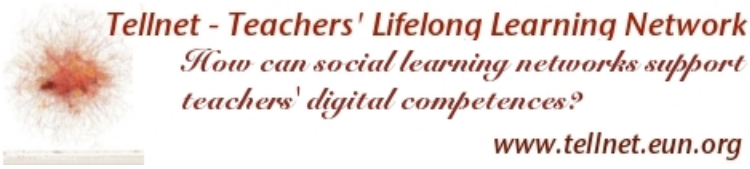


D.2.2
Social Network Analysis Methods for Lifelong Learning Communities

Authors:
Manh Cuong Pham, Yiwei Cao, Ralf Klamma
RWTH Aachen University

Project information

Project acronym:		TeLLNet
Project title:		Teachers' Lifelong Learning Network
Project number:		505594-LLP-1-2009-1-BE-KA4-KA4SRM
Sub-programme or KA:		KA 4
Project website:		http://tellnet.eun.org
Reporting period:	From	01.12.2009
	To	31.11.2012
Report version:		1.0
Date of preparation:		01.04.2012
Beneficiary organisation:		EUN Partnership AISBL (European Schoolnet)
Project coordinator:		Riina Vuorikari
Project coordinator organisation:		EUN Partnership AISBL (European Schoolnet)
Project coordinator telephone number:		+32 2 790 7537
Project coordinator email address:		riina.vuorikari@eun.org

This project has been funded with support from the European Commission.

Executive Summary

eTwinning[1] is the community for schools in Europe, which at the same time, creates a professional development network for European school teachers. It aims to promote European teachers' collaboration through the use of Information and Communication Technologies (ICT). TeLLNet Work Package Deliverable 2.2 aims to identify social roles in the learning network on the member level, on the sub-network structures, and on the network level. A dynamic data flow automatically triggers the re-computation of roles and structures over time, based on the previously established data warehouse. Time series analysis is executed on the eTwinning network data to observe the evolution of the teachers' professional development. We specify this set of social network analysis methods for future research.

Social roles are reflected through social capital which changes over time. In this report, we focus on employment of dynamic social network analysis (SNA) methods to explore eTwinning teachers' social capital via various interactions among them. The analysis is carried out on the established data warehouse in TeLLNet Deliverable 2.1. Accordingly, the data warehouse has been further improved and refined. We advance the research in two parts with the following questions.

First, what is the social capital of a European teacher in eTwinning? Driven by the idea that teachers collaborate across the borders with the support of an electronic platform called the eTwinning portal, we were able to find correlations between social network analysis measures like degree and betweenness centrality as well as the local clustering coefficient, activity statistics about usage of eTwinning, and the quality management through Quality Labels. A Quality Label is a semi-formal recognition granted by the eTwinning National Support Service to eTwinners for their work in collaborative school projects. Only the combination of these measures gives us indicators for teachers' attaining social capital.

Second, how does teachers' social capital change along a long term? A series of network centrality measures are computed at different time points from the starting date in 2005 onwards. We apply time series analysis as a dynamic network analysis method to track teachers' professional development patterns.

The research shows four discoveries based on network analysis methods as the main contributions of this deliverable.

1) The power-law distribution of node degree in eTwinning indicates the existence of complex network and its underlying community structures in the eTwinning professional development network. SNA methods can be well applied on those complex networks.

2) Dynamic analysis, the time-series analysis, of teachers' various communication and project networks reveals teachers' professional development path. Those teachers who are the "structural hole" in the network (connecting different communities) have more social capital than those who are located within one community.

[1] www.etwinning.net

3) The time-series analysis also explores that the development of the eTwinning network follows a network development model based on different phases from born, bonding, to emergence. The eTwinning network still proves to be a young network and has great potential for further development.

4) Social network analysis is able to esteem social capital through calculation of node and network properties. Especially, the structural hole position is in line with those active eTwinning teachers who are involved in more projects with more activities and more contacts. Thus, it is helpful for eTwinners to explore their roles in the eTwinning network and for eTwinning administrative staffs to promote and maintain the eTwinning network.

This deliverable is the outcome of task 2.2 of Work Package 2 with a focus on social network analysis based on an improved data warehouse. It is advanced on the results and feedback of the deliverable 2.1 report (see document D2.1 - Data management of large-scale lifelong learning data). All data analysis research is conducted on the anonymous data set extracted from the eTwinning network to guarantee eTwinning teachers' data privacy.

Table of Contents

1. Introduction 6
2. Social capital and dynamic social network analysis 8
   2.1. Social capital 8
   2.2. Dynamic social network analysis 10
3. Extension of the TeLLNet data model 13
   3.1. Review of the existing data model 13
   3.2. Extension of data warehouse for dynamic social network analysis 14
4. Properties of eTwinning networks 16
   4.1. Static properties 17
   4.2. Dynamic properties 19
   4.3. Summary 21
5. Social capital in eTwinning networks 21
   5.1. Structural hole and closure in eTwinning networks 21
   5.2. Local network structure and teacher's performance 25
   5.3. Success factors for projects 27
   5.4. Summary 29
6. Tracking development pattern of teachers 30
7. Conclusions and future work 32
Reference 33
Annex I: Data model 35

1. Introduction

The eTwinning[2] portal aims at providing European school teachers a common platform for communication and cooperation, as well as for professional development. In eTwinning teachers communicate and collaborate with each other using different information and communication technologies (ICTs) made available by the eTwinning portal. Through the use of ICT, teachers are building up a large-scale and dynamic social network for their professional development. As of November of 2011, 146,105 teachers from 34 countries have registered in eTwinning and involved in 23,641 projects.

The main actors of the portal are eTw*inners*, i.e. teachers from 32 different European countries that currently participate in the initiative. The platform is also used by the National Support Services (*NSS*) that function in each participating country and supports local users in eTwinning. The platform is managed by the Central Support Service (*CSS*), who is responsible for the implementation, and considered as a service provider. The Central Support Service, run by European Schoolnet, is in a position to collect and process data according to the data protection rules defined in the eTwinning privacy statement[3], as well as observe all communications and interactions between eTwinners (data processor). The platform is run under a service contract for the European Commission, the ultimate data controller is the Educational, Audiovisual & culture Executive Agency.

eTwinning offers teachers three main streams of activities:

1. Teachers can find partners to run *cross-border school collaboration projects* using Information and Communication Technologies (ICT) provided by the platform[4].
2. Various *formal and informal professional development* (PD) activities are offered. These include online Learning Events[5], a distance course for teachers, and more informal PD activities such as online interest Groups[6] and Teachers Rooms on topics of interest.
3. Additionally, the participating teachers have a set of *social networking tools* available, these include a profile page with personal and professional information[7], possibility to display connections to friends (i.e. contacts) and possibility to post on personal journal (e.g. status updates), but also post updates and comment on contacts' journals.

In order to improve cooperation among European school teachers, the TeLLNet project aims to explore how European teachers can use social networks to perform better in their carrier and to help them develop their professional competences. In TeLLNet, Social Network Analysis (SNA) is applied to understand the behavior and the usefulness of the network for teachers. This technique could help all stake-holders, e.g. policy-makers, the members of the eTwinning network (i.e. teachers, we also call them "eTwinners"), eTwinning Central Support

[2] www.etwinning.net
[3] http://www.etwinning.net/en/pub/misc/privacy_statement.htm
[4] http://www.etwinning.net/en/pub/tools/twinspace_tools.htm
[5] http://www.etwinning.net/en/pub/professional_development/learning_events.htm
[6] http://www.etwinning.net/en/pub/professional_development/etwinning_groups.htm
[7] http://www.etwinning.net/en/pub/tools/desktop_tools.htm

Service, 35 National Support Services and eTwinning Steering Committee, understand the underlying mechanisms for the transfer of good practices and innovation. More importantly, it eventually may help teachers understand their position in the network and use it to advance cooperation among other teachers and schools.

This deliverable reports on the social network analysis of eTwinning network performed in WP 2, "Social Network Analysis methods for lifelong learning communities". We proposed several social network analysis methods to deepen the understanding of the structure, characteristics and the dynamic of eTwinning network as well as teachers' social capital in the network. Concretely, we focus on the following objectives:

- **Data model**: In deliverable 2.1, a data model has been proposed to store information about teachers and their activities. Advanced data warehouse techniques are used to transfer aggregated data from the operational database to deliver diverse data views for further analysis. For the dynamic social network analysis, this data model was extended in order to store dynamic information, e.g. network snapshots and time series data.
- **The dynamic of eTwinning network**: Different social network analysis measures and time series analysis methods are applied to understand the structure, characteristics and the dynamic of the various eTwinning networks. Specifically, we focused on the community aspect of eTwinning networks created by different ICT mechanisms and their development trends.
- **Social capital in the eTwinning network**: Social capital (Coleman, 1988) is a concept that has been extensively studied in social science. It stands for the ability of actors to derive benefits from the membership in social networks or other social structures. We study social capital in eTwinning network from two points of view: as a property of the teachers and as a property of the working groups (here we consider members of a project as a working group). For the first view, we identify the position or local network structure of teachers who are effective in performing project related tasks (i.e. effectiveness is measured through Quality Labels). Based on time series analysis, we are able to track the development patterns of teachers over time and provide hints for their future career path within eTwinning.
  For the second view, we analyze the early communication and collaboration of project members that has taken place before the project starts and identify the most effective structure that can lead to high quality and achievement. The result is helpful for organizers to initialize and find partners for the project.

