affiliation networks which are naturally formed by E-Learning participants have to be re-constructed (i.e. re-engineered) so that the resulting communication network is efficient. Efficiency of the communication network in our particular case means that the tutors and the trainees have communicated sufficiently, and that the learning outcome shows a significant improvement of the knowledge of the trainees. Therefore, we need additional means which help us for this purpose. Control theory offers an approach which supports our idea: essentially, control theory contends that an object of observation is controllable if it can be manipulated to change its state from the initial to the desired one [1]. Furthermore, the applied control theory used in engineering defines controllability as the ability to transfer the system state from one state to another with an appropriate input in the finite time [2]. Applied to our situation, the controllability method should help us to construct the affiliation networks of our E-Learning participants so that the communication between trainees and trainers / tutors is enhanced – to transform our initial affiliation network into an efficient communication network.

II. Optimization Process

We have recognized in our research until now that trainees whose position (trainees' node degree in the graph representing the particular social network) in the communication network is "low" are likely to achieve weak results in the learning outcome. Therefore, the early prediction of such problematic nodes in the network and early prevention, i.e. putting such trainees into the tutored groups according to their most significant affiliations (as recognized for our purpose so far), is the topic of this paper. Intuitively, the matching of the trainee node with the tutored group according to its characteristic affiliations has to be applied. The matching in the social network represented by the graph [15] can be done according to the well-known, useful algorithms for matching in the arbitrary graph [3]. In our case, applying the matching algorithms is even more simple, the tutored groups and the trainees forming a bipartite graph that represents the affiliation network.

Furthermore, a closer investigation shows similarities with the controllability of complex networks approach [8], since this approach also utilizes matching algorithms in order to investigate the controllability of the networks. We show in this research paper that controllability enhancement methods are also useful in the case of E-Learning, especially when dealing with a complex industrial environment.

III. Related Work

The controllability of complex networks has been investigated in several research areas, however, the research that we would like to emphasize here was carried out by Liu et al. in [8]. It focused on applying matching algorithms to arbitrary graphs that provide the possibility to control a particular network in an efficient way. If the number of maximum matchings in the network increases, the number of so-called driver nodes, which are used to control the information flow through the network, tends to be minimized.

Applied to our case, this means that the communication network resulting from E-Learning course participation can be more efficient if the affiliation network for the grouping

978-0-7695-4702-2/12 $26.00 © 2012 IEEE
DOI 10.1109/ICALT.2012.209

of participants into tutored groups is made under the rule of maximal matching in a graph. Since the affiliation network forms a bipartite graph, the existence of maximum matching is guaranteed. Our research differs from other research in this field because it is not limited to either organizing E-Learning participants in virtual classrooms or grouping the participants into implicit social networks [4]. Even research attempts that investigate grouping E-Learning participants into social networks according to some behavioral rule [6] differ from our approach because they fail to investigate whether the learning outcome is improved [11].

Applying a matching algorithm to prevent the isolation of trainees with potentially weak results – by putting the particular trainee node into the corresponding tutored group – is one step beyond the research carried out until now in the E-Learning field on the influence of social networks. Our specific approach is called "engineering social networks" by using the controllability method.

IV. Research Method

The research presented in this paper follows the research agenda that we proposed in our previous work [12]. It investigates the best way to implement the algorithm for matching trainees with their appropriate tutors and groups. The implicit social network of the E-Learning participants is given as follows:

1) *Affiliation networks:*

a) *Language affiliation*

b) *Business branch (industry sector) affiliation*

c) *Geographical affiliation (country)*

d) *Organisational affiliation*

2) *Communication network:* This network is formed as a result of the communication between trainees and tutors. It emerges during the E-Learning course.

3) *The new network:* We add a third network, which is also an affiliation network (as indicated by its structure), but is actually the result of the engineering of the affiliation networks given in *1)*. This network is represented by the bipartite graph composed of the tutored group of participants on the one side, and trainees on the other side (Figure 1).

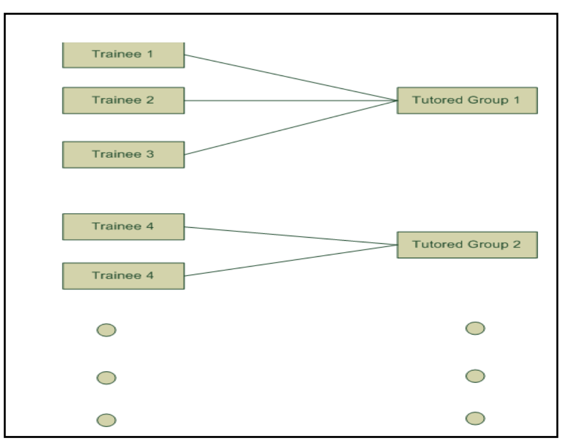

Figure 1. Bipartite gpaph of trainees and corresponding tutored groups.

To each tutored group, a trainer (tutor) is assigned. Within the tutored group, the graph has a so-called "star" structure – all nodes in the group are linked to the tutor (Figure 2).

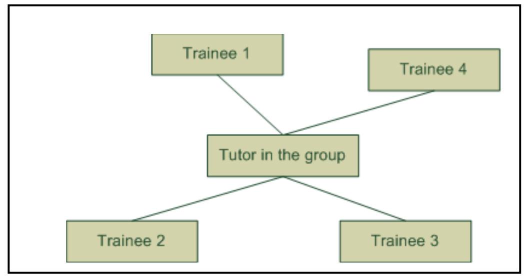

Figure 2. “Star” structure in the group.

Depending on the number of course participants, the way to group the trainees into balanced groups is modelled.

In this paper, we focus on the construction of the third, new network. Its construction is the result of applying the matching algorithm to the given bipartite graph. The links that are processed during the matching are the language, branch, country and organizational affiliation links, taking into account the categorization described in the next section. The target nodes for matching are the new trainees who have not yet been assigned to a tutored group. As regards the characteristics of the tutored group, the affiliations of the tutor are considered. They are compared with those of the trainee to detect their concordance. As regards the categorization of the affiliation, we have decided to give more weight to the language and branch affiliations since our experimental results in [12] have indicated that these affiliations are correlated with the learning outcome more strongly than others.

B. *Matching Method*

The matching categories are listed in the table below.

TABLE I. Matching Categories

Cat. No.	Categorization		
	Category	***Relation***	***Description***
1	Spoken langauge	n:m	A trainee may speak one or more languages
2	Business branches	n:m	A trainee may have expertise in one or more of the business branches
3	Country	n:1	A trainee lives/works in a single country
4	Organizational Unit	1	Free text containing the organizational unit the trainee works for

C. *Matching algorithm*

Each trainee is individually matched. The trainee is matched against all trainers in the system and is assigned to the candidate with the highest concordance.

1) *Calculate concordance individually for each category*
2) *Compute the score for each of those categories by considering their weights*
3) *Sum up the scores*
4) *Assign the trainee to the tutor with the highest score*
5) *If the tutor already has 20 participants in the tutored group, his/her node is excluded from the matching*

D. *Calculating concordance (single category)*

The concordance is calculated with the formula (1).

$$C_j = |l_i| \cdot w_c \,. \tag{1}$$

Where C_j denotes concordance, $|l_i|$ the number of links between j-th tutor and i-th trainee (for $i=1,..,n,$ where n is the number of trainees, and $j=1,..,m$, where m is the number of tutors), and w_c the weight of category, defined in the next section.

E. *Weighting*

Each category is weighted differently, since it is more important for some categories to be concordant than for others. The most important attribute of concordance is the language, followed by branches. The weight is calculated upon the correlation coefficients given as a result of our previous research paper [12]. We will briefly mention the methods that were used in [12] and provide the short overview of the results. We calculated Pearson product-moment correlation coefficients (PPMCC) between quantified characteristics of trainees' nodes in the network (trainees' node degree) and the learning outcome ratio (the ratio between the test results of the trainee before and after taking the E-Learning course). This told us about the existing of a linear dependence between trainees' positions in the network and their learning outcomes. In the experiment, we investigated the learning outcomes of 18 trainees tutored by 8 tutors. The trainees made an assessment before and after taking the E-Learning course. The aim was to reach the score of 75% or above. The ratio between the test result *after* the course and *before* the course was used as the learning improvement indicator (the learning outcome ratio). This indicator varied from 1.041 to 1.587, having an average of 1.191. An interesting fact was detected: 16 of 18 trainees scored 75% and above. The two trainees with the weakest test results also had the lowest learning outcome ratio (1.063 and 1.066) and, more importantly, 0 node degree in the outbound communication network, i.e. these trainees were not communicating with the tutor. After the calculation of PPMCCs, the following could be noted: between the affiliation networks, the language affiliation (number of commonly spoken languages of a trainee) had the highest coefficient of 0.408. The highest correlation coefficient of all networks with the learning outcome ratio was the one calculated for the outbound communication network node degree of trainees: 0.595

Since language affiliation had a correlation coefficient with the learning outcome of 0.408, and business branch affiliation was correlated with the coefficient of 0.084, their ratio (correlation coefficient of language affiliation divided by correlation coefficient of business branch affiliation) is 4.857, rounded to 5 to keep it simple for further calculation. Country (geographical affiliation) and organizational unit (organizational affiliation) are less correlated than the above two, but are included with a weight of 1 – for the cases where the first priority criteria (the first two categories) do not match.

TABLE II. WEIGHTING OF CATEGORIES

Weighting		
Category	***Weight***	
Langauge	10	
Business branch	2	
Country	1	
Organizational Unit	1	

V. PRACTICAL SETUP

In the industrial environment for the E-Learning system that we are observing [10], [9], the implicit social network has already been formed. Our next step is the implementation of the algorithm for the appropriate matching of trainees with tutored groups in the environment mentioned above.

A. *Example*

The example is given from the railway industry where business branches are denoted with:

- P = passenger traffic
- C = freight traffic (C for cargo)
- I = infrastructure

RCA (RailCargo Austria), SZ-Infrastructure (Slovenske železnice - Infrastruktura) and ÖBB-Infrastructure (Österreichische Bundesbahn Infrastrukturbetrieb) are the names of the particular companies used for the example.

TABLE III. EXAMPLE SITUATION

Category	**Given**		
	Trainee	***Tutor1***	***Tutor2***
Spoken language	German, English	English, Slovenian, Italian	German, English, French
Business branches	C, I	P, C, I	P, I
Country	Austria	Slovenia	Austria
Organizational Unit	RCA	SZ-Infrastructure	ÖBB-Infrastructure

The concordances of affiliations between the trainee and the two tutors is shown in the next two tables,where they can be compared. Concordance is calculated according to the formula (1), per category.

TABLE IV. SCORES FOR CONCORDANCE WITH TUTOR1

Category	Trainee vs. Tutor 1		
	Concordance	*Weight*	*Score*
Spoken language	1x10	10	10
Business branches	2x2	2	4
Country	0	1	0
Organizational Unit	0	1	0
Total score			**14**

TABLE V. SCORES FOR CONCORDANCE WITH TUTOR2

Category	Trainee vs. Tutor 1		
	Concordance	*Weight*	*Score*
Spoken language	2x10	10	20
Business branches	1x2	2	2
Country	1	1	1
Organizational Unit	0	1	0
Total score			**23**

In the example above, the Trainee is assigned to Tutor 2, according to the criteria given for the matching. However, if Tutor 2 is "saturated", i.e. his group already has 20 participants, Tutor 1 can still be used, if he/she is still not "saturated". If the new trainee node remains unmatched according to the criteria named above, an early recognition of the trainee with a potentially weak learning outcome is carried out.

B. Evaluation Results

This particular algorithm is applied "in the background" when assigning a trainee to the tutored group in the system. We have recently started applying this approach to the system that we are observing, thus this is work in progress. For the time being, the results show a significant enhancement of the learning outcome of the trainees compared to the previous time period.

TABLE VI. LEARNING OUTCOME PROGRESS

Time period				
	Number of tests	***Using tutoring***	***Using tutored group matching***	***Average learning outcome***
Apr 2009 – Oct 2009	8	No	No	55%
Jan 2010 – Aug 2011	91	Yes	No	81%
Sept 2011 – Jan 2012	70	Yes	Yes	88%

The learning outcome is measured on the test score of the trainees, given in percent. In Table VI we provide a comparison of the average learning outcomes (i.e. average test scores) at different periods of time, and using different E-Learning tutorship methods. In the beginning (from April 2009 to October 2009) no tutorship was provided. The trainees were supposed to prepare for the face-to-face training by reading the learning material in the E-Learning environment. The result was alarming, and it incited us to provide tutorship. In the year 2010 we coached 8 certified trainers (tutors) – these were former trainees with a test score greater than 75%. The new trainees were assigned to the tutors without any specific rule, simply following a personal decision by the E-Learning administrator. The average results were improved; however, there were still some trainees with weak results. In Table VII we show a comparison of the weakest test scores and the test results under the "watermark" of 75% between two "tutored" periods with different tutorship approaches.

TABLE VII. COMPARISON OF "WEAK" RESULTS

Time period				
	Number of tests	***Weakest score achieved by a trainee (in percent)***	***Number of tests with a score under 75%***	***Number of tests with a score under 75% in percent***
Jan 2010 – Aug 2011	91	25%	27	29,67%
Sept 2011 – Jan 2012	70	33%	13	18,57%

The investigation in [12] motivated us to implement the trainee – tutored group matching approach described in this paper. The main objective has been achieved: the average value of learning outcomes has increased, as shown in Table VI. Thanks to this new approach, the trainees are able to achieve better test results, i.e. to learn more from the particular E-Learning material by utilizing their position in the social network or, in other words, by putting them into the appropriate position in the network produced by the trainee – tutored group matching process. However, an interesting discovery was made, which has led us to change some of the quantifiers we used for the PPMCC calculation in [12] to investigate the dependencies between the position of the trainees in the communication network and their learning outcomes. We used there the trainees' outgoing communication node degree (the number of directed links from the trainee to the tutor in the communication network, when the network is represented by a directed graph). For that purpose, we counted the communication attempts of the trainee with the tutor (emails, forum entries directed to the tutor). However, a precise investigation of the behavior of trainees in the new network setup showed fewer communication attempts by the trainee directed to the tutor, compared to the number of communication attempts in the experimental setup given in [12]. In Table VIII we compare the average value of outgoing degree of the trainee as a node in the communication network in [12], which was calculated for a sample containing 18 trainees, with a sample of the same size taken out from the results provided in the latest time period (September 2011 to January 2012).

TABLE VIII. COMMUNICATION ATTEMPTS COMPARISON

Time period			
	Sample of trainees (trainee count)	***Tutored groups***	***Count of communication attempts pro trainee (average)***
Jan 2010 – Aug 2011	18	0	1,8
Sept 2011 – Jan 2012	18	6	0,5

Intuitively, the following explanation for this phenomenon can be provided: the trainees are now grouped into tutored groups where they have direct access to the common forum "hosted" by the tutor, targeted to the trainees with similar (concordant!) characteristics. One entry in the forum by one trainee is handled by the tutor and the feedback is evaluated by the targeted group. Hence, the whole group benefits from a single communication attempt by one of the group's members. This decreases the outgoing communication node degree of the trainee *without* damaging the learning outcome. Nevertheless, forum participation is the precondition for getting the appropriate information. Therefore, the new parameter for correlation calculation should be attendance at the forum (confirmation of receiving the information / tutor feedback); this should replace the outgoing communication node degree parameter.

VI. CONCLUSIONS

The controllability of complex networks approach can be applied to E-Learning social network setups of trainees and trainers/tutors affiliation networks. This is not without consequences – the results of the training can thereby be influenced. Beyond showing that controllability methods are useful for our research, we will continue in the following direction: generalizing the matching approach and matching criteria. That is to say, if the learning outcome shows that some parameters are more significant than others, we will use them as criteria for matching (weighting the graph). Additionally, the early prediction of weak results through regression analysis [7] by using the affiliation and communication network (or, more recently, the activity in the tutored group, such as attendance at the corresponding forum, as mentioned in the previous section) correlation to the learning outcome as prediction coefficients should be investigated. Early prevention (through the application of the matching to the social network) can also be achieved after investigation of the behavior of the prediction coefficients in the regression analysis. If the regression analysis shows that there are better indicators/predictors, we can take these over in the weighting and matching. If new predictors do not deviate much from the average score, we can continue using them for the matching. This will be the next step of our research.

ACKNOWLEDGMENT

The E-Learning system used in the experiments is provided, supported and hosted by RailNetEurope, Vienna, Austria. The contribution to the E-Learning training and testing experiment was provided by Forum Train Europe (FTE), Bern Switzerland. The author thanks both these organizations for their kind support.

REFERENCES

[1] E. D. Sontag, "Mathematical Control Theory: Deterministic Finite Dimensional Systems". Second Edition, Springer, New York, 1998.

[2] K. J. Åström and R. M. Murray, "Feedback Systems: An Introduction for Scientists and Engineers", Princeton University Press. 2008.