This deliverable is organized as follows. In Section 2, we provide an overview of basic concepts and SNA methods used in our social network analysis. Section 3 describes some extensions of the existing data model to handle dynamic and time series data. Section 4 presents a structural analysis of the eTwinning network. In Section 5, we identify and verify two forms of social capital in the eTwinning network: structural hole and closure. In Section 6, the development pattern of teachers is analyzed and presented. Section 7 wraps up the deliverable with some conclusions and future work.

2. Social capital and dynamic social network analysis

Social Network Analysis (SNA) methods have been applied in many domains, e.g. knowledge discovery in digital libraries (Pham & Klamma, 2010; Pham et al., 2011) and analysis of community of practice (Kienle et al., 2005, 2006; Hoadley, 2005). SNA is extensively used in sociology to study social behavior of actors, emerging phenomena, such as small-world effect (Travers & Milgram, 1969, Watts & Strogatz, 1998), and diffusion of innovation – the spread of new ideas and technology through cultures (Rogers, 2003). Among other applications of SNA in sociology, social capital has been studied by means of SNA methods. In this section, we review the notion of social capital and basic SNA measures used to identify the network structures that are more effective in creation of social capital. We also review basic techniques for dynamic network analysis to trace the development of networks. With this review, we set up a general framework for the upcoming analysis of the eTwinning community.

2.1. Social capital

Social capital stands for the ability of actors to derive benefits from the membership in social networks or other social structures. Social capital can be viewed as a property of a group, where some groups are more effective than others because of their social structure. Social capital can also be viewed as a property of an individual where a person can have more or less social capital depending on their positions in the network (Borgatti, 1998).

In social network research, studies are concerned with the identification of network structures that are the most effective factor for creating social capital. Two types are identified. Coleman (1988) emphasizes the benefits of being embedded into densely connected groups, as regards to the confidence, trust and secured relationship in the community. This form of social capital is referred to as *closure*. On the other hand, Burt (2001) discusses social capital as a tension between being embedded into communities and brokerage - the benefits arising from the ability to "broker" interactions at the interface between different groups. We refer to the form of brokerage as *structural hole*.

To measure the social capital of a node in the network, a set of centrality measures can be applied. In the following, a social network is defined as a graph $G= (V, E)$, where V is the set of nodes and E is the set of connections (edges) between nodes. If node v is connected to node u then u is called a neighbor of v. Network G can be undirected or directed, unweighted or weighted. If G is a directed network, then the direction of edges is considered. If G is a weighted network, each edge is associated with a real number to define the weight of the edge. The centrality measures of a node in the network are defined as follows (Wasserman & Faust, 1994):

- *Degree* of a node measures the number of connections to it. In a directed network, we define in-degree as the number of in-coming connections to a particular node and out-degree as the number of out-going connections from that node. For example, in eTwinning, if a teacher participated in many projects with many different teachers, her degree in the project collaboration network is high. In contrast, if a teacher repeatedly collaborated with the same teachers (even in many projects), her degree is low. *Local clustering coefficient* of a node measures the probability that two neighbors of that node are connected to each other:

$$C(v) = \frac{\left|(w,u \in N(v) : (w,u) \in E)\right|}{\frac{\left|N(v)\right| * (\left|N(v)\right| - 1)}{2}}$$

where *N(v)* is the set of neighbors of node *v*. Local clustering coefficient measures the extent to which a node is positioned in a dense-connected cluster. If a node is in a dense cluster, its local clustering coefficient is very high (e.g. node *A* in Fig. 1). For example, in eTwinning, if a teacher collaborates with teachers who have already collaborated with each other before, her local clustering coefficient in project collaboration network is high. On the contrary, if the node is positioned in the border of a cluster and connects different groups, its local clustering coefficient is relatively low (e.g. node *B* in Fig. 1). For example, in eTwinning, if a teacher collaborates with teachers who have not collaborated with each other before, her local clustering coefficient in project collaboration network is low.

- Betweenness measures the extent to which a particular node lies between the other nodes in the network:

$$B(v) = \sum_{i \neq j \neq v} \frac{\sigma^{v}(i, j)}{\sigma(i, j)}$$

where *B(v)* is the betweenness of node *v*, $\sigma^{v}(i, j)$ is the number of shortest paths between node *i* and *j* that pass through *v* and $\sigma(i,j)$ is the number of shortest paths between node *i* and *j*. In contrast to high clustering coefficient nodes, nodes with high betweenness have more power to control the information flow in the network and are normally the gatekeepers (or structural holes) who connect several dense groups (e.g. node *B* in Fig. 1). For example, in eTwinning, if a teacher collaborated with teachers in different projects and those teachers have not collaborated with each other before, her betweenness is high.

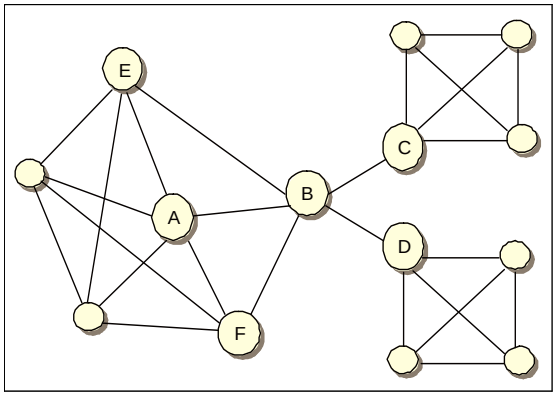


Figure 1: Structural property of social capital

In summary, the degree of a teacher measures the number of connections that she has in a particular network. It shows how active a teacher is in making contacts, collaboration and communication, but it does not show the position and communication as well as collaboration strategy of the teacher. Local clustering coefficient and betweenness reveal the structural properties of a teacher in the network. High local clustering coefficient and low betweenness values of a teacher indicate that she is positioned in a tightly connected community (closure),

while low local clustering coefficient and high betweenness values indicate that she is located at the interface of different communities or teachers as a star (structural hole).

With the above centrality measures, we are able to identify the position of teachers in the eTwinning network and how it relates to the performance of teachers, e.g. in carrying out projects and winning Quality Labels and awards. At promoting eTwinning Portal to more European teachers, eTwinning management staffs at the Central Support Service are highly interested in teachers' activities and possible progress on the eTwinning Portal. In order to recognize teachers' performance in project cooperation, additional values such as "Quality Label" and "European eTwinning Prizes" are applied to assess teachers' achievements within eTwinning. This additional information may help teachers and management staff learn about teachers' performance and/or professional development path. The correlation between teachers' position and their achievements helps verify the form of social capital in eTwinning network By the dynamical analysis of the centrality measures of teachers in the network over time, we can track and monitor teachers' development pattern and further give suggestions for the future development path.

2.2. Dynamic social network analysis

Dynamic network analysis (DNA) is an emergent research area that brings together traditional Social Network Analysis, link analysis and agent-based system into network science (Carley, 2003; Barrat et al., 2008). DNA deals with involving networks, i.e. networks that change over time (e.g., nodes and edges are added or removed, properties of nodes and edges change, etc.). The main difference between SNA and DNA is that DNA takes into account time dimension and the properties of the whole network, nodes and edges are supposed to change, while these properties in SNA are relatively static. Research in DNA are concerned with development of metrics, statistics, models and simulations to access and predict the evolution and changes of the networks as well as of nodes and links.

Methods in DNA can be grouped into two main categories. *Microscopic* approach (Barrat et al., 2008), which is well known in physics community, models the nodes as dynamic entities that change their properties while interacting with each other. *Macroscopic* approach focuses on the network evolution at global level, i.e. it studies the laws that govern the growth of the whole networks and then hypothesizes about the node's behaviors that would reproduce or result from the laws observed at global network structure (Leskovec, 2008). Here we follow the second approach: we first observe the structure and evolution of eTwinning network at global level then we analyze network structure of teachers at node level. Especially, we focus on the community structure and teachers' behaviors that makes the network be clustered. Social capital is a special property of teachers that results from global network structure.

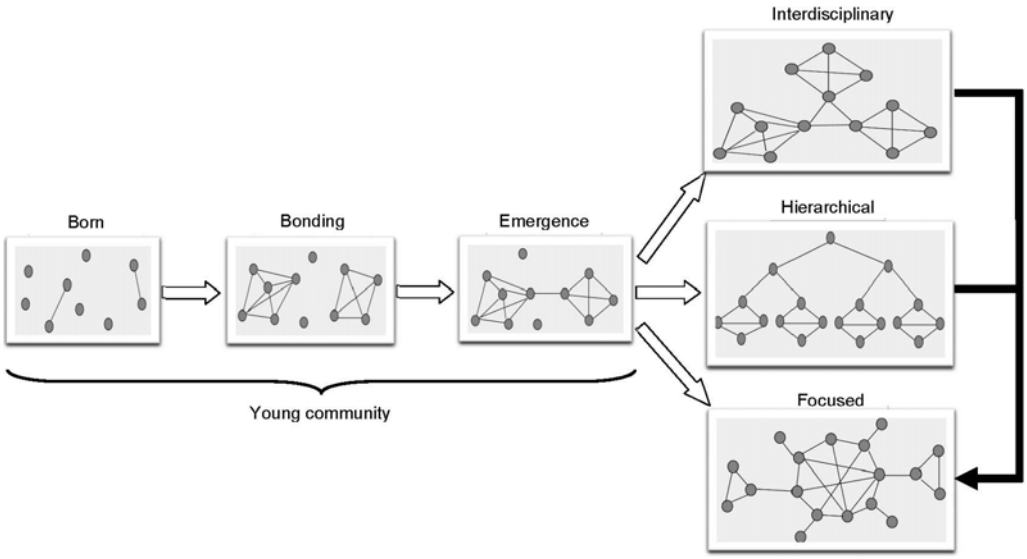


Figure 2: Community development model

To analyze the development of scientific community of practice, we have proposed a model as depicted in Figure 2 (Pham et al., 2011). For the teacher networks, initially there are few connections (*born* stage). After some time, teacher groups become apparent in the network (*bonding* stage) by carrying out projects together or communicating with each other via internal messaging and other communication tools. These groups are gradually integrated through projects that involve teachers from more than one group (*emergence* stage).

Over time, the teacher network can form a network topology that features a strongly connected core group of teachers that is connected to other smaller groups (*focused* stage). Alternatively, the network can develop toward an *interdisciplinary* typology where several groups are connected via some gatekeepers, but where there is no core group. It can also develop toward a *hierarchical* topology that exposes some "super gatekeepers" connecting a hierarchy of groups together. In workpackage 2, we used this model to describe and explain the development of teacher networks. We also used this model to analyze the optimal structure of member networks and its correlation to the achievement of recognition in eTwinning (e.g. Quality Labels). That reveals the form of group social capital as mentioned in Section 2.1.

To qualitatively characterize the development process of the teacher networks in relation to this development model, we applied time series analysis on the networks to reveal the following social network parameters of the networks over time: Densification law (Leskovec et al. , 2007), clustering coefficient, maximum betweenness, largest connected component, diameter, and average path length (Wasserman & Faust, 1995). These parameters are explained in the following paragraphs to describe the community building process in Figure 2. To interpret the shape of the community, one needs to use a combination of all of those parameters.

Formally, given the network $G = (V, E)$, where V is the set of vertices or nodes, and E is the set of edges, the above network metrics are defined as follows:

- *Densification law*: Leskovec, Kleinberg, and Faloutsos (2007) discovered that complex networks densify over time, with the number of edges growing super-linearly

with the number of nodes, meaning that the average degree (i.e., number of edges) of the nodes is increasing. In fact, the densification follows a power-law pattern:

$$e(t) = n(t)^{\alpha}$$

where (e(t) and n(t) are the number of edges and nodes at time t, respectively, and α is an exponent that lies between 1 and 2 (α = 1 corresponds to constant average degree over time, while α = 2 corresponds to a very dense graph where on average each node has edges to a constant fraction of all nodes). We use exponent α to differentiate the "speed" by which networks are densified.

- *Global clustering coefficient* measures the probability that two nodes are connected if they already have had a common neighbor:

$$C = \frac{3 \times Number\ of\ triangles\ in\ the\ graph}{number\ of\ connected\ triples}$$

Intuitively, during the first phase of development the clustering coefficient of the network is low, since nodes are disconnected from each other. In the second phase, the clustering coefficient tends to increase very quickly as nodes are clustered into very dense, yet disconnected components. When the disconnected components subsequently start to connect with each other, the clustering coefficient drops and stay relatively stable after some time.

- *Maximum betweenness* is the highest *betweenness* (defined in Section 2.1) of nodes in the network. For the overall network, the maximum betweenness is a good indicator of identifying whether there are strong gatekeepers (or structural holes) who connect different communities within the network. During the first two phases of the development process, the maximum betweenness is very low since the nodes are either completely disconnected or clustered in very dense yet disconnected groups (i.e., there are no controllers in the network). Maximum betweenness increases when more components become connected (*emergence* stage) and continues to increase when the network develops toward a *hierarchical* or *interdisciplinary* topology. However, maximum betweenness will reach a stable value when the network is at *focused* stage.
- *Largest connected component* (or giant component) measures the fraction of nodes that are connected with each other in the largest sub-network. As observed in Figure 2, this fraction is small in the first two phases and gradually increases as the network develops and actors from different sub-networks connect with each other. It reaches a stable state when the fraction of nodes that connect to the largest component is equal to the fraction of new nodes that stay disconnected from the largest component. Basically, *hierarchical*, *interdisciplinary* and *focused* topologies have a large connected component.
- *Diameter* is the length of the greatest geodesic distance (i.e., the length of the longest shortest path) between any two nodes. Intuitively, at the beginning, the diameter is small and then it increases. After some time, the diameter starts to shrink as new edges between existing nodes continue to be added. Note that the shrinking of the diameter is not caused by the emergence of the giant component (Leskovec, Kleinberg & Faloutsos, 2007). However, in our model, if the network develops toward a tree-like

topology (*hierarchical* stage) the diameter will be larger than in the focused and interdisciplinary topologies, respectively.

- *Average path length* is the average length of all shortest paths in the network. Clearly, at the first two phases, the average path length is small and increases when the network grows. Although teacher networks are not random networks — and the average path length should therefore be rather small (around six, according to the small-world effect) — there is a slight difference between *focused*, *interdisciplinary* and *hierarchical* topologies. In general, the average path length of a hierarchical network is larger than that of the other two topologies, which gives us more evidence to differentiate those topologies.

In summary, the emergence of the giant component (largest connected component) indicates the cohesiveness of teacher networks, while the betweenness shows the existence of the gatekeepers and their importance. The clustering coefficient measures the extent to which the network is clustered into sub-communities. Other parameters like diameter and average shortest path length show whether the community is still developing or whether it is stable. The densification law shows the speed by which the network is densified.

3. Extension of the TeLLNet data model

The basic data model has been established and reported in the report of Deliverable 2.1. In this section, this model is shortly reviewed and has been extended for dynamic social network analysis methods, i.e. the time-series analysis.

3.1. Review of the existing data model

The TeLLNet data model is established with the main entities of teachers, projects, project diary and messages on teachers' wall (i.e. journal). Among them, teachers are the actors who carry out interactive activities including cooperating on projects, writing and commenting on project diary, posting messages on their own or other teachers' journal (wall), posting in project guest book, commenting on the reception of project prizes' application forms, and collecting contacts, etc. Additionally, teachers have their own profiles such as subject, language, institution (school), and country information, and also this piece of information is taken into account.

Based on this data model, we were able to explore teachers' personal information and community information. Personal information is grouped into basic information and competence-related factors. Basic information includes teachers' ID while teachers' names are anonymous in the TeLLNet data set. Teachers' registration date, country, languages, and subjects are considered as well. Factors focus on the network data including numbers of joined projects, activities using social messaging systems, blogging, commenting on project guest books, and commenting on project prize's applications. At the level of community, an individual teacher's performance is compared to the other members of their communities. Two aspects are taken into considered to specify teachers' belonging communities. On the one hand, an individual teacher is a member of a school. On the other hand, the teacher is a member of projects with many other project members. In addition, the interactive activities among the community members are analyzed.

Temporal information of the aforementioned activities is recorded (i.e. a time stamp). It can reveal information about eTwinners activities and performance with a period of time on the eTwinning portal. However, social capital is not the analysis resulting of single performances on the teachers' individual level, but correlated to the network structure and the playing roles in the whole network or sub networks. Hence, the existing assessment based on single indicators such as activities on eTwinning portal (e.g. number of journal or blog posts, projects) is not able to show teachers' social capital from various aspects.

The dynamic network analysis considers various users' requirements in eTwinning and exploits the large-scale lifelong learning data set of eTwinning. There are different stakeholder groups who are interested in the data set:

- Decision makers in order to pursue a better future for European school education and for European teachers' professional development and lifelong learning, which influence school education greatly.
- Central Support Service (CSS) i.e. learning network management staff. They are interested in questions like "What is the relationship between eTwinning and professional development? How do they influence each other and how do they support each other? How can eTwinning contribute to professional development, and vice versa?" (Vuorikari, 2010)
- National Support Service (NSS) promotes eTwinning by providing training and support, organizing meetings and national competitions, and managing public relationships based in each country.
- School management staff is interested in teachers' activities on the eTwinning portal, especially regarding the development of teachers' ICT skills and project cooperation issues.
- Teachers from pre-school, primary, secondary and upper schools can all participate in eTwinning to exchange and collaborate, as well as to learn new ICT skills, communication skills, teaching skills, and interdisciplinary working skills.