[3] J. A. Bondy and U.S.R. Murty, "Graph Theory", Springer, New York, 2008.

[4] C. Haythornthwaite, "Social Network Methods and Measures for Examining E-learning", Social Networks, Citeex:10.1.1.135.6993, 2005.

[5] M.A. Chatti, M. Jarke, and D. Frosch-Wilke, "The future of e-learning: a shift to knowledge networking and social software", Int. J. Knowledge and Learning, Vol. 3, Nos. 4/5, 2007, pp.404–420.

[6] Z. Wang, L. Li, "Enable Collaborative Learning: An Improved E-Learning Social Network Exploiting Approach", Proceedings of the 6th WSEAS International Conference on Applied Computer Science, Hangzhou, China, April 15-17, 2007.

[7] A. Gelman, J. Hill, "Data Analysis Using Regression and Multilevel / Hierarchical Models", Cambridge University Press, 2007.

[8] Y. Liu, J. Slotine, A. Barabasi, "Controllability of complex networks", Nature Journal, Vol 473, Page 167-173, Nature Publishing Group, a division of Macmillan Publishers Limited, 2011.

[9] S. Maglajlic, D. Helic, C. Trattner, "Social Networks and eLearning: New Model for Learning at Workplace", Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces, Cavtat / Dubrovnik, Croatia, 2010.

[10] S. Maglajlic, D. Helic, "Integrating E-learning into work processes in industrial settings: a case study", Proceedings of the 9th international conference on Information technology based higher education and training (ITHET), Cappadocia, Turkey, 2010.

[11] S. Maglajlic, "On the Importance of the Impact Analysis of Social Network Methods in Elearning", In Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (pp. 749-752). Chesapeake, VA: AACE, 2011.

[12] S. Maglajlic, D. Helic, How do social networks influence learning outcomes? A case study in an industrial setting, IADIS International Conference WWW/Internet 2011, Proceedings, 2011, Page 203-213.

[13] G. Conole and J. Culver, "Cloudworks: Social networking for learning design", Australasian Journal of Educational Technology, 25(5), 2009, pp. 763–782.

[14] C. Dalsgaard, "Social software: E-Learning beyond learning management systems", European Journal of Open, Distance and E-Learning, 2006.

[15] S. Wasserman and K. Faust, "Social Network Analysis: Methods and Applications", Cambridge University Press, New York, USA, 2009.

3.3. Paper 3: Efficiency in E-Learning: Can Learning Outcomes Be Improved by Using Social Networks of Trainees and Tutors?

This paper was presented at the IEEE ICL 2012 conference in Villach, September 2012 and published under the same title in the proceedings. In October 2012 the first author was invited to extend the paper and submit the extended version to Addleton Academic Publishers Journal. It is the extended version that is provided in this thesis.

In this paper we realized that the social network engineering method for the construction of tutored groups – that relied on the concordance of the social network parameters of tutors and trainees – only partially succeeded in enhancing collaboration between E-Learning participants. It provided a static method for the detection of potentially weak trainees and for placing them into the corresponding social network – where such trainees could be tutored more efficiently. After an evaluation of the experimental data, we noted that the trainees who had participated in the tutored group programme enjoyed slightly better learning outcomes than the trainees who were tutored in the "old" way (as originally provided by the E-Learning system). However, the trainees who did not participate in the tutored group had a very weak learning performance. These comparisons were made by using the statistical *t* tests.

These results indicated us to that a dynamic method for the detection of potentially weak trainees during the E-Learning course needed to be invented. For this purpose – discovering alternative ways to increase collaboration – we examined the other SNA methods that could be used. We proposed to extend our SNA methods to the following items:

- Recognition of cliques
- Checking centrality
- Checking density.

We applied the PPMCC calculation to these new SNA quantifiers and evaluated the correlation between them and the learning outcomes of the trainees. The main findings were as follows:

- The cliques are automatically created through the construction of tutored groups in the observed social network
- PPMCC of density of the tutored group has a medium correlation with the learning outcome of the trainees of the group, therefore it can be used for the dynamic detection of the tutored groups that contain trainees with a potentially weak learning outcome.

Economics, Management, and Financial Markets
Volume 7(4), 2012, pp. , ISSN 1842-3191

EFFICIENCY IN E-LEARNING: CAN LEARNING OUTCOMES BE IMPROVED BY USING SOCIAL NETWORKS OF TRAINEES AND TUTORS?

SEID MAGLAJLIC
seid.maglajlic@rne.eu
RailNetEurope, Vienna

ABSTRACT. This article investigates improvements in the learning performance of trainees involved in an E-Learning programme in an industrial setting. To measure efficiency in E-Learning, our research develops a learning improvement indicator, defined as the difference in the learning outcome of a trainee before and after a certain period of time. The learning outcome, concretely, is measured by the test results of the trainees. In this regard, the test results should be evaluated both before and after taking the E-Learning course and / or before and after applying a specific action in the E-Learning environment. The experimental results in our industrial setting have indicated that there is room for improvement in the trainees' learning outcomes. We investigated the social networks that are implicitly formed by trainees and trainers in the E-Learning environment. The position of the trainee in these implicit social networks of the E-Learning setup proves to be correlated with the learning outcome. Methods that can improve the learning outcome by using the implicit social networks of trainees and tutors, as well as using social network engineering in order to influence the learning outcome - by placing the trainee in the appropriate position in the social network (i.e. by 'manipulating' the social network) – are currently being investigated. In this research we provide an overview of the results gathered until now and propose a framework for future discussion.

JEL Classification:

Keywords: E-learning, social networks, learning outcome, industrial setting

1. Introduction

The efficiency of E-Learning has been a topic of research for some time already. Several related aspects have been investigated such as: cognitive load theory (Clark et al., 2006), rapid dynamic assessment (Kaluga and

Sweller, 2005), and adaptive and collaborative learning (Ruiz et al., 2006), to name just a few. For our research purposes, efficiency in collaborative learning is especially interesting. The idea of collaborative learning in an E-Learning setting has also been recognized by some authors as a field where social network analysis (SNA) has its place (see (Haythornthwaite, 2005, Chati et al., 2007)). In our related research (Maglajlic et al., 2010) we applied the social network analysis approach to a setup of E-Learners in an industrial environment.

The history of the problem that we focus on in our research is as follows. The users of the E-Learning system that we are observing are not students from school or university. They are employees of several companies geographically distributed across Europe. They are supposed to participate in a new business process and to employ a completely new workflow support tool that they are not used to.

Teaching such E-Learning participants is a new challenge compared to teaching students in a classic educational environment. Discussion between students and teacher in the (virtual) classroom is quite a natural thing, as well as the utilization of discussion results leading to improved learning. However, in an industrial setting, collaboration between E-Learning participants cannot be taken for granted, as we will see in our experimental results further in this paper. The participants may differ widely in their skills and competences, or in their working experiences. This might affect their personal attitude as regards asking the tutors for help. Furthermore, one has to analyze whether their organizational, business sector-related, geographical or language-related background play a role in their learning process. Finally, a typical industrial, economic-efficiency question has to be considered: has the workforce's knowledge improved by using the particular E-Learning system, i.e. is there a benefit if a particular approach is applied? What if some of them haven't made any progress? How can we prevent the stagnation of some trainees, even if that does not affect the majority?

In other words, in an industrial setting, one has to provide simple and clear answers to these questions quite quickly. Therefore, our intention is:

- To measure learning improvement by comparing the testing results of the trainees in the different time intervals and after applying certain methods designed to enhance collaboration between participants.
- To observe the behaviour of trainees during the collaboration and its relationship with the learning outcome.
- To detect the potentially weak trainees early enough in order to be able to help them during the learning process.

A special emphasis is laid on the implicit and explicit social networks of trainees and tutors in our particular industrial setting. The investigation

of the impact of social networks on the learning outcome is ongoing and the current results of our research on this topic are given in this paper.
The rest of the paper is organized as follows. In Section 2 we describe our related research achievements until now. In Section 3 we propose a method for further investigation. In Section 4 we describe the experimental setup we used in our research. We discuss the results of our experiments in Section 5 and provide the statistical sample *t* tests that should help us underpin the conclusions about the methods applied. We conclude with Section 6.

2. Related Research

During the research carried out in the previous period, we reached the four milestones described below.

2.1 E-Learning System Setup

The E-Learning system was chosen according to the criteria relevant for the particular business sector and organization (Maglajlic and Helic, 2010). The organization at the centre of the research has more than 35 member-companies spread all over Europe, each of them containing up to several thousands of employees who are supposed to cooperate in various business processes and use some web-based tools that support their business workflows. The E-Learning system had to be provided and configured for one of these tools, which happens to support one of the mission-critical tasks in this particular business sector. The E-Learning system had to satisfy a number of criteria such as: learning path traceability, easy authoring of tests, reporting possibilities on the learning progress and test results, and configuration of attributes for characterization and grouping of trainees.

Hence, the selection and configuration of such a system was neither straightforward nor easy. More details can be found in (Maglajlic et al., 2010).

2.2 Detection of Implicit Social Networks

Implicit social networks in the E-Learning setup were detected (Maglajlic et al., 2010). More precisely, the existence of typical multi-modal social networks (Wasserman and Faust, 2009) could be observed: the relations between the trainees according to their company, country, language affiliation or even knowledge transfer between each other provided a good foundation for the application of SNA (social network analysis).

2.3 Impact Analysis of Implicit Social Networks

The analysis of the impact of the position of a trainee and his/her tutor in the implicit social networks started (Maglajlic, 2011, Maglajlic and Helic, 2011). In addition to our research, the results of (Haythornthwaite, 2005, Chatti et al. 2007, Wang et al., 2007) also clearly indicated the existence of social networks in E-Learning. Some of these research attempts, such as (Wang et al., 2007), even proposed the grouping / constructing of the trainees' social network units by applying a behavioral rule. However, an assessment of the impact of the trainees' grouping into social networks on the learning outcome – i.e. the application of SNA methods to the trainees and tutors –was missing. Therefore, in our research in (Maglajlic and Helic 2011) we observed a set of trainees who were tested both *before* and *after* taking an E-Learning course.

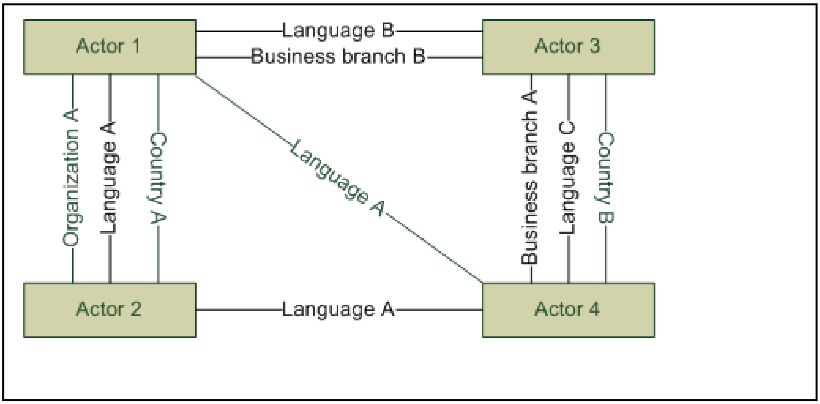

Figure 1: Implicit affiliation network represented as a graph

We proposed a theoretical framework comprising two types of implicit social networks: affiliation networks (language, country, organizational unit and business branch affiliation, see Figure 1) and communication network (inbound and outbound communication between trainee and tutor, see Figure 2).

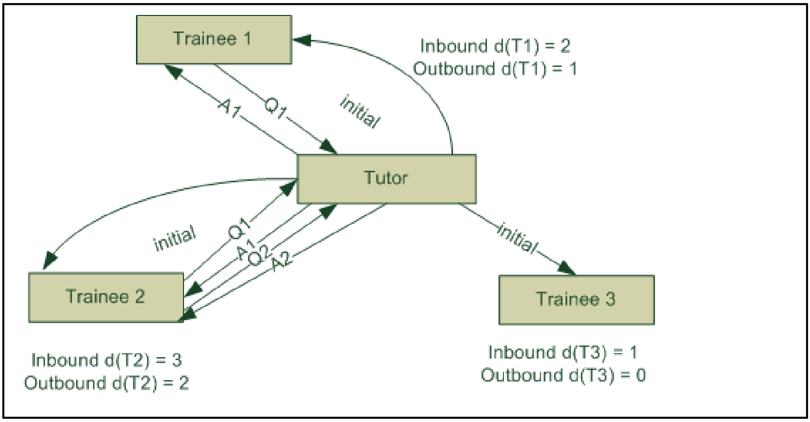

Figure 2 Implicit communication network represented as a graph

We used the basic SNA method, the graph representation of the social networks (Bondy and Murty, 2008, Wasserman and Faust, 2009), and noted the *node degrees* of the trainees and tutors in the affiliation and communication networks. We calculated the correlation coefficients (PPMCC - Pearson Product Moment Correlation Coefficient) between the trainee's node degree in the social networks and learning outcome.

Table 1 Results from the first experiment

1	Number of trainees	18
2	Number of tutors	8
3	Average learning improvement	1.1915
4	Trainees with learning outcome greater or equal to 75%	16
5	PPMCC (language, learning outcome)	0.4095
6	PPMCC (business branch, learning outcome)	0.0841
7	PPMCC (country, learning outcome)	-0.2043
8	PPMCC (organization, learning outcome)	-0.2035
9	PPMCC (outbound communication, learning outcome)	0.5949
10	PPMCC (inbound communication, learning outcome)	0.5558

Table 1 provides an overview of the experimental results which were gathered in (Maglajlic and Helic, 2011). The PPMCC indicates the linear dependency between the social network position of the trainee, measured by *node degree* of the trainee and trainer in the affiliation as well as in the communication network, and the learning outcome. The PPMCC value of the medium significance is the value of the correlation coefficient calculated between the language affiliation (how many languages has a trainee in common with other participants) and the learning outcome (Row

5 in Table 1). The highest correlation was detected for outbound communication of the trainee to tutor and learning outcome (Row 9 in Table 2). In other words, the trainees with a low node degree in the communication network, i.e. the trainees who had not communicated with their tutors and colleagues much, were likely to have weak test results, which means that their knowledge improvement (i.e. learning performance) was weak.

In spite of the fact that it might be intuitively clear that communication between trainees and tutors influences the learning outcome, a sufficiently precise analysis, especially of E-Learning environments in industrial settings, has not yet been provided that we could directly use in our model. A similar attempt to assess the influence of direct dialogue between student and teacher can be found in (Chen et al., 2011). That particular piece of research provides a very precise linear regression model for the prediction of learning outcomes, taking into account the one-to-one dialogue between students and teachers on computer science courses at a university. The experiment was carried out in the classic educational environment, i.e. not within an E-Learning system and not in an industrial environment.

For our present research purposes, we have to provide a precise set of results that can be compared with the results of our future experiments in our specific setup. Our approach seeks to analyze the best way to:

- Encourage trainees to communicate more with their tutors.
- Detect potentially weak candidates before the final test is carried out.

In the next section we describe the algorithmic method used to support the objectives mentioned above.

2.4 Engineering of Social Networks

The algorithm for social network engineering, i.e. for the placing of trainees into the appropriate tutored groups, has already been introduced (Maglajlic, 2012). The algorithm is modeled so as to discover the concordance between trainee and tutor based upon their affiliation network characteristics: the trainees and tutors are grouped according to their spoken languages, business branch affiliation as well as their geographical and organizational affiliations. By applying this method, the initial implicit social network is transformed into a network of tutored groups which can be seen as collaborative clusters (cliques), i.e. the new network is specifically designed to enhance communication between trainees and tutors. The tutored groups created by the algorithm are focused on communication within dedicated (corresponding) discussion forums.

The algorithm basically looks as follows:

- It starts when a new trainee is assigned to a course.

- The concordance between all course tutors and their trainees is checked.
- The concordance score between trainee and tutor is calculated according to the weighting categories presented in Table 2. If the trainee and the tutor share the same category, the category weights are added up to give the concordance score.
- The trainee is assigned to the tutor with the highest concordance score, if this trainer has less than 20 trainees in his/her group.

Table 2 Weighting of Categories

Weighting		
Category	***Weight***	
Language	10	
Business branch	2	
Country	1	
Organizational unit	1	

The category weights are derived from the results showed in 2.3 – since the language affiliation had the strongest correlation of all affiliations with the learning outcome, it received the highest weighting. The correlation coefficient value for business branch showed in Table 1 had approximately a five-fold lower value than that for language affiliation, therefore it was given the corresponding lower value (10 for language, 2 for business branch). The last two categories correspond to the affiliations that had negative correlation coefficients, and we use them with the lowest weighting: 1.