3.2. Extension of data warehouse for dynamic social network analysis

In order to advance the time series analysis, time point information is attached to the networks and centrality measures of the networks as a main extension of the existing data model. As depicted in Figure 3, for each network, an additional table is created. This network data models the edge and node information as usual, with explicit information about the temporal information, i.e. the date. We have started with project, project diary, and my journal messaging networks, as the contact network is lacking such temporal information. The time dimension is added to all parameters which can be grouped into node, cluster, and the whole network. To populate the data for dynamic analysis, data warehouse techniques such as data transformation and loading are applied. Basically, the data is processed as follows. First, raw data is loaded into the data warehouse. In the second step, networks are created using store procedures in the database. Then network parameters as well as node centrality measures are computed and loaded to database tables. To deal with dynamic networks, parameters are computed for each snapshot of the networks at a certain time point created by the temporal information which is stored together with the nodes and edges.

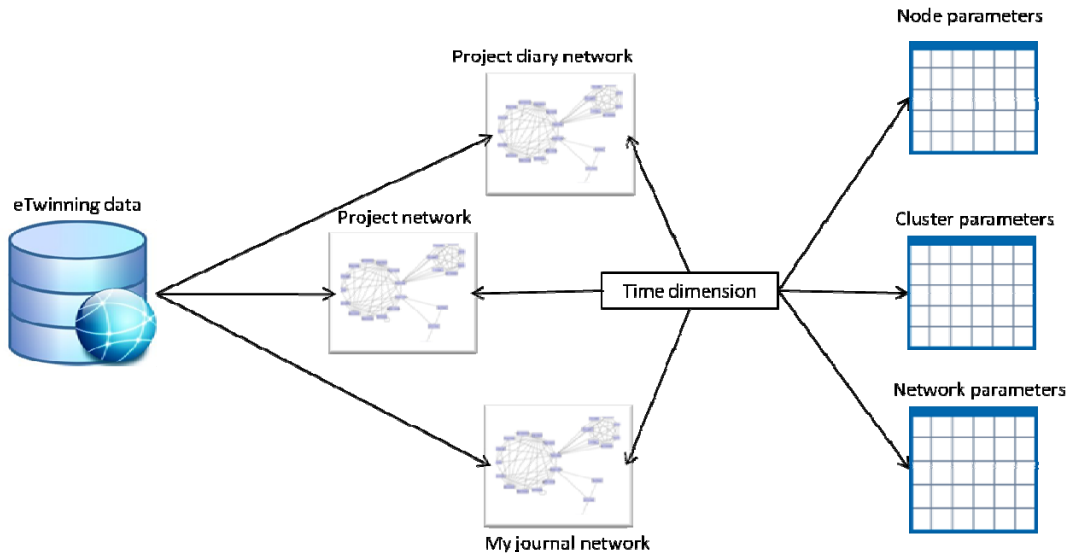


Figure 3: Extension of existing data model

The statistics of the eTwinning data related to our study is given in Table 1, specifically, data used to create networks including project collaboration, contact list, project diary and my journal messages. European Schoolnet provides the data set with anonymous teachers' information due to data privacy.

Table 1: Statistics on eTwinning data (as of 11.11.2011)

Data	**#data entries**	**Description**
Project	23641	Schools from at least two schools from at least two different European countries create a project and use ICT to carry out their work.
Contact	769578	Teachers are able to explore other teachers' profiles and add them into their own contact list. It is suggested to use internal messaging and other media to establish a relationship among teachers before accepting them as a contact.
Project diary	20963	Blog for project activities
Project diary post	49604	Each blog entry in project diary
Project diary comment	7184	Comments added to blog entries in project diary
My journal message	216704	Messages posted on teachers' wall which is part of teachers' profile
Teacher	146105	Registered teachers working in European schools and, namely "eTwinner"
Quality Label	8042	The QL is awarded to a teacher working on a project after they have applied for it through their National Support Service (NSS). The application includes information about the project and the work carried out by the teacher. Each NSS approves QL applications according to their contextualized procedures. There are National Quality Labels and European Quality Labels.
Prize	1384	eTwinning Prizes are awarded to projects. They are of European level and are called European eTwinning Prizes. Each Prize is connected to an application submitted by the eTwinner.
Institution	91077	Various European schools: pre-school, primary, secondary and upper schools

We conduct dynamic time-series analysis based on this data set with a focus on project, project diary and my journal messaging networks. Centrality measures of these networks are computed and stored along the time periods beginning with the starting time of the eTwinning portal. The contact network data reflects teachers' social capital even without the time dimension.

Teachers' performance is recognized in eTwinning in formal ways. Formal recognition consists of "Quality Label" and "European eTwinning Prizes" as specified in the above table. It is important to note that each NSS awards Quality Labels only after the application has been submitted by the eTwinner. The National Quality Labels are warded according to national and local requirements which might vary from NSS to NSS. The value of the Quality Label to its holder also varies from country to country; whereas in some countries they are well-valued by eTwinners, in some other countries only a few eTwinners apply for such recognition (this is usually the case in Northern part of Europe). Only in 2011, a common criterion was established for the European Quality Label. In addition, eTwinning Reach is a new measure concept for country assessment and specified as the registered users of a country / teacher population of this country (Vuorikari et al., 2011). Again, this assessment has its limitations and needs statistical support from SNA-based learning analytics.

4. Properties of eTwinning networks

The first questions we try to address are: what is the structure of the teacher networks? How do these networks evolve over time? What are the collaborative and communication behaviors of teachers? (Question 1)

Specifically we concentrate on the community aspect, i.e. whether teachers collaborate and communicate with teachers in their community or across different communities. The answers for these questions give us an overview of eTwinning network and are the basic for the analysis of social capital presented in the next section.

We create four networks using existing data from the eTwinning platform, namely project collaboration between teachers, contact information, project diary, and my journal (wall message). These mechanisms are offered in the eTwinning platform for collaboration and communication between members.

Figure 4: Communication tools in sTwinning

In eTwinning, teachers can take part in and carry out projects with other teachers. Usually, teachers post their ideas about new projects, or their willingness to participate in projects, in the Partner finding forum. This data is not part of the Tellnet data model. In the Partner finding forums, teachers view the proposals and contact each other to further develop the project idea using for example Internal messaging (also not part of the Tellnet data model). Additionally, the Journal can be used for messaging, either only for social network purposes, or also to invite others teachers who may be interested in joining the project. When a project starts, the Project diary is used to keep track and report about the progress. All the communication tools used in eTwinning are depicted in Figure 4.

All the information on the eTwinning portal about projects, teachers' contacts, project diary posts and comments as well as wall posts and comments are recorded. Using this information, we created four networks as follows:

- *Project network (G_P)*: nodes are teachers (eTwinners) and there is a connection (edge) between two teachers if they collaborate in at least one project. Edges in the network are undirected and weighted by the number of projects in which the two teachers collaborate.
- *Contact network (G_C)*: nodes are teachers and there is an edge between two teachers if at least one teacher is in the contact list of the other. Edges are undirected and unweighted.
- *Project diary network (G_B)*: nodes are teachers and there is an edge between two teachers if one teacher has commented on at least one blog post created by the other. Edges are directed and weighted by the number of comments.
- *My journal network (G_J)*: nodes are teachers and there is an edge between two teachers if one teacher has posted or commented on the wall of the other. Edges are directed and weighted by the number of messages.

For time series analysis, we aged each of the aforementioned networks by month, e.g., the project network at 01.10.2008 contains a set of teachers and the links between them in the projects that started from this time backward. The following time information was used:

- the Project network: time information from 14.01.2005 to 11.11.2011 (82 months)
- the Project diary network: time information from 28.09.2008 to 11.11.2011 (37 months)
- The Journal (wall) network from 28.09.2008 to 11.11.2011. (37 months)
- The contact network does not contain time information.

This data enables us to perform the analysis on the evolution of project, project diary and my journal networks within 82 and 37 -month periods accordingly.

4.1. Static properties

We compute global measures as defined in Section 2.2 for four networks. The results are presented in Table 2. Several observations can be made here. First, teachers are more active in project, contact, and my journal networks (26%, 75% and 30% of registered teachers, accordingly) than in project diary network (only 2.2% of registered teachers). Project and project diary networks are denser, while contact and my journal networks are quite sparse. Project and contact network are better connected than project diary and my journal networks. This can be seen by the average path length, diameter, the number of disconnected

components, and the size of giant component in these networks. The giant component exists in all four networks, but it is bigger in contact networks (99%).

Second, project network exposes a strong community structure with the clustering coefficient of 0.7308, while contact, project diary and my journal networks do not (clustering coefficient are 0.1135, 0.1317, and 0.028 respectively). Connection between teachers in contact, project diary and my journal networks are clustered in disconnected groups, but these groups are not well connected. These results raise our special interest in exploring the community structure of the eTwinning project collaboration network. Specifically, we would like to see whether project collaboration of eTwinners depends on a core community.