If the algorithm does not find the appropriate tutor for the trainee (or e.g. the trainee is assigned to a tutor with a score inferior to 10), according to our findings in Section 2.3 this means that we have detected a candidate with a potentially weak learning outcome, since communication of the trainee with any of the tutors is very limited. The experimental results, after applying the algorithm to the observed industrial E-Learning system, have shown the following (Table 3):

Table 3 Learning outcome improvement

	Number of tests	***Using tutored groups***	***Average learning outcome***
Jan 2010 – Aug 2011	91	No	81%
Sept 2011 – Jan 2012	70	Yes	88%

- The average learning performance of all trainees has increased compared to that obtained during the time period before applying the algorithm.
- Surprisingly, the outgoing communication node degree of the trainees belonging to the tutored group – which indicates their communication intensity – has not increased (in fact it has decreased, see Table 4).

Table 4 Communication Attempts Comparison

Time period	***Number of trainees in the sample***	***Tutored groups#***	***Count of communication attempts pro trainee (average)***
Jan 2010 – Aug 2011	18	0	1.8
Sept 2011 – Jan 2012	18	6	0.5

The improvement of the learning outcome after grouping the trainees as a consequence of social network engineering supports the claims expressed in (Trausan-Matu et al., 2012) and (Scardamalia, 2002), where the authors assert that knowledge advancement can be achieved as a community rather than as individual achievement.

To conclude this section, let us state that we succeeded in detecting the potentially weak candidates with this algorithm; however, communication intensity, which we hoped to boost, did not increase. This finding motivates us to carry out further research, which we wish to discuss in this paper.

3. Methods

In this section, we tackle the surprising phenomenon described in Section 2.4, namely that by grouping the E-Learning participants into tutored groups, we experienced an overall improvement of the learning performance (learning outcome increased, according to Table 3), but communication intensity between trainees and tutors did not increase, contrary to our expectations. The following facts can be observed:

- The matching algorithm for social network engineering depicted in Section 2.4 provides us with a *static* method for the recognition of potentially weak candidates
- A *dynamic* method for such recognition is not yet available, i.e. a method that would identify potentially weak candidates *during* the E-Learning course, before the examination, is lacking.

The hypothesis that the early detection of trainees with a weak learning outcome can be based on discovering a low communication intensity of those candidates with their tutors, which could be concluded from the description of the results in 2.3, has become harder to hold. Therefore, more investigation is needed on this topic.

For this purpose we have decided to apply additional tools of the SNA approach: we try to quantify the social networks of the E-Learning participants by applying SNA quantification methods. We would like to use this to look for further dependencies with the learning outcome in our experimental setting. In Sections 2.3 and 2.4 we used only the node degree from the SNA methodology (as mentioned in Section 2.3); however, for a deeper analysis of the phenomenon mentioned in Section 2.4, some other SNA tools may be useful, in addition to the node degree method, for example:

- *Cliques*: the aim of this method is to detect sub-networks inside a particular social network, e.g. concentrated communication between a group of nodes – in such a clique the nodes are more connected with each other than with other nodes of the network.
- *Centrality*: the centrality of an actor in the social network is calculated by dividing the node degree by the total number of edges in the network.
- *Density*: The density of a particular social network is calculated as the total number of edges in the network divided by the maximum number of edges possible for this network.

In the work (Silva and Figueira, 2012), the authors used the above-mentioned SNA methods (including *node degree* and *degree centrality,* which indicates the dependency of a network on a node) to analyze the interactions between students and teachers, i.e. for the graphic representation of these interactions. We use this method for a similar purpose, but we differ in the implementation and evaluation of the results from the conclusions given in the above-mentioned work. The authors of (Silva and Figueira, 2012) observed the online forum of teachers and students, and described how they visualized this interaction as a social network graph. Within that research, the authors applied the above-mentioned SNA methods in order to check the strength of the relations within a community, as well as to detect the members who played a central role in the network. The authors use so-called "stop words" to exclude irrelevant interactions from the communication network. More precisely, the messages containing sentences such as "Yes", "I agree" or "Thank you" were considered as irrelevant and excluded from the network.

However, we do not apply such a filtering in our research, since we rather support the ideas expressed in (Trausan-Matu et al., 2012) and (Scardamalia, 2002) about the socio-cultural paradigm of learning, as we already mentioned in Section 2.4. Hence, we interpret all communication attempts as equally valuable.

3.1 Cliques

Actually, since we apply the grouping algorithm described in Section 2.4 intentionally, the cliques are generated explicitly. Therefore, in our case, the detection of cliques is very easy: the tutored groups *are* our intermediate clusters, the cliques. We can observe the communication between trainees and the tutor within the dedicated discussion forum of the tutored group. Hence, our cliques are transformed into communication network clusters.

The postings in the discussion forum are counted as communication attempts from trainee to trainer and vice versa. In the terms of the graph theory applied to SNA, we observe the tutored group clique as directed graph (digraph) whereby the direction of the arcs in the graph is defined by the direction of the posting in the discussion forum between trainee and tutor.

3.2 Centrality

For the purpose of our investigation we slightly modified the standard approach to centrality:

- The calculation of centrality is provided as the ratio of the node degree of a tutor within the clique to the total number of all arcs generated in the whole network (for all cliques).
- On the other hand, the centrality of a trainee is calculated twice: as the ratio of the node degree of a trainee to the total number of arcs in the clique, and as the ratio of the node degree of the trainee to the total number of all arcs in the whole network.

3.3 Density

We have to be careful regarding the density calculation. The standard formula for density calculation for a digraph assumes that only one inbound and only one outbound arc may be considered. However, in our setting, every communication attempt between tutor and trainee is counted; hence, if the trainee has asked the tutor twice about two different issues, this is presented as two arcs from trainee to the tutor. Thus, the maximum number of edges (arcs) for our graph is theoretically infinite.

In such a case, it is advisable to treat our graph as a weighted (valued) graph, and calculate the density according to the rules given for weighted graphs (Wassserman and Faust, 2009). The weight of the arc is the number of arcs in the same direction between two nodes (i.e. the outbound or inbound node degree). In this case, the density is calculated by considering the maximum weight allowed. Hence, we have to make an assumption regarding the maximal number of communication attempts between trainee and tutor. Let us assume that the trainee will not attempt more than 20 times to contact the trainer during a single course. With this upper boundary it is possible to calculate the density as the ratio of the average weight of all arcs in the graph to the maximum weight possible for the graph. We will calculate the density at the clique level.

To sum up, we will calculate the density and centrality as defined above, and compare the values between the cliques, as well as check the dependencies of the learning outcome values on our new SNA quantifiers for the trainees. We will observe the learning outcome of each trainee as well as the average value of the learning outcome at clique level. The dependencies will be calculated, as described in Section 2.3, by using PPMCC.

4. Implementation

We observed a sample of 14 trainees and their behavior in the time period from January 2012 to August 2012 (i.e. in the same way as the sample mentioned in the lowest row of Table 4). They were divided into 4 tutored groups (cliques). 3 trainees were not participating in the tutored groups. As a consequence, the total number of trainees in the tutored groups is 11 (see Table 5).

Table 5 Cliques

Clique	***Tutor***	***Number of trainees***
1	TU-1	6
2	TU-2	2
3	TU-3	1
4	TU-4	2

According to what we stated in Section 3, each clique may be viewed as a communication network represented as a directed graph, with multiple arcs between nodes for the centrality calculation and a directed weighted graph (weighted but only one ingoing or outgoing arc between the nodes is possible) for the density calculation.

First, we calculated the centrality coefficients of tutors per clique (i.e. per his/her tutored group)

Table 6 Centrality of Tutors

Tutor	*Centrality*
TU-1	0.121
TU-2	0.061
TU-3	0.091
TU-4	0.091

Second, we calculated the centrality coefficients of trainees at the clique level.

Table 7 Centrality of Trainees at the Clique Level

Trainee	*Clique*	*Outbound node degree*	*Centrality*
T1	1	4	0.571
T2	1	2	0.286
T3	1	0	0.000
T4	1	1	0.143
T5	1	0	0.000
T6	1	0	0.000
T7	2	0	0.000
T8	2	4	1.000
T9	3	4	1
T10	4	4	0.667
T11	4	2	0.333

Third, we calculated the centrality coefficients of trainees at the general network level.

Table 8 Centrality of Trainees on General Network Level

Trainee	*Outbound node degree*	*Centrality*
T1	4	0.121
T2	2	0.061
T3	0	0.000
T4	1	0.030
T5	0	0.000
T6	0	0.000
T7	0	0.000
T8	4	0.121
T9	4	0.121
T10	4	0.121
T11	2	0.061
T12	0	0.000
T13	0	0.000
T14	0	0.000

Fourth, we calculated the density per clique.

Table 9 Density of the Cliques (Tutored Groups)

Clique	***Density***
1	0.046
2	0.15
3	0.35
4	0.225

In Table 10 we show the learning outcome values of the trainees as well as their tutored group index.

Table 10 Learning Outcome

Trainee	***Tutored group (clique)***	***Learning outcome***
T1	1	94
T2	1	88
T3	1	98
T4	1	90
T5	1	92
T6	1	80
T7	2	91
T8	2	93
T9	3	92
T10	4	100
T11	4	99
T12	Did not participate in the tutored group	62
T13	Did not participate in the tutored group	69
T14	Did not participate in the tutored group	58

As expected, the trainees that did not participate in the tutored group (T12, T13 and T14) had the weakest learning outcome.

In Table 11 we show the PPMCC calculated for (*i*) the centrality of the trainee at the clique level and the learning outcome of each trainee, (*ii*) the centrality of the trainee at the general level and the learning outcome of each trainee, (*iii*) the centrality of the tutor and the learning outcome of each trainee, (*iv*) the centrality of the tutor (at the general level of the whole graph) and the average learning outcome of each clique, (*v*) the density of each clique and the average learning outcome per clique, (*vi*) the density of each clique and the corresponding trainee's learning outcome.

Table 11 PPMCC Calculation

	PPMCC	***Value***
(i)	Centrality of trainee at the clique level, trainee's learning outcome	0.207
(ii)	Centrality of trainee at the general level, trainee's learning outcome	0.298
(iii)	Centrality of the tutor, trainee's learning outcome	-0.259
(iv)	Centrality of the tutor, average learning outcome pro clique	-0.110
(v)	Density of the clique, average learning outcome per clique	0.262
(vi)	Density of the clique, trainee's learning outcome	0.389

5. Discussion

The results shown in Table 11 indicate that the density of the clique has a *medium* correlation (according to the scale given by (Cohen, 1988)) with the learning outcome of each trainee (row (*vi*)) but, compared to other PPMCCs on SNA values (centrality PPMCCs are less than 0.3, suggesting the *low* correlation according to the above mentioned scale), it indicates the strongest correlation with the learning outcome. In spite of the fact that the outgoing node degree is still lower (if we calculate the average of the values shown in Table 8, the value of 1.5) than the outgoing node degree of the trainees in the period before September 2011, it seems that the communication intensity per clique, this time measured by density, does play a role in the impact of the social network activities of the trainees on their learning outcome. Furthermore, if we modify the value of the outgoing node degree in Table 6 from 0 to 1 for those trainees who only participated in the discussion forum but didn't send any messages to the tutor, the average outgoing node degree is getting close to the value we had in Table 4 for the period before September 2011 (see Table 12).

Table 12 Outbound Communication of Trainees (Node Degree)

Trainee	***Outbound node degree***	***Outbound node degree including participation***
T1	4	4
T2	2	2
T3	0	1
T4	1	1
T5	0	1
T6	0	1

Trainee	Outbound node degree	Outbound node degree including participation
T7	0	1
T8	4	4
T9	4	4
T10	4	4
T11	2	2
T12	0	0
T13	0	0
T14	0	0
Average	1.5	1.785

In other words, if we treat the trainees who read the content of the discussion forum as if they participated in the knowledge acquisition process, the results of our analysis explain the phenomenon that we were trying to understand in Section 2.4. Thus every feedback regarding knowledge acquisition made by the trainees is useful. But if we look at this from the system implementation point of view, in order to distinguish between the trainees who read and those who do not read the content of the forum dynamically (i.e. immediately, during the E-Learning course), we realize that this is a more complex task than if we had data about forum participation directly.

Furthermore, in order to analyze the average learning outcome after our experiment and to compare it with the average learning outcome that was achieved before the new methodology (involving specially formed tutored groups) was introduced, we made use of the sample t tests. Firstly, we took a sample of 18 trainees and their test results from the experiment we described in Section 2.3. We compared their results with the learning outcome results shown in Table 10 (14 trainees, of which 11 had participated in the tutored groups).

Table 13 t test results: t=-0.1091, 95% confidence interval <-∞,6.93432), p=0.4569 (>0.05)

Sample taken	*N*	*mean*	*t*
First experiment (Nov 2010)	18	85.667	t = -0.1091
Second experiment (Jan 2012 – August 2012)	14	86.143	

This comparison does not indicate a significant difference in the average learning outcome (Table 13). More precisely, according to the rules of statistical inference, we cannot conclude that the learning outcome after the implementation of the new methods that support collaborative learning is generally higher than before these methods were applied with a significance level of 5% (t=-0.1091, p=0.4569 > 0.05) . However, if we exclude the results of the trainees who did not participate in the tutored groups, i.e. those to whom the new methods have not really been applied, we get remarkably different results. Therefore, secondly, we calculated a t test with the first sample from Section 2.3 of 18 trainees and the sample of 11 trainees who participated in the experiment described in Section 4, i.e. who were involved in the tutored groups (Table 14).

Table 14 The second t test: t= -1.8825, 95% confidence interval <-∞, -0.6462717), p= 0.03529 (<0.05)

Sample taken	***N***	***mean***	***t***
First experiment (Nov 2010)	18	85.667	t = -1.8825
Second experiment (Jan 2012 – August 2012)	11	92.454	

These results of the t test show a difference between the two approaches in a different way. Basically, the values of the t test (t=-1.8825, p= 0.03529 < 0.05) indicate that trainees who were involved in the specially formed tutored group approach were able to achieve a better learning outcome than those who were not using the new methodology - with the significance level of 5%. We have therefore concluded that it is worth trying to improve the methods for supporting collaborative learning within the E-Learning system.

If we wish to use our findings, obtained thanks to this analytical approach, by implementing them in the E-Learning system in the same way that we implemented the algorithm described in Section 2.4, we need a practical method for the quantification of the trainee's characteristics and behavior within the clique. For instance, we could implement a mechanism in the discussion forum of the E-Learning system to notify the trainee about every new posting in the forum and require a confirmation from the trainee that he/she has read the content. This approach would enable us to acquire the data needed for further quantification – i.e. for the calculation of the outgoing node degree of the trainee and a more precise calculation of the density of the clique.

Finally, flowing from this analysis we can propose a method for the dynamic detection of potentially weak trainees:

- The density of the tutored group (clique) has to be checked permanently.

- In the groups with a lower density, the trainees with a lower outgoing node degree should be identified (confirmation messages about reading the content of the forum should also be considered).

Such trainees should be contacted by the tutor (or E-Learning administrator, if needed) and notified that there are contributions in the forum that might help them to improve their learning outcomes.

6. Conclusions

The *efficiency* of E-Learning can be monitored thanks to SNA methods and it may even be influenced with the help of appropriate social network engineering. The early detection of candidates with a potentially weak learning outcome and the prevention of disappointing results can be achieved through:

- Organizing tutored groups by applying the appropriate matching algorithm (static detection),
- Permanent application of SNA methods and corresponding result analysis (dynamic detection).

We will continue our research in this direction (*i*) finding ways to integrate SNA methods directly in the system in a customized way (*ii*) finding additional grouping models (social network engineering methods) which can also produce social networks of trainees (primarily trainees!) who help each other. Within this approach, the ontological parameters of the trainees could be evaluated and utilized for social network engineering purposes.

REFERENCES

Bondy, J. A., and Murty, U. S. R. (2008), *Graph Theory*. New York: Springer.

Chatti, M. A., Jarke, M., and Frosch-Wilke, D. (2007), "The Future of E-learning: A Shift to Knowledge Networking and Social Software," *International Journal of Knowledge and Learning* 3(4/5): 404–420.

Chen, L., Di Eugenio, B., Fossati, D., Ohlsson, S., and Cosejo, D. (2011), "Exploring Effective Dialogue Act Sequences in One-on-One Computer Science Tutoring Dialogues", IUNLPBEA '11, *Proceedings of the 6th Workshop on Innovative Use of NLP for Building Educational Applications*, Association for Computational Linguistics Stroudsburg, PA.

Clark, R. C., Nguyen, F., and Sweller, J. (2006), *Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load*. San Francisco: Pfeiffer.

Cohen, J. (1998), *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: L. Erlbaum.

Haythornthwaite, C. (2005), "Social Network Methods and Measures for Examining E-learning", *Social Networks*, Citeex:10.1.1.135.6993.

Kalyuga, S., and Sweller, J. (2005), “Rapid Dynamic Assessment of Expertise to Improve the Efficiency of Adaptive E-learning,” *Educational Technology Research and Development* 53(3): 83–93.

Liaw, S. S., Huang, H. M., and Chen G. W. (2007), “Surveying Instructor and Learner Attitudes toward E-learning”, *Computers & Education* 49(4): 1066–1080, 10.1016/j.compedu.2006.01.001.

Maglajlic, S., Helic, D., and Trattner, C. (2010), “Social Networks and eLearning: New Model for Learning at Workplace”, *Proceedings of the ITI*, 32nd International Conference on Information Technology Interfaces, Cavtat, Dubrovnik.

Maglajlic, S., and Helic, D. (2010), “Integrating E-learning into Work Processes in Industrial Settings: A Case Study,” *Proceedings of the 9th International Conference on Information Technology Based Higher Education and Training (ITHET)*, Cappadocia.

Maglajlic, S. (2011), “On the Importance of the Impact Analysis of Social Network Methods in Elearning,” in *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*. Chesapeake, VA: AACE, 749–752.

Maglajlic, S., and Helic, D. (2011), “How Do Social Networks Influence Learning Outcomes? A Case Study in an Industrial Setting”, *IADIS International Conference WWW/Internet, Proceedings*, 203–213.

Maglajlic, S. (2012), “Engineering Social Networks Using the Controllability Approach Applied to E-Learning”, *12th IEEE International Conference on Advanced Learning Technologies*, DOI 10.1109/ICALT.2012.209.

Ruiz, J. G., Mintzer, M. J., Leipzig, R. M. (2006), “The Impact of E-Learning in Medical Education,” *Academic Medicine* 81(3): 207–212.

Scardamalia, M. (2002), “Collective Cognitive Responsibility for the Advancement of Knowledge,” *Liberal Education in Knowledge Society*, Ch. 4, 67–98.

Silva, A., and Figueira, A. (2012), “Visual Analysis of Online Interactions through Social Network Patterns,” *12th IEEE International Conference on Advanced Learning Technologies*, DOI 10.1109/ICALT.2012.57

Sun, P. C, Tsai, R. J, Finger, G., Chen, Y. Y., and Yeh, D. (2008), “What Drives a Successful e-Learning? An Empirical Investigation of the Critical Factors Influencing Learner Satisfaction,” *Computers & Education* 50(4): 1183–1202, 10.1016/j.compedu.2006.11.007.

Trausan-Matu, S., Dascalu, M., Rebedea, T. (2012), “Computer-Supported Collaborative Learning, Dialogism, Chat, Polyphony, Natural Language Processing,” *12th IEEE International Conference on Advanced Learning Technologies*, DOI 10.1109/ICALT.2012.101

Wang, C., and Li, L. (2007), “Enable Collaborative Learning: An Improved E-Learning Social Network Exploiting Approach,” *Proceedings of the 6th WSEAS International Conference on Applied Computer Science*, Hangzhou, April 15–17.

Wasserman, S., and Faust, K. (2009), *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.

© IEEE Catalog Number: CFP1223R-USB, ISBN: 978-1-4673-2426-7. This article is an extended version of the paper which was presented and published at the IEEE

International Conference ICL (Interactive Collaborative Learning), Villach, Austria, 2012, September 26–28.

3.4. PAPER 4: IMPLEMENTATION OF A FRAMEWORK FOR COLLABORATIVE SOCIAL NETWORKS IN E-LEARNING

This paper has been submitted to the AACE Journal for E-Learning and was at the time of writing being reviewed by the editorial team of AACE. It contains the final proof-of-concept of the methods applied during our research. It presents an overview of all research results and detailed descriptions of the add-on modules of the E-Learning system that were developed following the specifications of the framework for collaboration enhancement described in Papers 2 and the paper (Maglajlic, 2012) which was published in the proceedings of the conference IADIS 2012 "WWW/Internet" in Madrid, Spain. It is important to mention that the later paper contains the concept of combining SNA and ontology engineering methods to provide an improved collaboration enhancement infrastructure within E-Learning systems. That paper proposes a way to implement an algorithm that takes into account the following elements:

- Social network concordance between the trainees
- Density of the tutored group
- Outgoing node degree of a trainee
- Trainees' ontology containing their knowledge level class
- Learning ontology containing course, lesson, progress and score classes
- Relations between trainees' and learning ontology (relations between single classes across the ontologies).

This algorithm helps to build trainee-to-trainee relationships between E-Learning participants and utilizes the ideas of semantic social networks. It results in the recommendation of those trainees (high achievers) that could help other trainees (with a potentially weaker learning outcome). Finally, in the referenced paper (Maglajlic, 2012) it is recommended to develop the algorithm as the second add-on to the E-Learning system that promotes collaboration within the E-Learning community. The Paper 4 contains once again the description of the basic principles of the proposed algorithm as well as the implementation guideline for the framework for its utilization.

The final comparison of the learning outcomes of the trainees who participated in E-Learning courses *before* the collaboration enhancement add-ons had been implemented and those who were supported by the new collaboration methods in the E-Learning system is provided in the form of a statistical *t* test. The results of the statistical *t* test actually indicate that the learning outcome of the trainees who participated in the E-Learning courses according to the new “standards” is significantly better than that of the participants during the initial period.

Implementation of a Framework for Collaborative Social Networks in E-Learning

Abstract: This paper describes how we have implemented a framework for the construction and utilization of social networks in E-Learning. These social networks aim to enhance collaboration between all E-Learning participants (i.e. both trainee-to-trainee and trainee-to-tutor communication are targeted). E-Learning systems that include a so-called "social component", i.e. enable forming links between trainees and tutors, thereby creating a social network, already do exist; however, sophisticated networking approaches are lacking in these systems. Indeed at the present time, E-Learning systems containing social components do not automatically consider the results of social network analysis (SNA) of a specific dataset of trainees and tutors; neither do they take into account the knowledge level of the trainees according to their learning progress and/or test results. Furthermore, potential ontological relations between the network of participants and their knowledge – for the purpose of interlinking the participants – are, unfortunately, commonly ignored. Therefore, we have created a framework that makes use of the results of SNA concerning the implicit and explicit social networks of E-Learning participants, and their learning outcome, as well as taking into account the relationships between learning and learners' ontology. The result of the implementation of this framework are E-Learning system modules that support special dedicated learning (tutored) groups of trainees gathered around the tutor, and peer-to-peer links between trainees that encourage them to help each other during the learning process.

Introduction

Quite a few research results indicate that knowledge advancement can best be achieved as a community rather than as an individual (Scardamalia 2002, Trausan-Matu et al. 2012). Such results have inspired other researchers to continue investigating specific topics related to collaborative learning. For example, the concepts underpinning collaborative learning are investigated in (Jonson and Johnson 1994, Slavin 1995, Guillies 2004) and also used in experiments such as (Hwang et al. 2012). In the experimental approach described in (Hwang et al. 2012), it has been shown that the student's cognitive load is reduced and their learning outcome is better if computerized collaborative concepts are used.

Hence, recent research results keep confirming that collaborative learning is useful and can be further promoted. Our special emphasis in this paper lays on possible methods to improve collaboration in E-Learning settings. Research results such as those shown in (Silva and Figuera 2012) support the idea of applying the SNA (social network analysis) methodology to collaboration between trainees and tutors. By visualizing the activities of the participants in a discussion forum in the chosen E-Learning setup, the authors show how the social component of E-Learning can be closely investigated.

We started our investigation on collaborative learning by executing several experiments in the industrial setting where the sample was taken – a pan-European organization with member companies all over Europe that share a common business process for cooperation – and analyzed their results. It is important to emphasize that our research and experiments were carried out within an industrial setting; indeed the profile of our E-Learning participants largely differs from the profile of users in typical educational environments such as a school or a university: they are very likely to differ in age, as well as in educational, business, geographical, cultural and language background and skills. Furthermore, the trainees' motivation for learning and collaborating in an industry sector differs from students' motivations in a more classical educational environment. Quite intuitively, one expects that students in a classical educational environment will communicate with their teachers and fellow students; however, trainees in an industrial setting may encounter barriers due to their potentially very dissimilar backgrounds and skills.

In our own research (Maglajlic and Helic 2011) we noticed that knowledge improvement was weak for some users (albeit a minority) because, we supposed, they had been hindered from communicating with other participants for various reasons: poor knowledge of the common language used, business branch diversity, geographical or organizational differences. We investigated these parameters and found undeniable dependencies between some of these parameters and the learning outcomes of the trainees. These findings allowed us to identify the problem more clearly: the learning performance through E-Learning in our industrial setting was weak if communication between trainees and tutors was lacking.

The main objective of this paper is to describe the implementation of a framework for the construction and use of collaborative social networks within an E-Learning environment. The general objective of our research was to discover methods that can enhance collaboration between trainees and tutors, i.e. all E-Learning participants within the specific industrial setting in which the E-Learning took place in order to help their knowledge improvement. For that purpose, firstly we investigated whether there were any social networks of participants, and if so, what did they look like: could one say that there were implicit social networks, if so, which characteristics did they have? Secondly, we asked ourselves the following questions:

- Could we utilize these implicit social networks in order to enhance collaboration?
- In order to construct new social networks (i.e. communication networks), could the participants be encouraged to connect with each other and cooperate?
- What could be the preconditions for that to happen?
- What are the consequences: did the learning outcome of participants improve after they were provided with the facilities to collaborate?

In order to answer these questions, we applied several methodologies such as SNA (Wasserman and Faust, 2009), controllability theory (Sontag 1998, Åström 2008), as well as ontology engineering (Pernas et al. 2012) and statistical independent sample *t* tests. This resulted in a framework that is described in more detail in the next section. Basically, we chose a suitable E-Learning system, modified the system to support the construction of tutored groups (first step of social network engineering), provided the utilities for tutored groups management in the system, and finally implemented a novel method of production of peer-to-peer links between trainees by applying a new algorithm, which combines SNA and ontology rules (ontology engineering).

The rest of the paper is organized as follows. In the next section, we explain the underlying principles of our research framework. In the second and third sections, we describe the implementation of the framework and its functional architecture. In the fourth section, we evaluate the results. In the fifth section, we discuss the achievements of the implementation. We then provide conclusions in the fifth section.

A framework based on previous research

The research framework that we propose is founded on the following research findings. Firstly we analyzed the impact of social networks on the learning outcome of the trainees in the E-Learning environment in an industrial setting (Maglajlic and Helic, 2011). We investigated the existence of implicit social networks of trainees according to their affiliations: geographical, organizational, linguistic, and business sector-related. We applied one of the basic SNA methods: *node degree*. For example, if two trainees speak the same language, they will be related in the implicit social network. The number of such relations of the trainee is calculated as node degree, where the trainee, as an actor of the social network is actually the node in the graph representation of the network, and the above-mentioned relation represents the link in the graph. The number of links going from or to the node in the graph representation is defined as node degree. The other affiliations have been quantified in the same way. Furthermore, during the E-Learning course (i.e. during the learning process), communication between tutors and trainees is measured by counting the communication attempts from trainee to tutor as *outgoing node degree* of the trainee, and communication attempts from tutor to trainee as *ingoing node degree*. The quantification with node degree was used for the calculation of Pearson Product Moment Correlation Coefficients (PPMCC), in order to check the correlation of the values provided by the SNA method with the learning outcome of the trainees. For the quantification of the learning outcome, the test score of the trainee in the corresponding E-Learning course was taken. The results of the experiment on a sample of 18 trainees and 8 tutors are shown in Table 1.

Table 1

The results of the experiment with 18 trainees and 8 tutors

PPMCC	Value
Language, learning outcome	**0.4095**
Business branch, learning outcome	0.0841
Country, learning outcome	-0.2043
Organization, learning outcome	-0.2035
Outgoing communication node degree, learning outcome	**0.5949**
Ingoing communication node degree, learning outcome	0.5558

We noted that the correlation between outgoing node degree and the learning outcome of the trainee was strongest among all implicit social networks (communication network and affiliation network). It was also noted that language displayed the strongest correlation among the affiliation network relations. Hence, the experimental results showed a strong correlation between the communication intensity of the trainees and tutors and the learning outcome of the trainees. This inspired us to try to construct social networks (or, rather to control social network creation) which would serve for communication and collaboration enhancement between tutors and trainees.

Consequently, the second step of our research was the attempt to construct collaborative clusters within the social network of trainees and tutors (Maglajlic 2012). More precisely, the idea was to detect and prevent the emergence of isolated actors in the social network, because we had seen in the previous research phase that lack of communication had an impact on the learning outcome. To make it easier to detect potentially isolated actors (nodes), we applied the experiences gained from control theory (Sontag 1998, Astrom 2008). The newest findings of (Liu et al. 2011) have shown that the detection of those nodes in the network that are critical for the control of the information flow within the network can be efficiently done by applying algorithms for matching in the arbitrary graph (detailed description of matching algorithms in graph theory can be found in (Murty and Bondy, 2008)). Therefore, we developed a special algorithm that connects the trainees and tutors by applying matching rules, thus producing the tutored groups: bipartite graphs of tutor and the trainees assigned to her/him. The tutored groups can be viewed as collaborative clusters since the matching was made by considering the preconditions for communication between actors – such as common spoken language, common business branch, geographical and organizational similarities. The concordance scale looks as follows (Table 2).

Table 2

Weighting of categories

Category	Weight
Language	10
Business branch	2
Country	1
Organizational unit	1

Basically, the algorithm works as follows: when a trainee is added to an E-Learning course, the concordance is calculated for all the course tutors (the weights are summed up) and the tutors' concordance scores are compared. The tutor with the highest concordance score is chosen, and the trainee is assigned to her/his tutored group.

In our third step of the research (Maglajlic and Gütl 2012), we analyzed the communication intensity of trainees and tutors in tutored groups and the learning outcomes of the trainees. We wished to find out whether these collaborative clusters had helped with the learning improvement of the trainees. However, we detected the following paradox: the learning improvement in the period of time since the collaborative clusters had been introduced in the E-Learning environment had increased, but the outgoing node degree of the trainees within the tutored groups had, in average, decreased, contrary to our expectations (Table 3).

Table 3

Communication attempts comparison

Time period	Number of trainees in the sample	Tutored groups#	Count of communication attempts pro trainee (average)
Jan 2010 – Aug 2011	18	0	1.8
Sept 2011 – Jan 2012	18	6	0.5

More precisely, the collaboration, according to the node degree measurements, decreased.

In order to analyze this paradox more closely, we applied additional SNA methods (Wasserman and Faust, 2009):

- *Cliques* – the parts of the network where the interlinking between the nodes is concentrated, i.e. is higher than in the rest of the network. In our case, the cliques are the tutored groups – the sub-networks of trainees gathered around the tutor.
- *Centrality* – the ratio between the node degree of a particular actor in the social network and the total number of links in the network. Since our tutored groups play an important role in the whole concept, the centrality of the trainees and tutors was calculated both at the clique and general levels. The number of communication attempts between trainees and tutors was taken as the number of links between trainees and tutors. Entries in the discussion forum of a given tutored group made by a given trainee and/or tutor were interpreted as communication attempts.
- *Density* – the density of the social network is defined as the ratio between the number of links in the network and the maximum possible number of links in this network. In our case, we calculated the density of cliques (tutored groups), again by interpreting the entries in the discussion forum of the tutored group as communication attempts, i.e. links. We limited the maximum number of communication attempts between two participants to 20 in order to be able to calculate the density. Practically, this number has never been exceeded in the system that we were observing.

We used the quantifiers (centrality and density) in PPMCC calculation again in order to detect the level of dependency between these values and the learning outcomes (Table 4).

Table 4

The results of the experiment with 14 trainees and 4 tutors (i.e. 4 cliques = 4 tutored groups)

PPMCC	**Value**
Centrality of trainee at the clique level, trainee's learning outcome	0.207
Centrality of trainee at the general level, trainee's learning outcome	0.298
Centrality of the tutor, trainee's learning outcome	-0.259
Centrality of the tutor, average learning outcome pro clique	-0.110
Density of the clique, average learning outcome per clique	0.262
Density of the clique, trainee's learning outcome	**0.389**

The analysis showed that, among these quantifiers, density had the strongest correlation to the learning outcome. Furthermore, trainees who did not participate in the tutored groups had the weakest results (Table 5).

Table 5

The test score of the trainees including the outgoing node degree and the number of visits to the discussion forum of the group. Trainees who did not participate in the tutored group did not visit the discussion forum either.

Trainee	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
Tutored group	1	1	1	1	1	1	2	2	3	4	4	Did not participate in the		

Trainee	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
Outgoing node degree	4	2	0	1	0	0	0	4	4	4	2	tutored group		
# of visits to the forum	7	4	4	6	3	4	5	4	5	5	4			
Learning outcome	94	88	98	90	92	80	91	93	92	100	99	62	69	58

Hence, trainees who participated in the common discussions within their groups, even as simple readers, achieved a better learning outcome than those who were isolated. The detailed result overview of this particular experiment can be found in our work (Maglajlic and Gütl 2012). This was an indication for us to continue our investigation towards increasing the density of the cliques (i.e. collaborative clusters = tutored groups).