Table 2: Properties of eTwinning networks

Network	**#nodes**	**#edges**	**Clustering coefficient**	**Average path length**	**Diameter**	**Number of disconnected components**	**Largest component size**
Project	37893 (26%)	804825 (0.11%)	0.7308	3.948996	19	2493	31480 (83%)
Contact	109321 (75%)	573602 (0.01%)	0.1135	4.319833	13	506	108140 (99%)
Project diary	3162 (2.2%)	3211 (0.06%)	0.1317	6.429112	21	453	2032 (64%)
My journal	43863 (30%)	56138 (0.006%)	0.028	6.54689	21	2789	37179 (85%)

We detect the communities in project network using a clustering algorithm proposed by Newman (2004) and get 3086 clusters (including disconnected components) with their sizes as listed in Table 3. The modularity calculated by the algorithm indicates the quality of clusters. It is the fraction of any node's connections within its cluster and its connections to other clusters. Empirical observation indicates that the modularity greater than 0.3 corresponds to significant community structures. With the algorithm eTwinning project network receives the modularity equals 0.47, corresponding to a significant clustering of the project network.

This result shows that eTwinning collaboration depends on a main core of four large communities (LC1-LC4). There are 23578 (16%) eTwinners who are part of these four clusters. These clusters are formed by eTwinners who have collaborated with each other in a high number of projects over a long period. Note that the project network has a big number of disconnected components (2493 components, Table 2), these represent project partnerships of teachers who are engaged in project collaboration for the first time. Therefore from Table 3 we can see that 589 small communities (3086-2493-4=589)are part of the giant component and they are connected to the core (four clusters) via many gatekeepers.

Table 3: eTwinning project network clusters

Cluster size (number of eTwinners)	**10567 (LC1)**	**6277 (LC2)**	**4362 (LC3)**	**2372 (LC4)**	**100-1000**	**10-100**	**2-9**	**Total**
Number of times identified	1	1	1	1	12	166	2904	3086

4.2. Dynamic properties

To track the development of eTwinning networks, we analyze the network parameters of each network over time. At each time point, we have taken a snapshot of each network and computed its parameters, as defined in Section 2.2. We plot these parameters over the course of observation, which in the corresponding graphs is called "Age of the network". The age of the network is 82 months for project network, 37 months for project diary and also for my journal network. This allows us to observe the evolution of the networks and the stage to which they tend to develop toward, according to the model depicted in Figure 2.

The plots of six parameters of project, project diary and my journal networks are given in Figure 4 (we do not consider contact network here since it does not have time information). Figure 4 (a) plots the number of edges against the number of nodes over time. Figures 4 b, c, d, e, f plot the clustering coefficient, the maximum betweenness, the fraction of largest connected component, the diameter and the average shortest path of the networks against their age . Age 1 means the first month, 2 the second month and so forth.

From Figure 5a, we see that project network is densified at a very high speed (exponent α = 2), while project diary and my journal networks are densified much slower (exponent α = 1.2 and 1.1 respectively). Moreover, a closer look at other parameters reveals that the project network was developing toward a cohesive and well-connected network quite early, around month 9^{th} This is indicated by the following observations: fast increasing of clustering coefficient (see Figure 5b) and largest connected component (see Figure 5d); the decay of diameter (see Figure 5e) and average shortest path (see Figure 5f); and low betweenness (see Figure 5c). In fact, the project network developed to the *focused* stage by month 25^{th}, as depicted in the model of Figure 2. This means that it is now mature.

On the other hand, Figure 5 also shows that project diary and my journal networks are developing toward a tree-like topology (*hierarchical* stage) without any community structure. This is shown by very low and no-increasing clustering coefficient which indicates no community structure (see Figure 5b), a very high and fast increasing betweenness which represents the root of the tree (see Figure 5c), high diameter and average path length (see Figure 5e and 5d). Last, from Figure 5 we can observe that all three networks are stable now; this can be seen in the stable values of five network parameters (Figures 5(b, c, d, e, and f)).

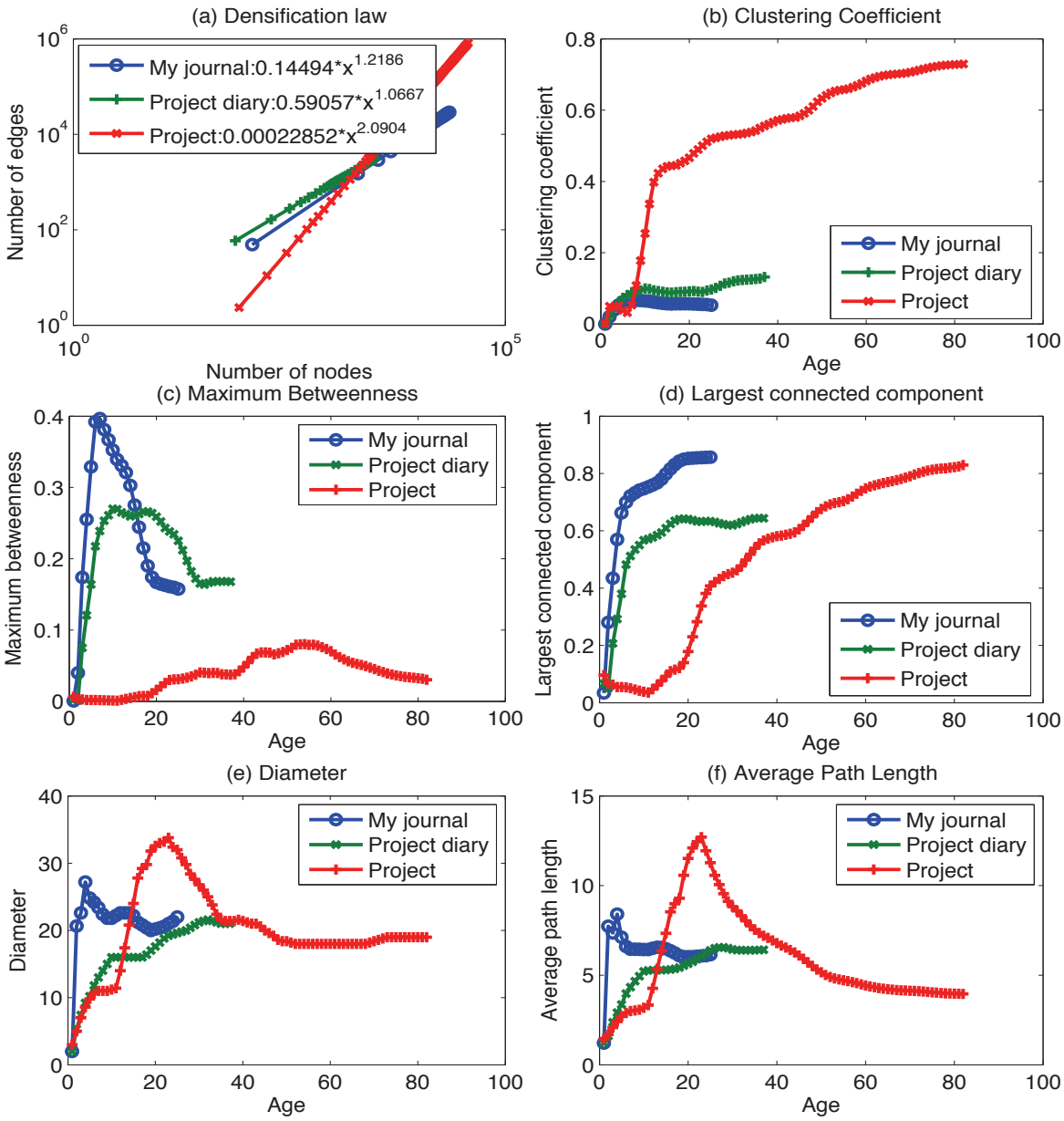


Figure 5: Properties of teacher networks over time

The observations above allow us to draw some conclusions about teachers' communication and collaboration in eTwinning. Although teachers are very active in project collaboration, their communication via eTwinning mechanisms is quite limited (it is known that teachers also use other ways to communicate, for example, their email, evidence of which is not traceable on the eTwinning platform, as well as the internal messaging system. Communities have not emerged in project diary and my journal networks, though the phenomena can be clearly seen in project collaboration network.

A possible explanation is that eTwinning platform lacks effective tools to keep teachers updated about their colleagues. For example, in Facebook, friends' status and activities are automatically updated on users' wall and they can easily see what is happening within their

community. Such notification system also exists in eTwinning, but due to privacy constraints, it is an opt-in service which is currently little used.

4.3. Summary

To summarize, teachers are more active in project collaboration and contact networks than in project diary and my journal networks. The Project network exposes a strong community structure, while other networks are clustered into disconnected components each component is in a star topology and not well connected. The community structure analysis of project network shows that the network depends on a core of 5 large clusters. A large number of small clusters connect to the core via many gatekeepers.

The time series analysis the project network is developing well and it is mature now. With the time, it most likely will keep evolving as it is. While the other two networks, project diary and my journal networks, are still growing. This result can be interpreted to show that the use of ICT mechanisms offered by eTwinning platform in communication between teachers is quite limited and it can be speculated that majority of communication takes place without using these two tools (e.g. either outside of the platform or by using the internal messaging system, for which Tellnet data does not exist). It could be suggested that tools which can support community building, could be helpful for teachers and help them to align themselves in the right community and projects, as well as increase the quality of projects by selecting the right members. Such tools are for example status currently are offered on the platform, but only as an "opt-in" service due to privacy concerns. Other such tools could be recommendations on people and projects.