For this purpose, we combined in our fourth step of the research (Maglajlic 2012) certain findings in the field of semantic social networks and ontology engineering with our own achievements. The construction of semantic social networks is done by applying ontology engineering (Mika 2005, Jung and Euzeant 2007), i.e. through the investigation and detection of the interrelations between ontologies that are used by the actors, or that are used to describe the actors and their actions in a social network. In addition to this research approach, a very detailed research on context-sensitive, situation-aware ontology engineering in E-Learning (Pernas et al. 2012) seemed very promising and we combined its findings with our ideas. This resulted in the modeling of learning and learners' ontologies (see Figure 1), relating these ontologies in an ontology network and exploring the rules needed for the purpose of connecting the trainees as actors in the social network and classes in the ontology.

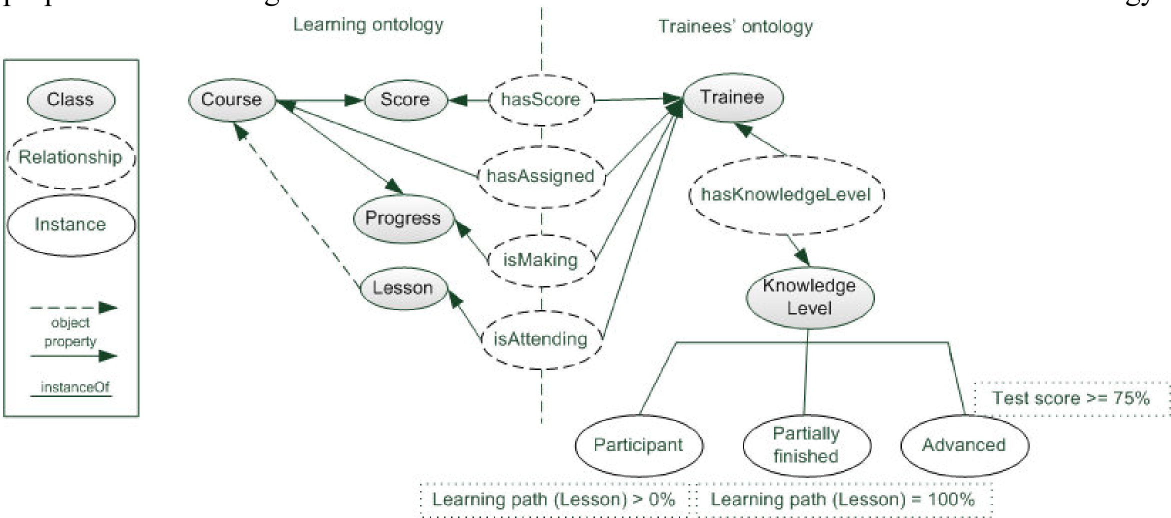

Figure 1: Ontology network consisted of learning and learners' ontology

The connecting of the trainees on a peer-to-peer basis is designed to help them find colleagues to contact in case they might need help but are not feeling ready to make contact with the tutor at that particular stage of the learning process. More precisely, we seek to enhance collaboration not only between trainees and tutors, but also between the trainees themselves in order that they help each other. However, one may ask where is the benefit for the more advanced trainees if they are contacted and asked for help by those who learn more slowly? Actually, the findings stated in (Hwang et al., 2012) and (Gillies, 2004) do confirm that peers like to help each other within a collaborative learning context. The benefit for those who were helping their colleagues lies in their ability to recapitulate and improve their knowledge too.

The ontology-based rules that we applied basically answer the following queries:

- Find all trainees in the tutored group of a particular trainee who scored more than 75% on the course test
- Find all trainees in the tutored group of a particular trainee who have read all of the course material (i.e. have reached 100% of the learning path)
- Find all trainees in the tutored group of a particular trainee who have read more course material than the trainee (i.e. find the trainees in the tutored group who have reached a higher percentage in the learning path).

Again, we developed a special algorithm, which starts by finding the tutored group with the lowest density. Within that group, the trainee with the lowest outgoing node degree has to be found. The above rules are applied to that particular trainee. The trainees found by applying these rules to the trainee dataset of the tutored group are then proposed to that trainee as *recommended contacts* for help.

To sum up, our framework consists of (1) applying social network engineering to create tutored groups that prevent the isolation of the trainees and (2) applying SNA methods and ontology rules in order to provide peer-to-peer infrastructure for collaboration and help between trainees.

Implementation

At first, we found and used an E-Learning system fulfilling the specific criteria described in (Maglajlic 2010). Basically, we were looking for a system that enables hierarchical structuring of the courseware, tutoring (grouping into virtual classroom, discussion forum, chat), learning path tracking (also referenced as learning progress, or learning trajectory in (Pernas et al. 2012)), assessment of each learning object, common calendar with event announcements, full text search, linking of internal and external content, and multilingualism (internationalization). After a comparative analysis of contemporary, mainly open-source E-Learning systems, we chose the system (Dokeos 2012). Technically speaking, the system is deployed according to the typical LAMP (Linux, Apache, MySQL and PHP, (LAMP 2012)) approach. It is constantly populated with E-Learning material (content) and new users are registering in increasing numbers. An overview is shown in Table 5.

Table 5

The basic figures of our E-Learning system

E-Learning system figures
6 Courses
44 Lessons (learning paths)
1051 Registered users (concurrently active up to 20)
15 Trainers

However, in order to utilize the system for our purpose, the trainees' and tutors' profile records in the systems' database have been extended so as to contain not only basic information such as name, login, password and email, but also affiliations such as company, country, spoken languages and business branch. The trainees can be assigned to several courses, but for each course, the trainee can belong to only one tutored group per course (i.e. he/she has only one tutor per course). The system records the trainees' learning paths, their test scores as well as their contributions to the discussion forum of their tutored groups. Both the learning path and test score are available for each user in the form of reports to the E-Learning administrator. At the beginning, before we applied the changes according to our framework, the functional architecture of the system looked as shown in Figure 2.

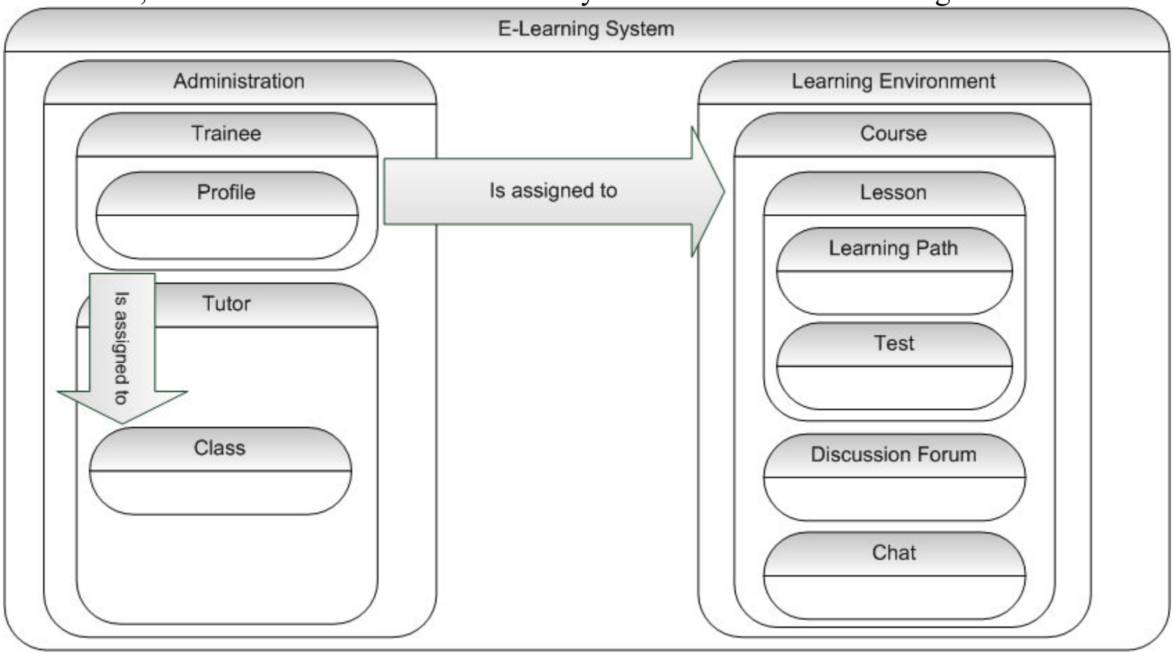

Figure 2: Functional architecture before the modifications

Hence, in the second phase of the implementation, in order to support our approach – namely the application of social network engineering and ontology rules – the functional architecture had to be extended. We already mentioned that the trainee / tutor profiles had been extended with affiliation data. In accordance with our proposed framework, a first major extension of the existing E-Learning system was carried out: the algorithm for the creation of the tutored groups was implemented as an additional system module. This module is invoked when a new user is added to a course. The following diagram (Figure 3) shows the simplified flow-chart of the module.

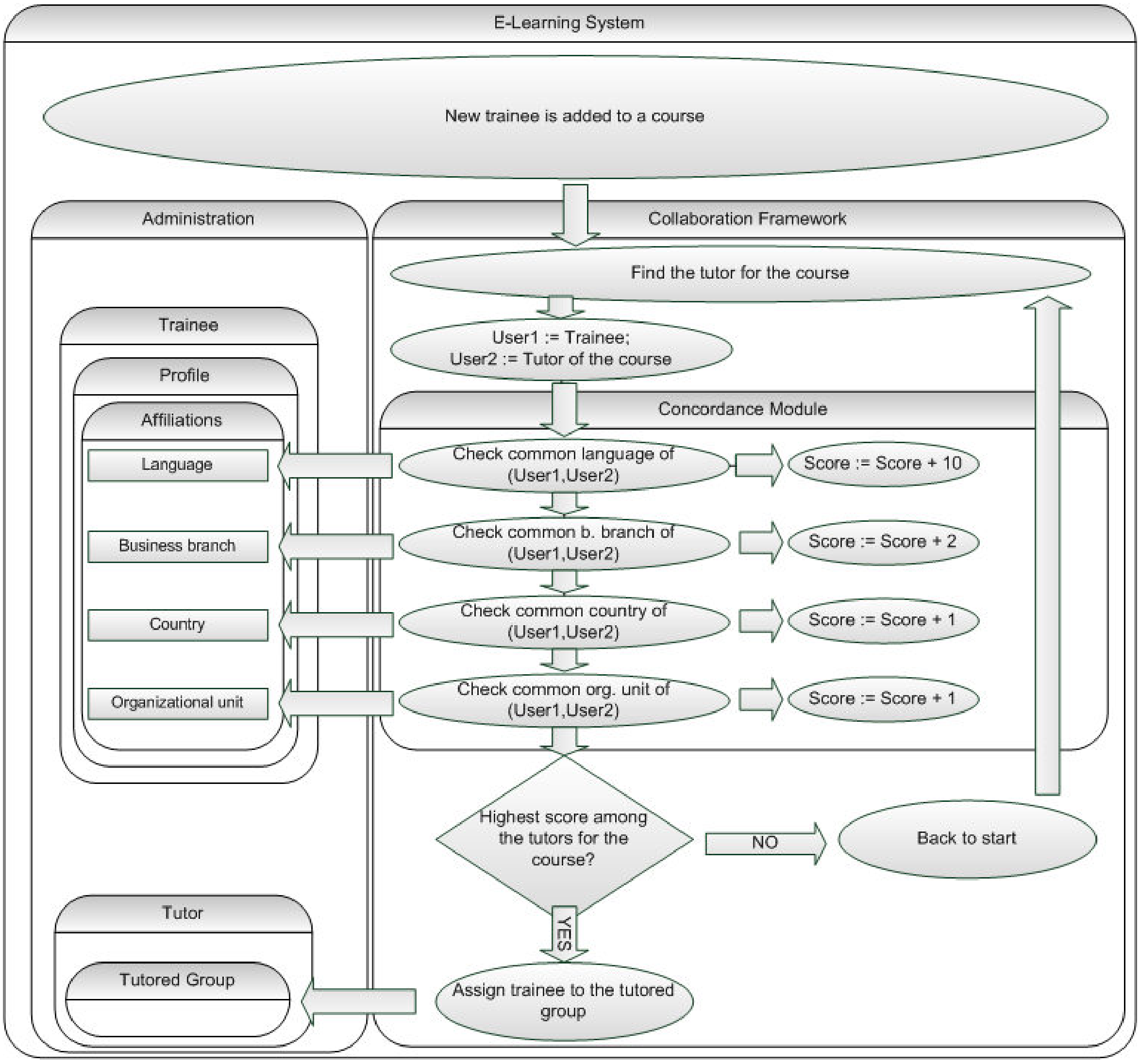

Figure 3: Invocation of the concordance algorithm

The primary aim of the invocation of this module is to assign the trainee to a tutored group, where communication between trainee and tutor takes place within the dedicated discussion forum. However, we note that the invocation of the parts of this module (sub-module for concordance calculation) is not limited to the creation of the tutored groups. It can also be invoked with a changed input parameter set: it need not be only the trainee and tutor who need to be connected, it can also happen that a trainee needs to be linked with another trainee. Such a case is described in the section below.

The third phase of the implementation consisted of the second major extension of the existing E-Learning system: an algorithm for the construction of peer-to-peer links (for trainee-to-trainee collaboration). This algorithm was somewhat more complex than the one used for tutored group construction. In the previous section (framework description) we briefly mentioned the algorithm steps. Now, let us go into more detail. The algorithm consists of the following steps:

- Calculation of the density of each tutored group in the system: this is done by summing up the number of entries in the discussion forum of the group, multiplying the number of trainees in the group by 20 and dividing these two values.

- Calculation of the node degree for each trainee in the group: this is done by counting the number of entries in the discussion forum.
- Calculation of the concordance between the trainee and his/her fellow trainees from the group: this is done through the invocation of the sub-module for the concordance calculation, but with the two trainees as input parameters instead of trainee and tutor, as was done for the module dealing with the construction of the tutored groups.
- Application of the ontology rules to search for the best matching colleague, who will be proposed for peer-to-peer help. The ontology rules are written in the SWRL (Semantic Web Rule Language, (Horrocks et al. 2004)), as shown in Figure 4, for the precise software design purposes. In the software implementation, the rules are interpreted as modules, and the relations as functions – with input and output parameters according to the ontology rule definition. The classes of the ontology are also represented accordingly as data records in the system's database.

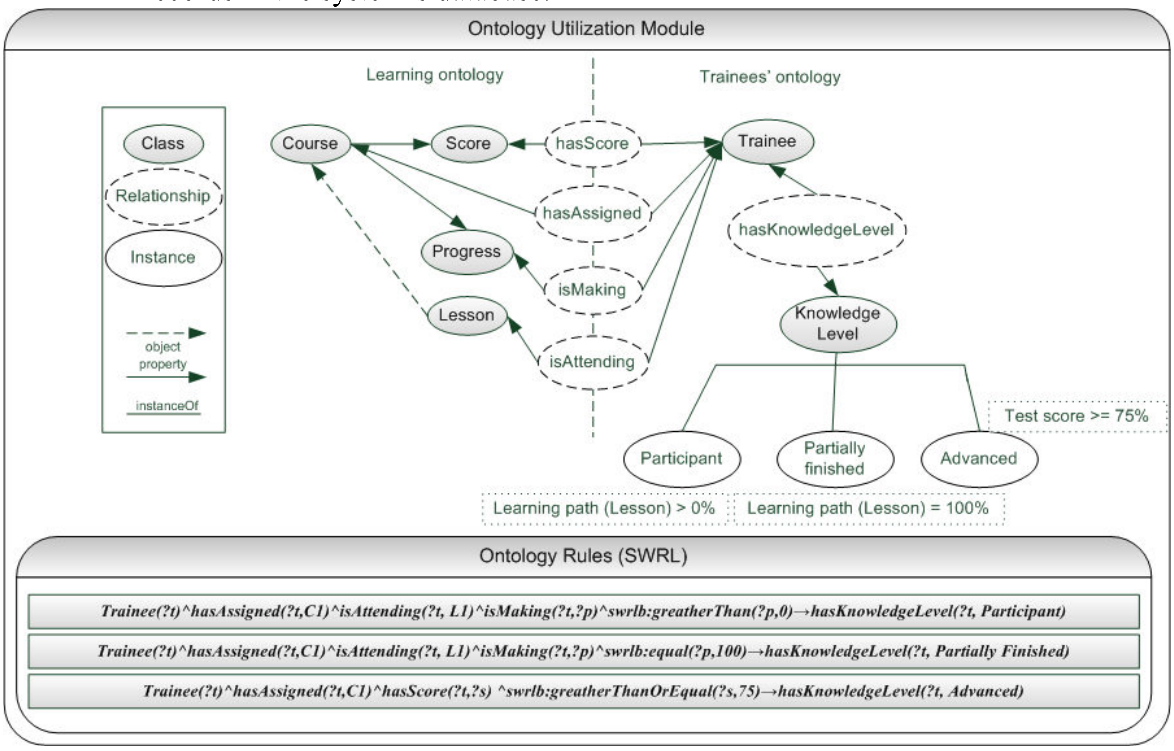

Figure 4: Ontology utilization module with rules written in SWRL

In the same way as for the matching algorithm regarding the tutored groups, this ontology utilization algorithm is implemented in a separate module. It is also worth mentioning that the sub-modules that correspond to the ontology rules can be independently invoked on demand. Execution of this module results in a list of names and e-mail addresses which is shown to every trainee as a "recommended contact". The final functional architecture of the extended system can be summarized as shown in Figure 5.

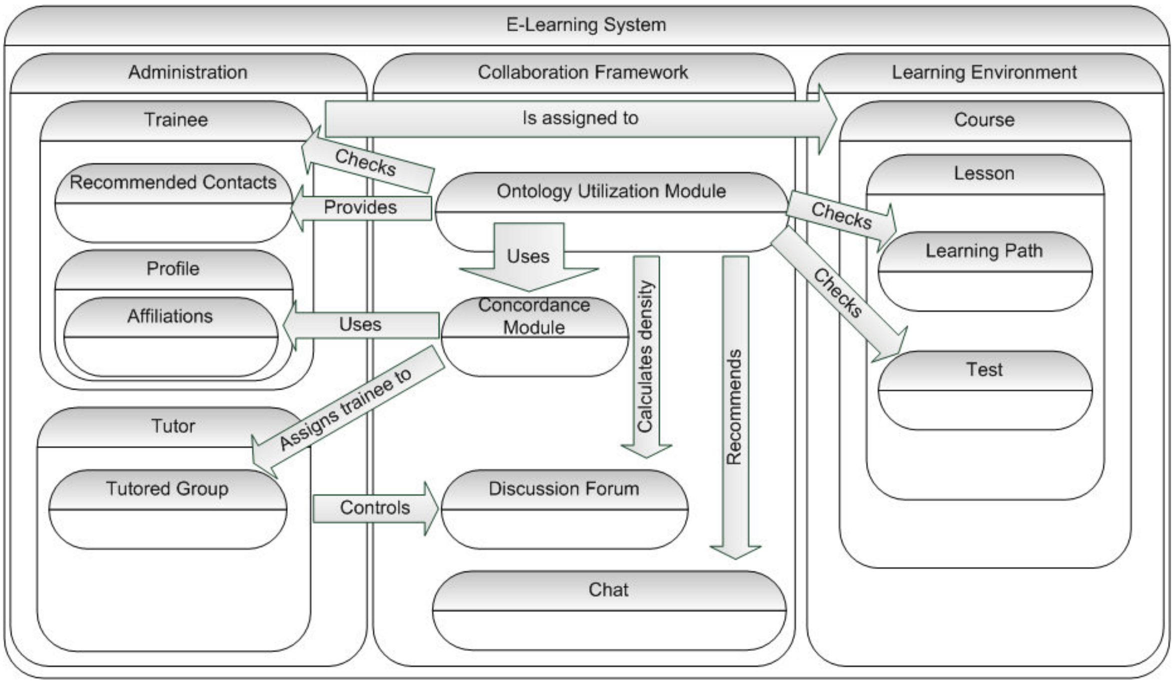

Figure 5: The functional architecture after modifications

Evaluation

In order to evaluate the new methodology when applied to E-Learning, we utilized the sample *t* test. The aim of the evaluation was to see whether the trainees had increased their knowledge more than before the E-Learning system was modified (as described above). Basically, we compared two approaches: using the E-Learning system as it was originally implemented and using the E-Learning system *after* the implementation of the framework for enhanced collaboration. We took two independent samples. The first sample consisted of 18 trainees and their test scores, obtained from January 2010 until August 2011 – in fact, the same sample that had been used for the PPMCC calculation shown in Table 1, before modifications to the E-Learning system were made. The second sample consisted of 19 trainees and their test scores in the period from January 2012 to October 2012, i.e. *after* the system had been upgraded thanks to the new framework for collaboration enhancement. All the trainees in the second sample participated in the corresponding tutored groups. The results of the *t* test are shown in Table 6.

Table 6

t test results of learning outcome: $t=-2.2188$, $p= 0.01645 <.05$; 95% confidence interval <-∞, -1.460363)

Sample taken	N	mean	*t*
Before system modifications	18	85.67	-2.2188
After system modifications	19	91.78	

The *t* test revealed a remarkable difference between the two approaches ($t=-2.2188$, $p<.05$), indicating that trainees are able to achieve a greater learning improvement within a collaborative framework that takes into account their social network and ontology parameters than those without such an infrastructure in their E-Learning settings with the significance level of 5%.

Discussion

This approach is original, compared to other research attempts in this field, since it employs automatic social network engineering, implements SNA quantification methods directly, uses the results of the SNA quantification and applies ontology rules dynamically. The E-Learning systems that are currently popular (WRD 2012) mostly do provide support for social networking (Dokeos 2012, Moodle 2012), but there is no indication that these systems allow our particular combination of methods: automatically-controlled social network engineering to create the tutored groups, and ontology methods and rules for the peer-to-peer connecting of trainees in a flexible way so that these methods can be directly used for our specific purpose. Actually, in spite of the fact that ontology is strongly related to E-Learning, there are not many systems that integrate and make use of ontology rules. One of the rare examples is the system (ILIAS 2012), which employs ontology rules for full-text search. In that particular system, the learning objects are described with RDF (Resource Description Framework, (Manola and Miller, 2004)) and a set of ontology rules is implemented for the search. In the same way as in our framework implementation, in the (ILIAS 2012) system, the ontology relationships are implemented as functions (methods) in the software modules, and the rules as well. The ontology classes influenced the software design – they are interpreted through the object model of the software as well as data records in the system's database.

Although there are already some E-Learning systems that support direct integration of social network functionalities and ontology rules, these do not provide built-in social network engineering: automatic social network construction (or link proposals for the E-Learning participants based upon sophisticated ontology rules), collaborative clusters for trainee-tutor communication, and enhancement of peer-to-peer communication.

However, we note that there is still some room for improvement with the implementation of our framework. Currently for the application of SNA methods we can only take into consideration collaborations carried out within the discussion forum. It would be useful to automatically take into account other collaboration attempts such as chat or e-mailing or, if trainees and tutors agreed to this, to note the phone calls (just counting them and entering the count in the system). If we go one step further in the future, collaboration through some other social networks might be considered, such as messaging through business or private networking facilities such as LinkedIn or Facebook (LinkedIn 2012, Facebook 2012).

Furthermore, from the technical point of view, the following improvements are worth considering. (1) An administrative table of concordance categories where new categories and weights can be added and/or updated. With this approach, the behavior of the concordance algorithm could be influenced, as well as the resulting structure of the tutored groups. This would work since the algorithm implementation in the module concerned is flexible enough to treat all given categories dynamically. (2) We could not locate a fully functional (stable) interpreter of SWRL that could be directly deployed within our system environment; the only resource which is publicly available was found in – (see (Ball et al. 2005)) – but seems not to have been developed further.

We compared our implementation to that shown in (ILIAS 2012), and detected the following similarity In the case of (ILIAS 2012) system and our implementation, the ontology rules are not dynamically interpreted by the system, i.e. no SWRL interpreter is used to execute the ontology rules generically. Such an interpreter could be applied / invoked in the module for ontology rules interpretation. With this approach, one could execute any given rule written in SWRL that has been defined in the ontology network for a particular system's dataset. Currently we have to program/add a new sub-module each time the new rule is explored manually, but with such an interpreter, the rules would be automatically executable on demand.

Conclusions

It is important to mention that this framework is extensible – after implementation of the flexible algorithms for social network engineering that use weighting of the social network affiliations for the basic construction of tutored groups, and combine ontology rules in order to enhance peer-to-peer communication – it is possible to (1) add new weighting categories, (2) vary the values used in the weighting, (3) add new ontology rules, (4) vary ontology rules, depending on which results one wishes to achieve, i.e. how should the collaboration within the social network in the E-Learning setup be enhanced.

Ontology engineering offers a lot of resources worth considering. For example, we could model a situation-aware ontology network, in a similar way as in (Pernas et al. 2012), i.e. by adding both the technology ontology to the existing ontologies for learning material and for learners. The technology ontology would contain classes that help distinguish between devices that trainees use for E-Learning (such as smartphones or laptops) and relations that indicate which device is used by a trainee. Thus we could think of linking the smartphone user to someone who uses the standard computer screen (e.g. 17" laptop). It might be the case that some images or flash animations in the E-Learning course material require a high-quality display, and that a specific lesson containing multimedia material can be explained better by a person who has browsed the material on a screen with higher resolution (quality).

As has already been mentioned in the discussion section, the implemented framework is currently suited to our industrial setting, but could be further improved to make it more universal by, for example, introducing a SWRL interpreter that would execute arbitrary ontology rules dynamically.

However, to conclude, we may state that the framework basically fulfills the objectives of our research: it provides the means for the creation of collaborative social networks within the E-Learning environment and it helps, as the independent sample *t* tests show (see fourth section), to improve the learning outcome of the trainees. This framework is not only suitable for our specific industrial setting;, it could also be extended and deployed in any other E-Learning environment.

References

Åström, K.J. and Murray, R.M. (2008), *Feedback Systems: An Introduction for Scientists and Engineers*, Princeton University Press.

Ball, M., Boley, H., Hirtle, D., Mei, J., Spencer, B. (2005), Implementing RuleML Using Schemas, Translators, and Bidirectional Interpreters, W3C, retrieved in October 2012 from http://ruleml.org/w3c-ws-rules/implementing-ruleml-w3c-ws.html.

Bondy, J.A. and Murty, U.S.R. (2008). *Graph Theory*, Springer, New York.

Dokeos eLearning System (2012), retrieved in December 2012 from http://www.dokeos.com/.

Facebook (2012), retrieved in December 2012 from http://www.facebook.com/.

Gillies, R.M. (2004), The effects of collaborative learning on junior high school students during small group learning, Learning and Instruction, Vol. 14, No. 2, 2004, pp. 197-213.

Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., Dean, M. (2004), A Semantic Web Rule Language Combining OWL and RuleML, retrieved October 2012 from http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/.

Hwang, G.J., Chu, H.C., Liang, Y.R. (2012), Effects of Computerized Collaborative Concept Map Approach on Students' Learning Achievements and Cognitive Loads, 12th IEEE International Conference on Advanced Learning Technologies, ICALT 2012, Rome, Italy, July 4-6, 2012, DOI 10.1109/ICALT.2012.233.

ILIAS (2012), retrieved in December 2012 from http://www.ilias.de/.

Johnson, D.W. & Johnson, R.T. (1994), *Learning together and alone: collaborative, competitive, and individualistic learning*, (4th ed.), Boston: Allyn and Bacon press.

Jung, J.J. and Euzenat, J. (2007), Towards Semantic Social Networks. In Proceedings of the 4th European conference on The Semantic Web: Research and Applications (ESWC '07), Enrico Franconi, Michael Kifer, and Wolfgang May (Eds.), Springer-Verlag, Berlin, Heidelberg, 267-280. DOI=10.1007/978-3-540-72667-8_20, retrieved in October 2012 from http://dx.doi.org/10.1007/978-3-540-72667-8_20.

LAMP (Software bundle) (2012), retrieved in October 2012 from http://en.wikipedia.org/wiki/LAMP_(software_bundle).

LinkedIn (2012), retrieved in December 2012 from http://www.linkedin.com/.

Liu, Y., Slotine, J., Barabasi, A. (2011), Controllability of complex networks, Nature Journal, Nature Publishing Group, a division of Macmillan Publishers Limited, Vol. 473, pp. 167-173.

Maglajlic, S., Helic, D. (2010), Integrating E-learning into work processes in industrial settings: a case study, Proceedings of the 9th international conference on Information technology based higher education and training (ITHET), Cappadocia, Turkey.

Maglajlic, S., Helic, D., Trattner, C. (2010), Social Networks and eLearning: New Model for Learning at Workplace, Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces, Cavtat / Dubrovnik, Croatia.

Maglajlic, S., Helic, D. (2011), How do social networks influence learning outcomes? A case study in an industrial setting, IADIS International Conference WWW/Internet 2011, Proceedings, pp. 203-213.

Maglajlic, S. (2012), Engineering Social Networks Using the Controllability Approach Applied to E-Learning, Proceedings of 12th IEEE International Conference on Advanced Learning Technologies, DOI 10.1109/ICALT.2012.209.

Maglajlic, S., Gütl, C. (2012), Efficiency in E-Learning: Can Learning Outcomes Be Improved by Using Social Networks of Trainees and Tutors?, Proceedings of International Conference on Interactive Collaborative Learning (IEEE ICL), Villach, Austria, 2012; printed version: paper number 260.

Manola, F., Miller, E. (2004), RDF Primer, W3C, retrieved in October 2012 from http://www.w3.org/TR/rdf-primer/.

Mika, P. (2005), Ontologies are us: A unified model of social networks and semantics, In Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A., eds.: Proc. of the 4th Int. Semantic Web Conf. Volume 3729 of Lecture Notes in Computer Science., Springer pp. 522–536.

Moodle E-Learning System (2012), retrieved in October 2012 from http://moodle.org/.

Pernas, A.M., Diaz, A., Motz, R., de Oliveira, J. P. M. (2012), Enriching adaptation in e-learning systems through a situation-aware ontology network, Interactive Technology and Smart Education, Vol. 9, Iss: 2, pp. 60-73.

Scardamalia, M. (2002), Collective cognitive responsibility for the advancement of knowledge, Liberal Education in Knowledge Society, Chapter 4, pp. 67-98.

Silva A., Figueira, A. (2012), Visual Analysis of Online Interactions through Social Network Patterns, 12th IEEE International Conference on Advanced Learning Technologies, 2012, DOI 10.1109/ICALT.2012.57.

Slavin, R.E. (1995), *Collaborative learning: theory, research, and practice,* (2nd ed.), Boston, MA: Allyn and Bacon press.

Sontag, E.D. (1998), *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Second Edition, Springer, New York.

Trausan-Matu, S., Dascalu, M., Rebedea, T. (2012),"Computer-Supported Collaborative Learing, dialogism, chat, polyphony, Natural Language Processing", 12th IEEE International Conference on Advanced Learning Technologies, 2012, DOI 10.1109/ICALT.2012.101.

Wasserman, S. and Faust, K. (2009), *Social Network Analysis: Methods and Applications*, Cambridge University Press, New York, USA.

Web Resources Depot (WRD) (2012), retrieved in October 2012 from http://www.webresourcesdepot.com/7-widely-used-and-open-source-e-learning-applications/.

4. Final Evaluation

This chapter provides an overview of the answers to the research questions that were presented in the introductory chapter, as well as a final statistical analysis of the research findings. It describes the experimental results and analyses the dependencies between the learning outcome and the methods applied within our scientific framework in relation with the research questions. Firstly, we will briefly mention the dependencies between social network parameters (SNA quantifiers) such as communication intensity (i.e. collaboration) in the E-learning environment that were indicated by applying PPMCC calculation and analyzing the values in the "Papers" chapter (Chapter 3). Secondly, we will show the two sample t tests made on the observed samples during the work at different points in time. Thirdly, we will provide an analysis of the learning outcome progress with chi-square tests by comparing the results of the experiments before, during and after the full implementation of the new framework for enhancing collaboration in our E-Learning system.

RQ1: What are the structure and properties of implicit social networks in E-Learning?

The existence of social networks in the E-Learning setting has been investigated and their correlation to the learning outcome given by PPMCC calculation was presented in Paper 1 in Chapter 3. Basically, we have identified two types of implicit social networks within the E-Learning setting: the affiliation network and the communication network. The first SNA quantifier that was used in Paper 1 was the node degree. The results in Paper 1 showed that, in our specific case, the language node degree in the affiliation networks and the ingoing and outgoing node degree in the communication networks of the trainees have the highest correlation values to the learning outcome among all node degrees measured on both networks.

RQ2: How can we control and adapt social networks in the E-Learning setting to enhance collaboration?

The experiences gained during the investigation of RQ1 provided us with a significant input for the concept of the utilization of implicit social networks that is presented in Paper 2 (Chapter 3). Briefly speaking, the relations within the affiliation network of the trainees and tutors have been weighted. Their weighting categories have been used for the calculation of concordance between trainees and tutors resulting by the forming of special tutored groups. In this work we also use a synonym for this methodology: "Social Network Engineering". The creation of tutored groups aims to enhance collaboration between tutors and trainees through a dedicated communication channel (discussion forum). In Papers 2 and 3 we noted that the implementation of the tutored group concept within the E-Learning setting did help to improve the learning outcome, however, the collaboration intensity between trainees and tutors did not increase, i.e. it remained at the same level. More precisely, as shown in Table 1, the sample t test results indicate that the learning outcome did improve after the implementation of the new tutoring infrastructure within the E-Learning system, but, as shown in Table 2, the communication intensity remained at more or less the same level as before the implementation.