5. Social capital in eTwinning networks

In this section, we study the social capital in teacher networks. As mentioned in Section 2.1, social capital can be viewed as either a property of a group or a property of an individual.

- For group social capital, we are concerned with the network structures of the group members that make it function more effectively than others do.
- For individual social capital, we study and identify the location of the individual that has more advantage in performing tasks.

In eTwinning, we consider projects as groups of members and study the correlation between the group's member network structures and the achievement of the projects, a manifest of which are Quality Labels. The performance of teachers is also recognized by the Quality Labels that they received through projects.

5.1. Structural hole and closure in eTwinning networks

As described in Section 2.1, structural hole and closure are two important forms of social capital.

- Structural hole stands for the location of a node that is positioned at the interface between different communities and is identified by high betweenness.
- Closure is the node that is clustered in a densely connected group and therefore has high local clustering coefficient.

Moreover, in network theory, there are three main types of networks:

- random network where links are randomly placed between nodes;
- small world network that has small average shortest path length and high clustering coefficient; and
- scale-free network where the degree distribution follows power law.

Here we study structural hole and closure as properties of teachers (individuals) as well as projects (groups). Question 2: do structures such as structural hole and closure exist in eTwinning networks?

First, we identify the class of teacher networks to answer this question. Figure 6 depicts the degree distribution of nodes in our four networks; the project, contact, project diary and my journal networks respectively. We see that all four networks are scale-free networks with a power law degree distribution (a fat-tail distribution), with exponent $\alpha = 1.455, 1.933, 1.625, 1.723$ for project, contact, project diary and my journal networks respectively. The power law degree distribution indicates that super connectors (or hubs) exist. Super connectors (or hubs) refer to those nodes that connect many nodes or communities. They play an important role to ensure the connectivity, the information spreading and behavior cascading in networks. They also have more power and control over the network than the other nodes which lie in the tail of the distribution.

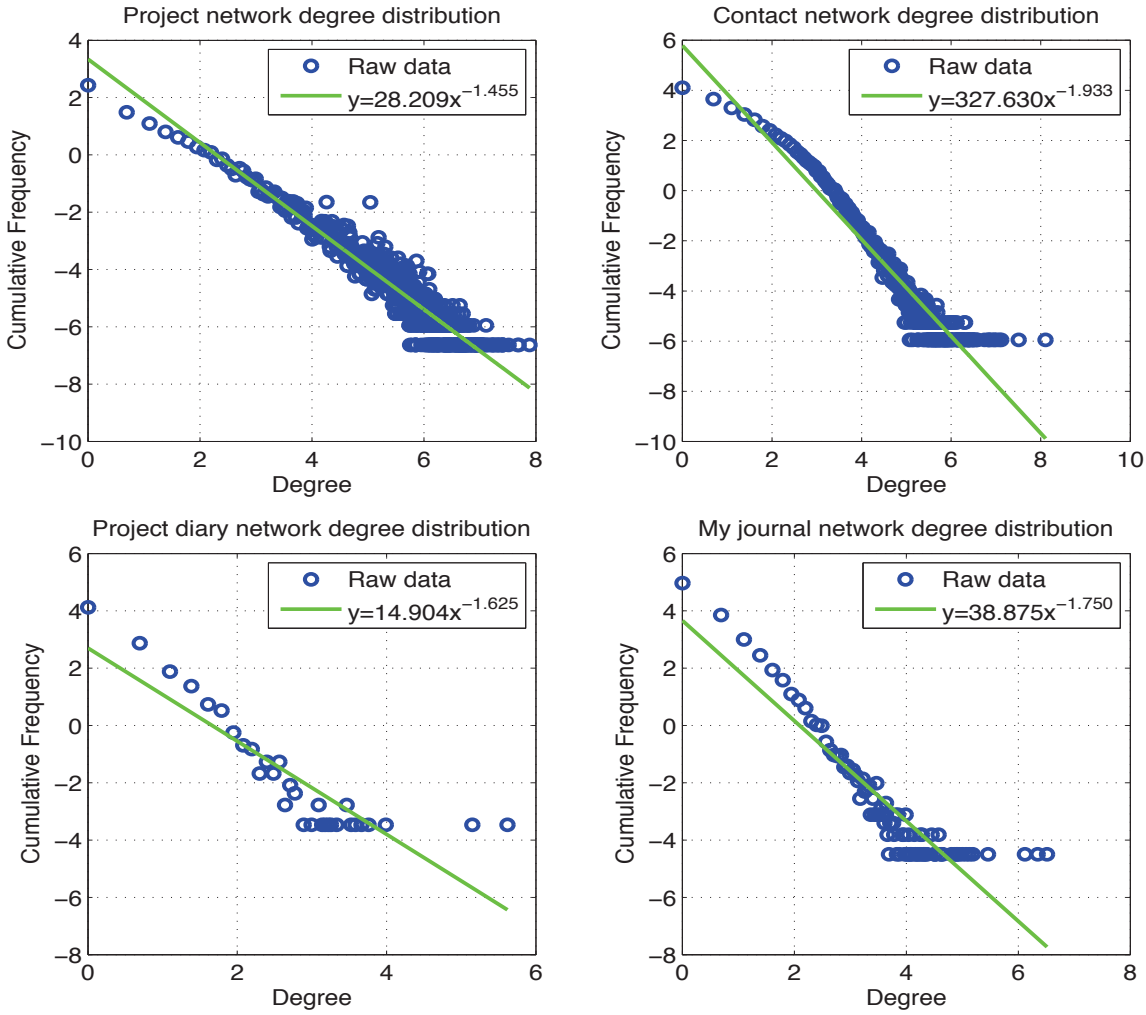


Figure 6: Power law degree distribution of teacher networks

Moreover given the hubs in the networks, what is their local structure? Are they structural holes or closures? Figure 7 plots the betweenness and local clustering coefficient as a function of node degree. We can see a clear trend in project, contact and my journal networks (Figure 7a, b and d): betweenness increases and clustering coefficient decreases as degree increases. It means that teachers in project and contact networks with high degree of centrality are positioned at the interface of different communities. These teachers can be seen as structural holes in the network.