Table1: The second *t* test: ***t*= -1.8825**, 95% confidence interval <-∞, -0.6462717), **p= 0.03529 (<0.05)**

Sample taken	**N**	**mean**	***t***
First experiment (Nov 2010)	18	85.667	*t* = -1.8825
Second experiment (Jan 2012 – August 2012)	11	92.454	

Table2: Average outbound communication node degree of the trainees who participated in the experiments

Sample taken	**N**	**Average outbound node degree**
First experiment (Nov 2010)	18	1.8
Second experiment (Jan 2012 – August 2012)	11	1.9

We note here that the outcomes which were used in the sample t test for the period from January 2012 to August 2012 were taken only from trainees who participated in the tutored groups. This case is described in detail in Paper 3 –trainees who did not participate in the tutored groups achieved significantly weaker results.

Actually, thanks to this approach, we found a new method enabling the early detection of the potentially isolated nodes among the trainees. We called this method **static** in Paper 3. We indicated that a **dynamic** method for the detection of trainees who could be potentially isolated during the E-Learning course was still lacking. More precisely, the mechanism in the E-Learning system that would help to recognize those trainees who communicate less than others or, to put it in another way, who use the communication infrastructure less intensively than they should in order to improve their learning outcome, was still not in place. In Paper 3 we applied new SNA methods in addition to the node degree: cliques, centrality and density. Identifying cliques in social network graphs is usually not an easy task; however, in our case it was rather straightforward since our tutored groups had already been constructed to serve as cliques. Following the PPMCC calculation, we realized that the centrality of the nodes within the social network of E-Learning participants was weakly correlated with the learning outcome; however, the research results in Paper 3 showed that the density of the tutored groups (cliques) was medium-level correlated with the learning outcome.

Therefore, in Paper 4, we proposed the implementation of a framework for collaboration enhancement based not only on the concordance between the E-Learning participants but also on their learning and trainee ontologies (i.e. ontology engineering) and on the density of their tutored groups. More precisely, trainees belonging to tutored groups with a lower density who display a weak learning progress at the moment of the automatic system check are registered by the system according to the implemented framework. The system finds the trainees who have the highest concordance with the registered "weak" trainee **dynamically** during the course attendance, and recommends the trainees to be contacted (i.e. to obtain explanations and tips). The contacting is recommended by providing the link to the chat platform of the learning system as well as the e-mail address of the relevant trainee. The system cannot currently provide any specific information about a trainee contacting another because it can only register the chat; the sending of an e-mail by one trainee to another cannot currently be registered.

We also noted in Paper 4 that the precise evaluation of the communication intensity between trainees after the contacts have been recommended by the system was difficult and that currently the only way to analyze the results in the productive system was to observe the usage of the tutored groups' discussion fora. After the latest investigation we noted the data provided in Table 3.

Table 3: We took the information about the count of the visits to the dedicated discussion fora of the tutored groups during the course from the E-Learning system administration data storage

Period	Average number of visits to fora during courses
Jan –Aug. 2012	11.2
Sept – Oct 2012	30.5

Hence, the number of visits to the discussion forum of the tutored groups dramatically increased after the implementation of the collaboration enhancement framework in the system. After a few informal interviews with E-Learning participants, we assume that the "more advanced" trainees have indicated to their colleagues who have contacted them that some answers to the "frequently asked questions" are already provided within the discussion forum of the tutored group.

The consequences for the learning outcome are discussed under RQ3.

RQ3: How adaptations to social networks within E-Learning settings improve collaboration and learning outcomes?

In Paper 4 we showed that the learning outcome did improve after the implementation of modules that are part of the new framework for collaboration enhancement within the E-Learning system. The result of the sample *t* test with the data used in Paper 3 is shown in Table 1. This data was gathered from the system from January 2012 to August 2012, that is, after the implementation of the social network engineering method for the creation of tutored groups, but before the implementation of additional collaboration enhancement facilities based on ontology engineering. This result indicates that the trainees were able to achieve a better learning outcome when they were organized into special tutored groups generated by social network engineering.

In Paper 4, the sample *t* test is provided; it indicates that, in general, after the full implementation of the framework for collaboration enhancement, the learning outcome of the trainees is better, with a significance of 5% (Table 4).

Table 4: t test results of learning outcome before and after E-Learning system enhancements: ***t*=-2.2188**, **p= 0.01645 <.05**; 95% confidence interval <-∞, -1.460363)

Sample taken	**N**	**mean**	***t***
Before system modifications	18	85.67	*t* = -2.2188
After system modifications	19	91.78	

Additionally, we will provide here one additional sample t test and one chi-square test, which will tell us more about the learning outcome improvement both during and after the implementation of the framework for collaboration enhancement.

Table 5 presents the results of the sample *t* test of the data, which were collected until August 2012 during the implementation of the new framework (i.e. before implementation was complete), when trainees and tutors were brought together in tutored groups according to the social network engineering rules described in Paper 2, and the data after August 2012 were gathered after the full implementation of the framework.

Table 5: *t* test results of learning outcome during the implementation of enhancements and after full implementation of system enhancements: ***t*=-1.6294, p= 0.05814 (>.05)**; 95% confidence interval <-∞, 0.3368899)

Sample taken	N	mean	*t*
August 2012 (during the implementation)	14	86.14	*t* = -1.6294
September 2012 (after the implementation)	12	92.88	

This t test does not indicate difference within the significance level of 5% (p-value is greater than .05) between the learning outcomes of trainees tested during the framework's implementation and after full implementation.

Hence, after analyzing the *t* tests results, we can conclude that trainees have been able to achieve a better learning outcome since the new methods for collaboration enhancement have been implemented in the E-Learning system. However, the difference in the degree of efficiency between the different methods for the enhancement of E-Learning collaboration is still not indicated significantly with this statistical test. More precisely, the result of the t test shown in Table 5 indicates a difference between the tutored group method and the fully-implemented framework (including social network engineering of tutored groups and ontology engineering for contact recommendation for the trainees) in the mean, but not as significant as the learning outcome results collected after the implementation of both these methods separately differ from the learning outcome results gathered within the initial E-Learning setup (Table 1 and Table 4).

In order to investigate the impact of the new methods on the E-Learning setting more precisely, additionally we applied the chi-square tests. For this purpose we took samples from three different periods and categorized the learning outcomes of the trainees in the following way. In the period from April 2009 to August 2011, we took a sample of 21 trainees who had participated in 80 tests concerning 8 lessons. These trainees had no enhanced collaboration infrastructure. The E-Learning system did not contain any additional module, only the initial setup. In the rest of this chapter we will refer to this period as the initial period. In the period from September 2011 to August 2012, we took a sample of 14 trainees who participated in 62 tests concerning 7 lessons. During that time the new module for the social network engineering of special tutored groups (as described in Papers 2 and 3) was deployed in the E-Learning system. The first additional method to enhance collaboration was applied to these trainees. We emphasize here that 3 trainees did not participate in any of the tutored groups (their learning outcomes were excluded in the *t* test provided in Table 1). In the rest of this chapter, we will refer to this period as the middle period. In the period from September 2012 to December 2012, we took a sample of 12 trainees who participated in 74 tests concerning 11 lessons. This is the period after the full implementation of the collaboration enhancement framework in the E-Learning setting. We will refer to this period as the final period.

Table 6: Sample taking periods, number of participants, lessons and tests

Period	**Participants #**	**Lessons#**	**Tests#**
Initial Period	21	8	80
Middle Period	14	7	62
Final Period	12	11	74

The categorization of the learning outcome was done according to the scale given in Table 6.

Table 7: Categorization of learning outcomes

Category	**Score**
1	< 75
2	75 – 90
3	> 90

As the table shows, the learning outcome under 75 scores belongs to Category 1. The category of the outcomes between 75 and 90 scores is 2, and those scores between 90 and 100 belong to Category 3, which is the highest. The chi-square tests provided us with the following findings.

Firstly we compared the learning outcomes during the initial period with those during the middle period. The result of the corresponding chi-square test is given in Table 8.

Table 8: Chi-square (**χ2**) test result: **p = 0.7821**

Sample taken	**Category 1**	**Category 2**	**Category 3**	**χ2**
Initial Period	18	20	41	0.4916
Middle Period	11	17	33	

These results show no significant difference between the samples taken in the initial and middle periods, since p = 0.7821. The *t* test results shown in Table 1 do provide different information, however, only after the exclusion of the 3 trainees who did not participate in the tutored group activities.

Secondly, if we compare the outcomes during the initial period with those during the final period, the chi-square test shows some significant differences (Table 9) even more clearly than the *t* test shown in Table 4.

Table 9: Chi-square (**χ2**) test result: **p = 0.01525 (<.05)**

Sample taken	Category 1	Category 2	Category 3	χ2
Initial Period	18	20	41	8.3669
Final Period	6	14	53	

Hence, all the trainees in the final period sample were integrated into tutored groups, and their learning outcomes were likely to be significantly improved, since the p-value is 0.01525 – which is less than 0.05 (i.e. a significance level of 5% is clearly reached). In other words, the efficiency of the learning of the trainees when they were integrated in the collaboration enhancement framework significantly improved.

The third and the last comparison we made was between the outcomes during the middle and final periods. It can be seen (Table 10) that the significance level has not been reached (p-value is 0.06786) but the p-value is getting close to it.

Table 10: Chi-square (**χ2**) test result: **p = 0.06786**

Sample taken	Category 1	Category 2	Category 3	χ2
Middle Period	11	17	33	5.3806
Final Period	6	14	53	

Actually, the p-value of the chi-square test shown in Table 10 points to a difference concerning the efficiency of the trainees in improving their learning outcomes during the middle and final periods, in almost the same way as the p-value shown in Table 5. Hence, we can conclude that the learning outcomes during the middle and final periods do differ, but not dramatically (the significance level of the t test shown in Table 5 is 6% and the significance level of the chi-square test shown in Table 10 is 7%) as was the case when we compared the outcomes in the initial and final periods.

Thanks to the results of the analysis presented above, the RQ3 has also been answered: the full implementation of the framework as described in Paper 4 help to increase the efficiency of the trainees (i.e. their learning outcomes) who attend the E-Learning courses.

5. Conclusions

The first conclusion that can be drawn from this piece of research is that a framework for the solution of the problem stated in the introduction chapter has been developed successfully. We have established mechanisms to detect potentially isolated E-Learning trainees, as well as to enhance their inclusion (maybe, we can say "de-isolation") in order to help improve their learning outcomes. We invented two methods for the de-isolation of the trainees: static and dynamic.

Static method is based on the implicit social network position of the trainee. It is implemented as a novel matching (grouping) algorithm for modeling of special social networks for dedicated tutoring. We call this method static because it is applied before the trainee starts to attend the E-Learning course.

Dynamic method relays on the combination of the SNA quantification and ontology engineering applied on the social networks produced with the above mentioned static method. We call this method dynamic because it is applied during the trainees' attending the E-Learning course. It results with the permanent recommendation of the contacts to the trainees for the collaboration and E-Learning outcome improvement.

Our second conclusion is that the research resulted in the creation of a flexible framework for collaboration enhancement within the E-Learning system which combines both static and dynamic method. As the main result after the statistical analysis provided in the previous chapter we may emphasize the following. Implementation of the method for static de-isolation of the trainees alone is not sufficient for significant E-Learning outcome improvement. Only after the implementation of the dynamic de-isolation method by combining SNA and ontology engineering, the E-Learning outcomes started to differ significantly from those before the new methods have been applied.

5.1. Contributions

The following characteristics of the study make it original and specify its scientific contribution:

- We investigated the impact of implicit social networks within E-Learning systems on the learning outcomes
- We introduced a social network engineering method for the grouping of trainees around a tutor according to the concordance of their social network characteristics
- Introduction of an ontology engineering method designed to enhance the collaboration between trainees
- We specified and implemented a framework for the utilization of social network engineering and ontology engineering methods whose aim is to enhance collaboration

within an E-Learning system between trainees and tutors for the purpose of improving the trainees' learning outcomes.

Furthermore, it is rather important to mention that in the technical implementation of the extensions of the E-Learning system it provides a novel approach of combination of usage of SNA methods such as density calculation and implementation of ontology rules.

This work also provides an example of the utilization of the control theory findings and their application to the social network by inventing a matching algorithm which helps construction of the cliques within the social network for dedicated tutorship.

5.2. IMPLICATIONS OF SOCIAL NETWORK ENGINEERING METHOD (THE GROUPING ALGORITHM, THE STATIC METHOD)

The grouping algorithm for the trainees and tutors (which is described in Paper 2 and further utilized according to the rules given in Paper 4) can be reconfigured with other concordance categories. According to the particular situation at hand, the weights of the concordance categories can be varied, and concordance categories can even be added or removed – the basic principle remains the same. This principle relays on assortative mixing (Newman, 2002), i.e. the nodes in the network are related (are put in relation) according to their similarities. In other words, nodes displaying similar characteristics within the social network are put into relation with each other.

5.3. IMPLICATIONS FOR COLLABORATION ENHANCEMENT OF ONTOLOGY ENGINEERING METHOD (THE DYNAMIC METHOD)

The ontology engineering method is, of course, open to further changes, depending on the requirements of a given E-Learning setting and expectations of its owners and /or administrators regarding the E-Learning system and methods.

In this study, non-assortative mixing was applied according to the same source as in the previous case (Newman, 2002), i.e. the nodes with "opposite" characteristics were put into relation within a given social network. Furthermore, although the ontology engineering method was used to create a recommendation mechanism that enhances collaboration (recommending particular people), it can be modified so that it recommends learning material consistent with the context of the learner (Pernas et al. 2012, Duval et al. 2012) – not only the learning context but also the social network context.

5.4. LIMITATIONS

We have to mention several limitations we took into account during this research.

This research has been carried out in a very specific industrial setting. The methods we invented here have not yet been applied to other business sectors with different industrial settings. The implicit social networks and their influence to the learning outcome and collaboration between the participants may differ in some other business sector.

We already mentioned in the Chapter 4 the limitations regarding the result gathering: during the measurements of the implications of the dynamic method we could only efficiently observe the activities of trainees and tutors within the tutored group discussion fora. Although the chat attempts between the trainees are registered by the system, the direct communication between the trainees via email (email is indicated in the contact data of the "recommended" trainee), telephone, Skype or any other communication channel except discussion forum could not be registered. Under "other communication channel" we may assume currently popular social network facilities such as Facebook or LinkedIn. The main reason not analyze the communication intensity between trainees and / or tutors within these social networks is the very restrictive web security policy in our industrial setting (railways) which very often forbids the users to access these tools.

5.5. RECOMMENDATIONS FOR FURTHER RESEARCH

Both of these methods can be further extended and/or their parameters can be changed. In both cases, the mixing direction can be changed from assortative to non-assortative, and vice versa. For example, in our particular case – an industrial setting – this mixing was appropriate. However, in another setting, another mixing philosophy might be more appropriate. The framework we developed allows such variations, which would be worth experimenting further. The most accurate research results (Vaquero and Cebrian, 2013) confirm the claims of this thesis and indicate the need for extension of the research in this field. The authors have investigated the influence of the social network based collaboration in E-Learning to the learning outcome of the students in an academic environment. Their experimental results are based on a larger sample than the one used in our experiments and the PPMCC values indicate even stronger positive correlation of the collaboration intensity between the trainees to their learning outcome improvement. The authors depict a similar idea to ours: to "catch" those trainees with low communication intensity as potential candidates for weak learning outcome in order to help them. However, the authors still do not provide any proposal for a theoretical or practical framework which would support their idea. Our achievements can therefore serve as a complementary approach to such research attempts and be even used as a guideline for further investigation. The simplified view to our approach is given in Figure 5.