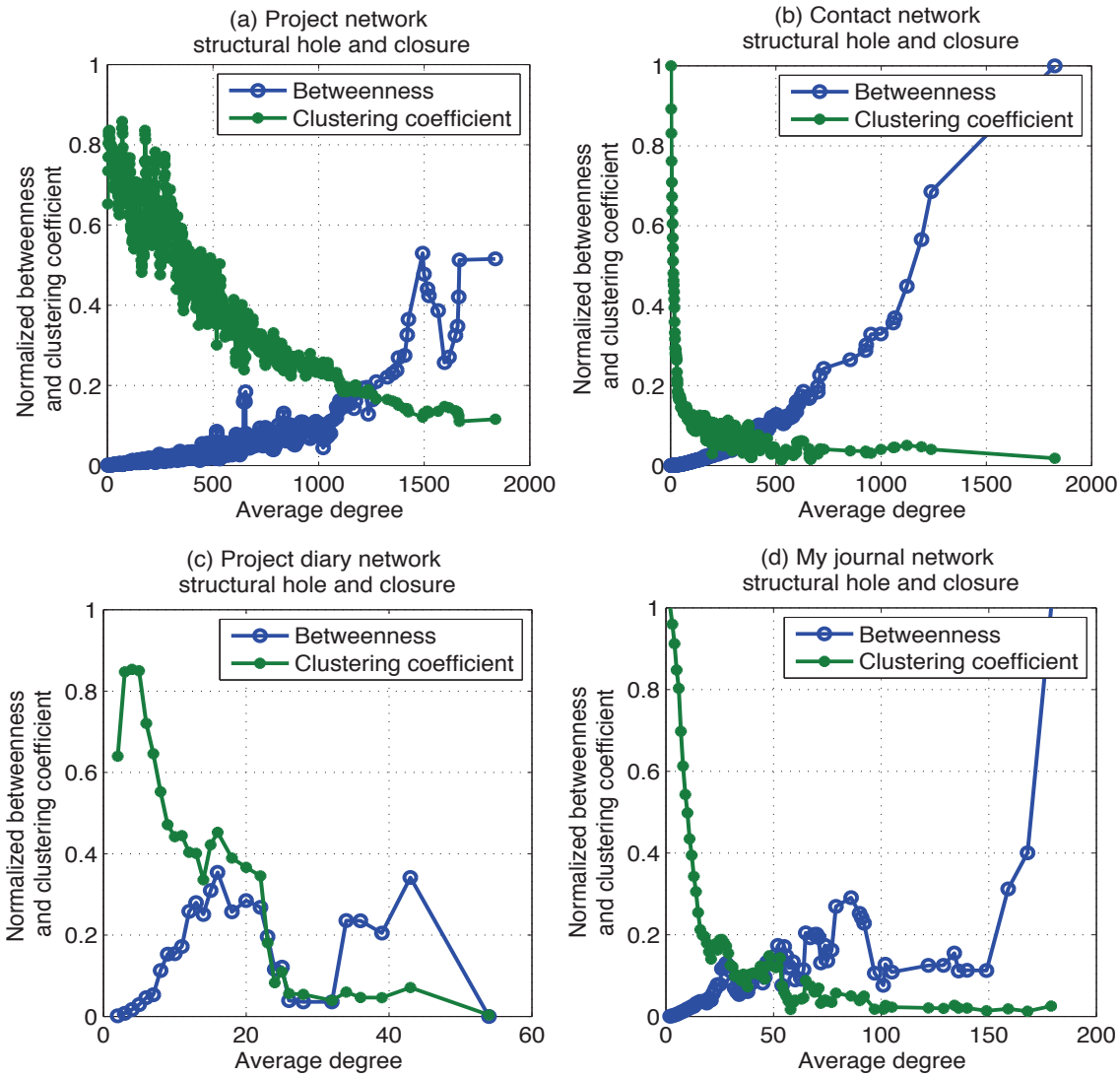


Figure 7: Structural hole and closure of teachers

We can also observe this trend in project diary network for teachers with low degree (less than 20). As for greater degree in this network, betweenness and clustering coefficient do not follow this trend: although the clustering still decreases, the betweenness does not increase as degree increases. We can see the reason if we connect this observation with the properties of project diary network as discussed in Section 4. In this network, high degree teachers are placed in the center of a local group like a star, where many other teachers connect to the stars, but few of them connect to each other. Stars also do not connect to other stars in the network. Few connections between members surrounding a star create shortcuts and therefore decrease the betweenness, while they are not enough to increase the clustering coefficient.

This observation is another evidence for the limitation in using project diary in communication within eTwinning community. Although my journal network also develops towards a tree-like topology, but the stars in this network are connected to each other. Therefore, clustering coefficient and betweenness of my journal network still follow the same trend as of project collaboration and contact networks.

Question 3 investigates the structure of the project member network. Specifically, we attempt to understand the strategy to find partners for a new project: whether teachers look for the partners that they have already collaborated with. It helps gain an insight into the preparation phase of the projects which leads to the high level accomplishment of the project. We process the data as follows. For each project, we take its members and find the links between them in the project network that are created *a priori* to the starting time of the current project. For example, let's take a project that starts on 01.10.2009. The *member network* for this project is created by using the links between the members in the project network until the given point of time (Oct 1, 2009). We then compute the network parameters to identify the shape of this sub-graph.

Since we have project collaboration data from 14.01.2005, we monitor projects that start from 01.01.2007 in order to have sufficient data about the previous collaboration of teachers for the analysis. Furthermore, we only consider projects that have more than 2 members since some network parameters (e.g. clustering coefficient and betweenness) only make sense in the networks with more than 2 nodes. Overall, we analyzed 7612 projects which have 26479 members in total.

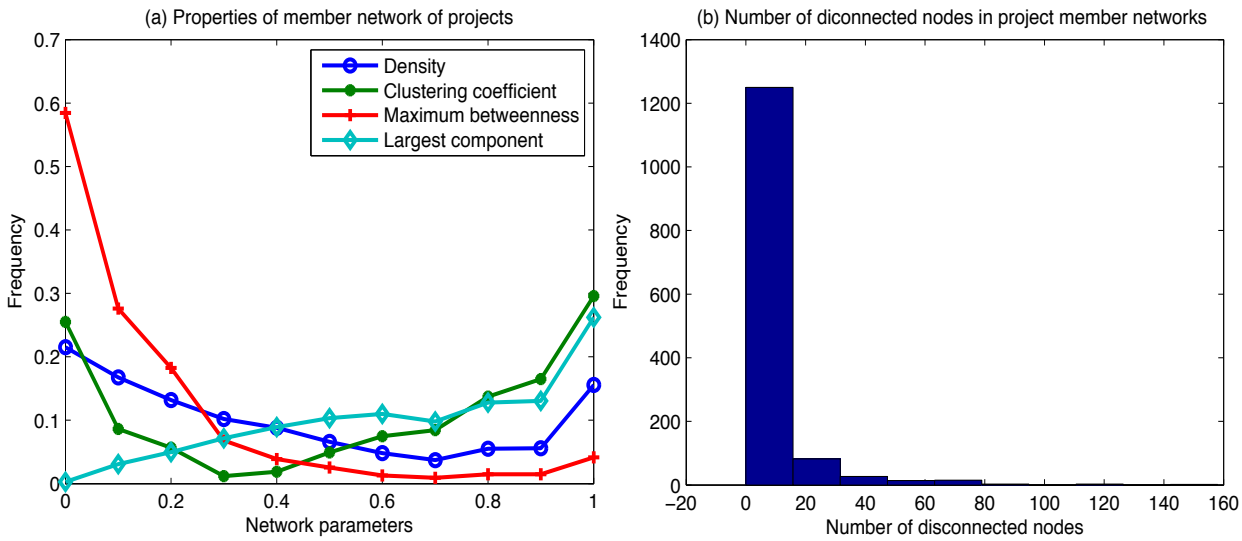


Figure 8: Structure of project member network in eTwinning

Figure 8(a) depicts the normalized frequency of each network parameter, including density, clustering coefficient, maximum betweenness and the largest connected component. Figure 8(b), on the other hand, plots the distribution of disconnected teachers of projects (i.e. the teachers who have not collaborated with any teachers in the new projects). It shows that most of the projects have small number of disconnected members (from 1 to 20), i.e. members that do not have connections to any other members because they are involved in the project for the first time. These are eTwinners who for the first time participate in the project collaboration and have therefore not had possibility to create other connections within the project network

yet. Therefore they are excluded from the network parameters calculation. It is good to mention that since this number is small, it does not affect the conclusions we draw from the observation based on network parameters.

From Figure 8(a) we can see the three following options: 1) that most of the *project member* networks are either in a *focused* stage (as described in the model in Figure 2) with a dense and large connected component, low betweenness and high clustering coefficient, or 2) in a *bonding* stage without community structure (the small components are in star topology) or 3) with a community structure (small components are densely connected), described by low/high clustering coefficient, low betweenness, low/high density and small connected components.

Connecting this observation with the notions of social capital, we can conclude that projects in eTwinning expose either *closure* or *structural hole* properties. The *focused* stage in Figure 2 presents a dense and well connected community structure, while *bonding* stage describes unconnected communities. Projects in the first form can benefit from the secure and trust relationship between members. On the other hand, projects in the second form can have many new ideas from different groups when they are brought together. In the next sections, we will verify these forms of social capital based on the achievement of teachers and projects.

5.2. Local network structure and teacher's performance

Given the existence of two forms of social capital in eTwinning, how can we identify the network structures that characterize them? Furthermore, which form are eTwinners (teachers) following? (Question 4)

Here we take the aforementioned "Quality Label" as an indicator for teachers' performance and reputation. With this indicator, we are able to find the correlation between the performance and the teachers' positions in the eTwinning network. We compute network properties as well as teachers' activities, such as number of project diary posts and comments, contacts, project, and journal posts and comments as functions of the number of Quality Labels, as depicted in Figure 9.