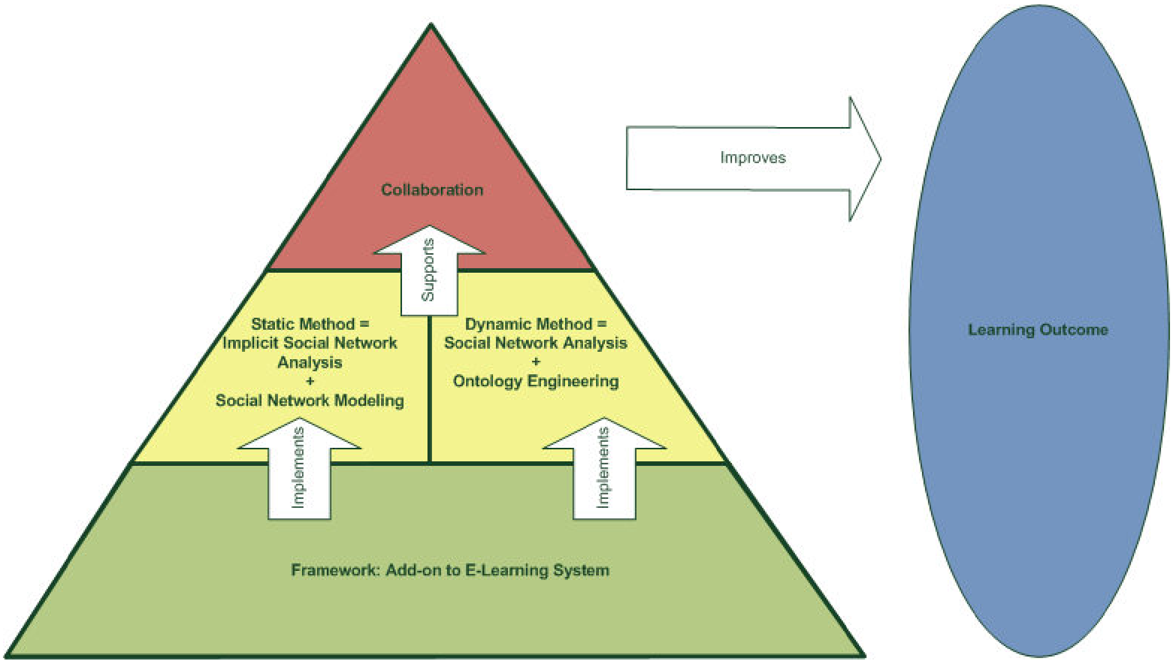

Figure 5: Approach that we propose: combination of SNA and ontology engineering implemented in the framework as an add-on to the E-Learning system supports the collaboration of the E-Learning participants and therefore helps improving the learning outcome.

In any case, we recommend the following research agenda in order to overcome the limitations mentioned in the previous section and support the ideas such as the one mentioned in (Vaquero and Cebrian, 2013):

- Experimenting with changes in concordance categories and ontology relation interpretation and evaluation, as well as experimenting with different mixing directions.

- Registering and evaluation of the direct messaging (email, telephone, Skype and similar) between the trainees, and between the trainee and the tutor, especially after the applying of the dynamic method.

- An evaluation of the influence of potential relations between trainees and tutors (trainee-to-trainee, trainee-to-tutor, tutor-to-tutor) in other (external) social networks, even those that are widespread such as Facebook or LinkedIn, would be interesting. The eventual utilization of such networks and their integration by using ontology rules within the system-based approach sounds challenging.

The formation of new cliques within the social network could also be investigated. In our case, cliques were automatically created through the grouping algorithm. However, it would be interesting to evaluate how the cliques of trainees would be built if we were to apply the assortative and/or non-assortative mixing rule in the ontology engineering method to:

- Relate only trainees with the highest score category

- Relate trainees with a high score across different courses
- Relate trainees who have not participated in the same course at all but have a high score in common with others from the clique
- Relate tutors with the most visited fora
- Relate those tutors with a high visit frequency with those with a low tutored group density
- Relate tutors according to the test score of their trainees – those with high scores with similar ones , and as a second variant, non-assortative.

In our opinion, further research into all the steps listed above would be valuable, not only in our chosen industrial setting, but in any other suitable one as well.

6. References

Åström, K.J. and Murray, R.M. (2008), *Feedback Systems: An Introduction for Scientists and Engineers*, Princeton University Press.

Ball, M., Boley, H., Hirtle, D., Mei, J., Spencer, B. (2005), Implementing RuleML Using Schemas, Translators, and Bidirectional Interpreters, W3C, retrieved in October 2012 from http://ruleml.org/w3c-ws-rules/implementing-ruleml-w3c-ws.html.

Bondy, J.A. and Murty, U.S.R. (2008). *Graph Theory*, Springer, New York.

Breiger, L.R., (1974) *The Duality of Persons and Groups*, Social Forces, 53(2):181–190.

Chatti, M.A., Jarke, M., and Frosch-Wilke, D. (2007), "The future of e-learning: a shift to knowledge networking and social software", *Int. J. Knowledge and Learning*, Vol. 3, Nos. 4/5, pp. 404–420.

Chen, L., Di Eugenio, B., Fossati, D., Ohlsson, S., and Cosejo, D. (2011), "Exploring Effective Dialogue Act Sequences in One-on-One Computer Science Tutoring Dialogues", IUNLPBEA '11, *Proceedings of the 6th Workshop on Innovative Use of NLP for Building Educational Applications*, Association for Computational Linguistics Stroudsburg, PA.

Cho, H., Gay, G., Davidson, B. & Ingraffea, A. Social networks, communication styles, and learning performance in a CSCL community. Computers & Education 49, 309–329 (2007)

Clark, R. C., Nguyen, F., and Sweller, J. (2006), *Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load*. San Francisco: Pfeiffer.

Cohen, J. (1988) *Statistical power analysis for the behavioral sciences*, L. Erlbaum Associates, Hillsdale, N. J., USA

Conole, G., Culver, J.,(2009), Cloudworks: Social networking for learning design. *Australasian Journal of Educational Technology*, 25(5), pp. 763–782

Dalsgaard, C. (2006) "Social software: E-Learning beyond learning management systems", European Journal of Open, Distance and E-Learning.

Dokeos eLearning System (2012), retrieved in December 2012 from http://www.dokeos.com/.

d'Aquin (2006), Gangemi, A., Haase, P. (2006) "Definition of ontology networks", *NeOn Book – NeOn Methodology in a Nutshell*, available at: www.neon-project.org/nw/NeOn_Book

Facebook (2012), retrieved in December 2012 from http://www.facebook.com/.

Gillies, R.M. (2004), The effects of collaborative learning on junior high school students during small group learning, Learning and Instruction, Vol. 14, No. 2, 2004, pp. 197-213.

Griol, D., Molina, J.M, de Miguel A.S., Callejas, Z. (2012), A Proposal to Create Learning Environments in Virtual Worlds Integrating Advanced Educative Resources, Journal of Universal Computer Science, vol. 18, no. 18, 2516-2541

Guillaume, J.-L. and Latapy, M. (2006), “Bipartite graphs as models of complex networks”, *Physica A: Statistical and Theoretical Physics*, Vol. 371, Nr. 2, pp 795 – 813.

Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., Dean, M. (2004), A Semantic Web Rule Language Combining OWL and RuleML, retrieved October 2012 from http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/.

Hwang, G.J., Chu, H.C., Liang, Y.R. (2012), Effects of Computerized Collaborative Concept Map Approach on Students' Learning Achievements and Cognitive Loads, 12th IEEE International Conference on Advanced Learning Technologies, ICALT 2012, Rome, Italy, July 4-6, 2012, DOI 10.1109/ICALT.2012.233.

Haythornthwaite, C. (2005), “Social Network Methods and Measures for Examining E-learning”, *Social Networks*, Citeex:10.1.1.135.6993.
Gelman, A., Hill, J. (2007), *Data Analysis Using Regression and Multilevel / Hierarchical Models*, Cambridge University Press

ILIAS (2012), retrieved in December 2012 from http://www.ilias.de/.

Johnson, D.W. & Johnson, R.T. (1994), *Learning together and alone: collaborative, competitive, and individualistic learning*, (4th ed.), Boston: Allyn and Bacon press.

Jung, J.J. and Euzenat, J. (2007), Towards Semantic Social Networks. In Proceedings of the 4th European conference on The Semantic Web: Research and Applications (ESWC '07), Enrico Franconi, Michael Kifer, and Wolfgang May (Eds.), Springer-Verlag, Berlin, Heidelberg, 267-280. DOI=10.1007/978-3-540-72667-8_20, retrieved in October 2012 from http://dx.doi.org/10.1007/978-3-540-72667-8_20.

Kalman, R. E. (1963), Mathematical description of linear dynamical systems. J. Soc. Indus. Appl. Math. Ser. A 1, 152–192.

Kalyuga, S., and Sweller, J. (2005), "Rapid Dynamic Assessment of Expertise to Improve the Efficiency of Adaptive E-learning," *Educational Technology Research and Development* 53(3): 83–93.

Karrer (2008), http://www.workliteracy.com/knowledge-workframework

Knoke, D., Yang, S., & Kuklinski, J. H. (2008). *Social network analysis* (Vol. 2). Los Angeles, CA: Sage Publications.

LAMP (Software bundle) (2012), retrieved in October 2012 from http://en.wikipedia.org/wiki/LAMP_(software_bundle).

LinkedIn (2012), retrieved in December 2012 from http://www.linkedin.com/.

Littlejohn, A., Milligan, C., Margaryan, A. (2011), Collective Learning in the Workplace: Important Knowledge Sharing Behaviours, International Journal of Advanced Corporate Learning. ISSN: 1867-5565 Vol 4, No 4 (2011)

Liu, Y., Slotine, J., Barabasi, A. (2011), Controllability of complex networks, Nature Journal, Nature Publishing Group, a division of Macmillan Publishers Limited, Vol. 473, pp. 167-173.

Lundvall, B-A., Rasmussen, P., Lorenz, E. (2008) Education in the Learning Economy: A European Perspective. Policy Futures in Education, 6(6), 681-700. http://dx.doi.org/10.2304/pfie.2008.6.6.681

Maglajlic, S., Helic, D. (2010), Integrating E-learning into work processes in industrial settings: a case study, Proceedings of the 9th international conference on Information technology based higher education and training (ITHET), Cappadocia, Turkey, pp.151-157, ISBN 978-1-4244-4810-4.

Maglajlic, S., Helic, D., Trattner, C. (2010), Social Networks and eLearning: New Model for Learning at Workplace, Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces, Cavtat / Dubrovnik, Croatia, pp. 373-378.

Maglajlic, S. (2011), On the Importance of the Impact Analysis of Social Network Methods in Elearning, Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. Chesapeake, VA: AACE, 749–752.

Maglajlic, S., Helic, D. (2011), How do social networks influence learning outcomes? A case study in an industrial setting, IADIS International Conference WWW/Internet 2011, Proceedings, pp. 203-213.

Maglajlic, S., Helic, D. (2011), How do social networks influence learning outcomes? A case study in an industrial setting, Emerald Journal on Interactive Technology and Smart Education, Vol. 9 Iss: 2, pp.74–88.

Maglajlic, S. (2012), Engineering Social Networks Using the Controllability Approach Applied to E-Learning, Proceedings of 12th IEEE International Conference on Advanced Learning Technologies, pp. 276-280, DOI 10.1109/ICALT.2012.209.

Maglajlic, S., Gütl, C. (2012), Efficiency in E-Learning: Can Learning Outcomes Be Improved by Using Social Networks of Trainees and Tutors?, Proceedings of International Conference on Interactive Collaborative Learning (IEEE ICL), Villach, Austria, 2012, pp. 1-8., DOI 10.1109/ICL.2012.6402088.

Maglajlic, S. (2012), Efficiency in E-Learning: Can Learning Outcomes Be Improved by Using Social Networks of Trainees and Tutors? Addleton Academic Publishers Economics, Management, and Financial Markets, Volume 7(4), 2012, pp. 121-137, ISSN 1842-3191

Maglajlic, S. (2012), Social Network Engineering and Ontology Engineering For E-Learning: How Do These Work Together? IADIS International Conference WWW/Internet 2012, Proceedings, in printing.

Maglajlic, S. (2013), Implementation of a Framework for Collaborative Social Networks in E-Learning, AACE International Journal on E-Learning (IJEL) Corporate, Government, Healthcare, & Higher Education, in review.

Manola, F., Miller, E. (2004), RDF Primer, W3C, retrieved in October 2012 from http://www.w3.org/TR/rdf-primer/.

Mann, H. B., & Wald, A. (1942). On the choice of the number of class intervals in the application of the chi square test. *The Annals of Mathematical Statistics*, 306-317.

Myers, R. H. (1990), *Classical and modern regression with applications* (Vol. 2). Belmont, CA: Duxbury Press.

Mika, P. (2005), Ontologies are us: A unified model of social networks and semantics, In Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A., eds.: Proc. of the 4th Int. Semantic Web Conf. Volume 3729 of Lecture Notes in Computer Science., Springer pp. 522–536.

Moodle E-Learning System (2012), retrieved in October 2012 from http://moodle.org/.

Newman, M. E. J. (2002), Assortative Mixing in Networks, Phys. Rev. Lett., Vol. 89, Issue 20, pp. 208701, American Physical Society, DOI 10.1103/PhysRevLett.89.208701, http://link.aps.org/doi/10.1103/PhysRevLett.89.208701 (visited in Januar 2013)

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist*, *38*(1), 63-71.

Pernas, A.M., Diaz, A., Motz, R., de Oliveira, J. P. M. (2012), Enriching adaptation in e-learning systems through a situation-aware ontology network, Interactive Technology and Smart Education, Vol. 9, Iss: 2, pp. 60-73.

Plass, J. L., Moreno, R., & Brünken, R. (Eds.). (2010). *Cognitive load theory*. Cambridge University Press.

R-Project (2013), www.r-project.org, last visit in March 2013

Ruiz, J. G., Mintzer, M. J., Leipzig, R. M. (2006), "The Impact of E-Learning in Medical Education," *Academic Medicine* 81(3): 207–212.

Stocker, A., Granitzer, G., and Tochtermann, K., (2009) Can Intra-Organizational Wikis Facilitate Knowledge Transfer and Learning? Am Explorative Case-Study, *Proceedings of eLBa – eLearning Baltics*, Fraunhofer Verlag, Stuttgart.

Stocker, A., Strohmaier, M., and Tochtermann, K., (2008) Studying knowledge transfer with weblogs in small and medium enterprises: An exploratory case study, *Scalable Computing Practice and Experience*, Vol. 9, Nr. 4, pp 243–258.

Schachner, W., Tochtermann, K. (2008) *Corporate Web2.0 – Web2.0 in Unternehmen - Das paßt das zusammen!* Shaker Verlag, Aachen , ISBN 978-3-8322-7447-4, ISSN 1438-8081

Scardamalia, M. (2002), Collective cognitive responsibility for the advancement of knowledge, Liberal Education in Knowledge Society, Chapter 4, pp. 67-98.

Scott, J. (1988), Social network analysis. *Sociology*, *22*(1), 109-127.

Silva A., Figueira, A. (2012), Visual Analysis of Online Interactions through Social Network Patterns, 12th IEEE International Conference on Advanced Learning Technologies, 2012, DOI 10.1109/ICALT.2012.57.

Slavin, R.E. (1995), *Collaborative learning: theory, research, and practice,* (2nd ed.), Boston, MA: Allyn and Bacon press.

Smith, B. (2001), Objects and Their Environments: From Aristotle to Ecological Ontology, The Life and Motion of Socio-Economic Units (GISDATA 8), London: Taylor and Francis, 79-97.

Sellen, A.J., Murphy, R. M., & Shaw, K. (2002) How knowledge workers use the Web. Proceedings of CHI 2002, Minneapolis, MN. New York: ACM Press, pp. 227-234. Available from http://research.microsoft.com/~asellen/publications/knowledgerworkers and the web 02.pdf

Sontag, E.D. (1998), *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Second Edition, Springer, New York.

Steel, R. G. D. and Torrie, J. H., (1960), *Principles and Procedures of Statistics,* McGraw-Hill, New York, pp. 187, 287.

Trausan-Matu, S., Dascalu, M., Rebedea, T. (2012),"Computer-Supported Collaborative Learing, dialogism, chat, polyphony, Natural Language Processing", 12th IEEE International Conference on Advanced Learning Technologies, 2012, DOI 10.1109/ICALT.2012.101.

Ullrich, C., Borau, K., and Stepanyan, K. (2010) Who Students Interact With? A Social Network Analysis Perspective on the Use of Twitter in Language Learning, M. Wolpers et al. (Eds.): EC-TEL 2010, LNCS 6383, pp. 432–437, Springer-Verlag Berlin Heidelberg

Vaquero, L. M., Cebrian, M. (2013), "The rich club phenomenon in the classroom", Scientific Reports 3, Article number: 1174, Macmillan Publishers Limited. All rights reserved, 2013/01/30/online, doi:10.1038/srep01174, http://dx.doi.org/10.1038/srep01174

Weibelzahl, S., De Bra, P., Paramythis, A., Ertmer, P., & Desjardins, F. (2008, June). Adaptive or Collaborative Learning?. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (Vol. 2008, No. 1, pp. 5474-5477).

Wang, C., and Li, L. (2007), "Enable Collaborative Learning: An Improved E-Learning Social Network Exploiting Approach," *Proceedings of the 6th WSEAS International Conference on Applied Computer Science*, Hangzhou, April 15–17.

Wasserman, S. and Faust, K. (2009), *Social Network Analysis: Methods and Applications*, Cambridge University Press, New York, USA.

Web Resources Depot (WRD) (2012), retrieved in October 2012 from http://www.webresourcesdepot.com/7-widely-used-and-open-source-e-learning-applications/.