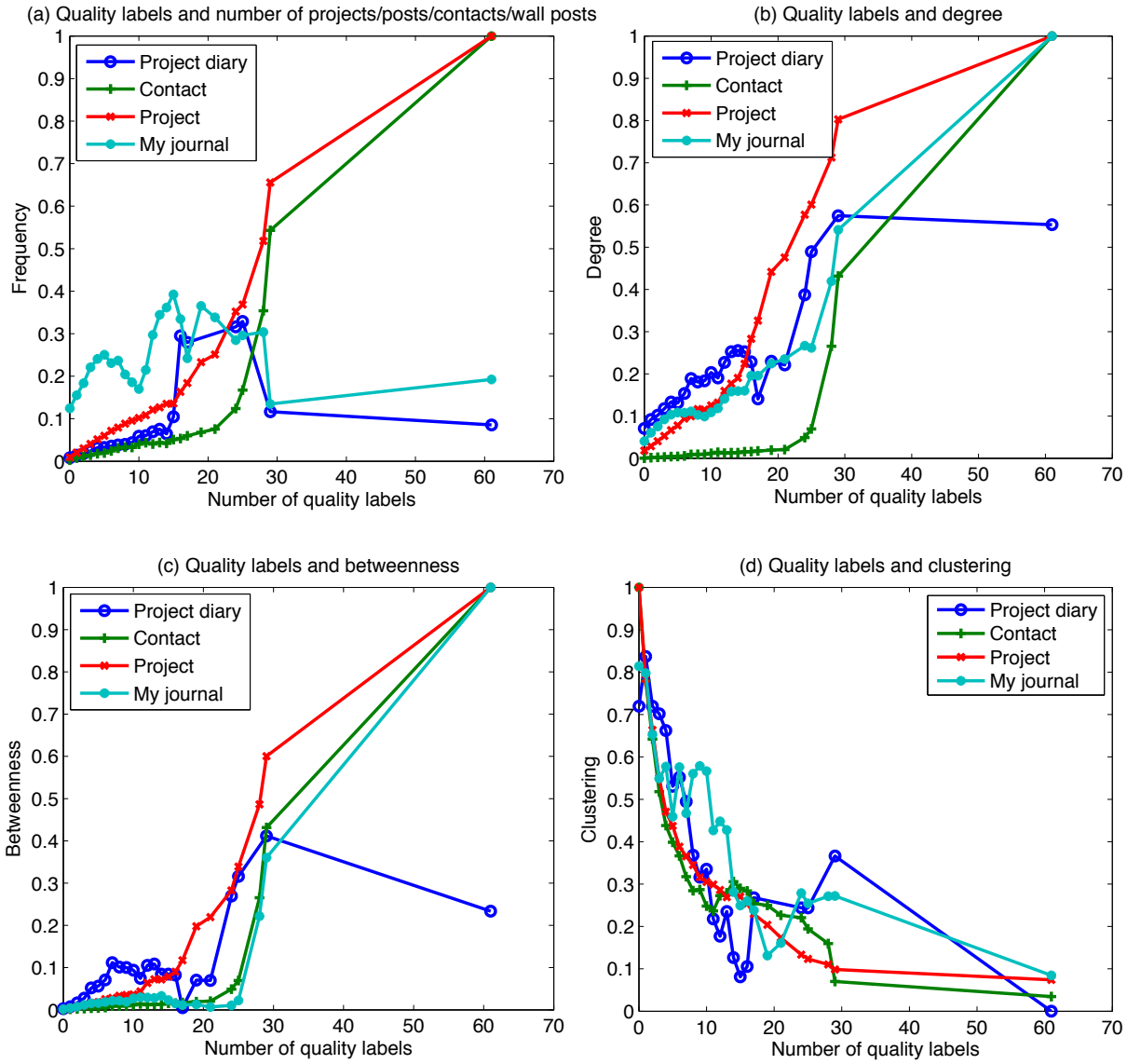


Figure 9: Performance of teachers as functions of the number of project diary posts and comments, contacts, project, and journal posts and comments, degree, betweenness and local clustering coefficient in eTwinning networks

Figure 9 shows that the properties of local network structure of a node clearly state that being positioned at the interface between communities has a big advantage. In detail, the nodes (eTwinners) with a high number of Quality Labels have very high betweenness and low local clustering coefficient, which indicates that they are structural holes connecting different communities together. Similarly, the nodes (eTwinners) with a low number of Quality Labels have low betweenness and very high local clustering coefficient, which indicates they are closure being clustered and located within the communities. Additionally, the pure number of project diary posts and comments, contacts, project, and journal posts and comments, and the degree of a node also correlate with the Quality Label.

This observation suggests that structural hole is more effective than closure in performing project related tasks in eTwinning and receiving Quality Labels. Connecting this result to the discussion of two forms of social capital in Section 2, we see that teachers who are structural

holes benefit from the diversity of information and knowledge sources from different communities. This diversity gives them more options in finding projects, partners and novel ideas. The structural hole also has more control power of information flow between communities in eTwinning.

5.3. Success factors for projects

One of primary goals of the eTwinning platform is to facilitate the communication and collaboration of teachers in Europe. It is a virtual meeting point for the exchange of information between schools, and it provides all the tools and services necessary for schools to find partners, resources, advice, help, and information, and even to build eTwinning projects in partnership with each other. Teachers from different schools and countries carry out projects in eTwinning. Teachers use ICT to exchange information, ideas and find partners for projects. A project can use different languages as well as be on one or multiple subjects.

The success of a project is recognized by various means, including the "Quality Label" as mentioned earlier. There are several factors that can lead to the award of a Quality Label, including the use of ICT, the internationality as well as the number and variety of institutions involved. One of the important factors that lead to the success of the project is the use of ICT for communication in carrying out the projects. We are interested in the early phase of a project, i.e. when project partners set up and prepare for the project. Communication between (potential) partners of a project happen before the project actually starts (e.g. partnership invitation, exchange of the ideas for the project, making contacts, and setting up communication channels). Hence, we investigate whether this early communication has an effect on the achievement of the project. It is important to note, though, that within the Tellnet data model, not all the possible communication tools are tracked (e.g. data from the internal messaging system is missing, as well as communication using tools such as email).

We begin with some subjective factors, including the number of teachers, institutions countries involved, languages used and subjects of a project. Since these factors are properties of a particular project and are decided by potential partners before the project starts, they somehow reflect the early communication between members of the project. Figure 10 depicts the fraction of projects that have received Quality Labels as a function of the number of countries involve, languages used, subjects (10a), and teachers and institutions (10b). Figure 10a clearly shows that the more countries involved, the better the project is. The number of subjects has almost no effect on the quality of the projects. Other factors follow this trend to a certain degree (10 languages, 30 teachers and 30 institutions).

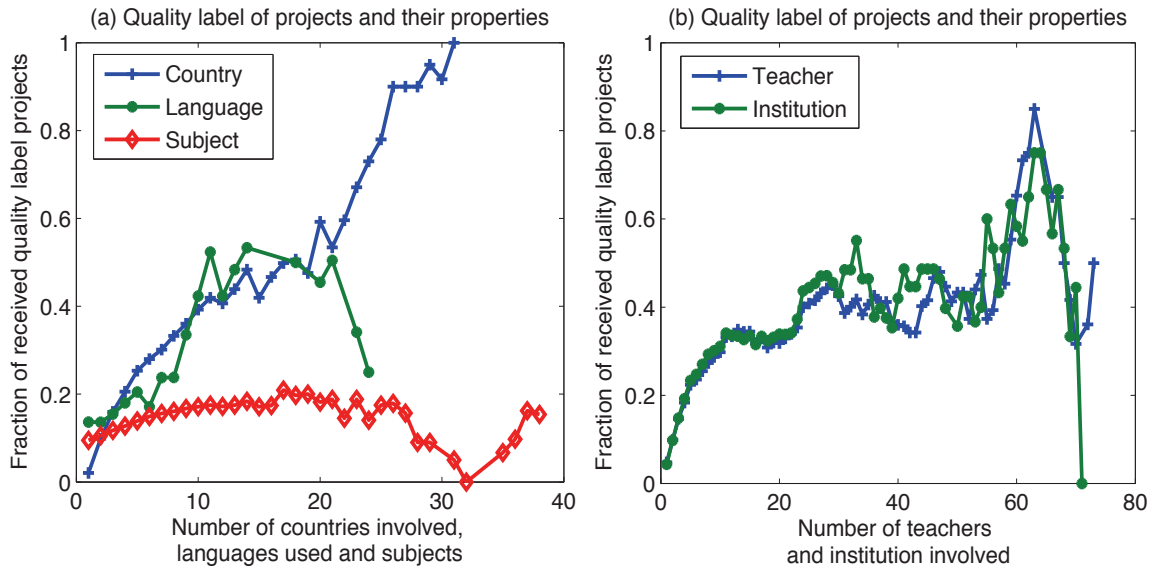


Figure 10: Project achievement and properties

Now given the early communication and collaboration between partners of a project, what is the structure of the sub-network that they form and how does this structure correlate to the quality of the project? (Question 5)

As mentioned earlier, within the Tellnet data model, not all the communication channels of teachers are recorded. We have data about the use of journal post/comment, project diary and project collaboration with time information. However, we do not have data about other possible communication tools such as internal messaging and email, which might be used extensively by teachers for communication in the early phase of a project. Therefore, analytical results based on this incomplete data might lead to imprecise conclusions. Instead, we investigate the strategy to find project partner, as we already mentioned in Section 5.1. In particular, we are interested in whether teachers look for new project partners who they have already collaborated with in previous projects and the correlation between this strategy and the achievement of the projects. This analysis partly answers Question 5: how teachers find the partners to invite and exchange ideas about a new project.

We processed data as described in Section 5.1. As presented in Figure 8 (Section 5.1), the member sub-networks of projects in eTwinning expose either structural hole or closure properties. Concretely, they are in the *bonding* stage (with or without community structure) or *focused* stage of the model in Figure 2. To identify the social capital form that projects in eTwinning follow, we find the correlation between the Quality Label and the structural properties of the project *member network*.

Figure 11 plots the fraction of the projects that have received Quality Labels and the (normalized) network parameters of their membership. It shows that high quality projects (those that received Quality Label) have member network in *bonding* stage with community structure (shown by high fraction of the projects that received Quality Label and have member network characterized by medium clustering coefficient, small connected components and low betweenness in Figure 11). Those projects that have member networks in the *focused* stage or *hierarchical* stage are less effective (shown by low fraction of the projects that received Quality Label and have member network characterized by low/high

clustering coefficient, large connected components and high betweenness in Figure 11). That means high quality projects favor the structural hole and the diversity in the membership is the key success factor. This behavior is in line with the goals of eTwinning platform and projects: to facilitate the communication and collaboration of teachers with ICT mechanisms. The membership in bonding stage has two advantages. On the one hand, members in densely connected groups have close and trust relations with other members of the group. On the other hand, when the project is carried out, it brings these groups together and combines their ideas, resources and information. Memberships in *focused* stage do not gain much benefit and new knowledge when they do a project together, since they have known or collaborated with each other before.

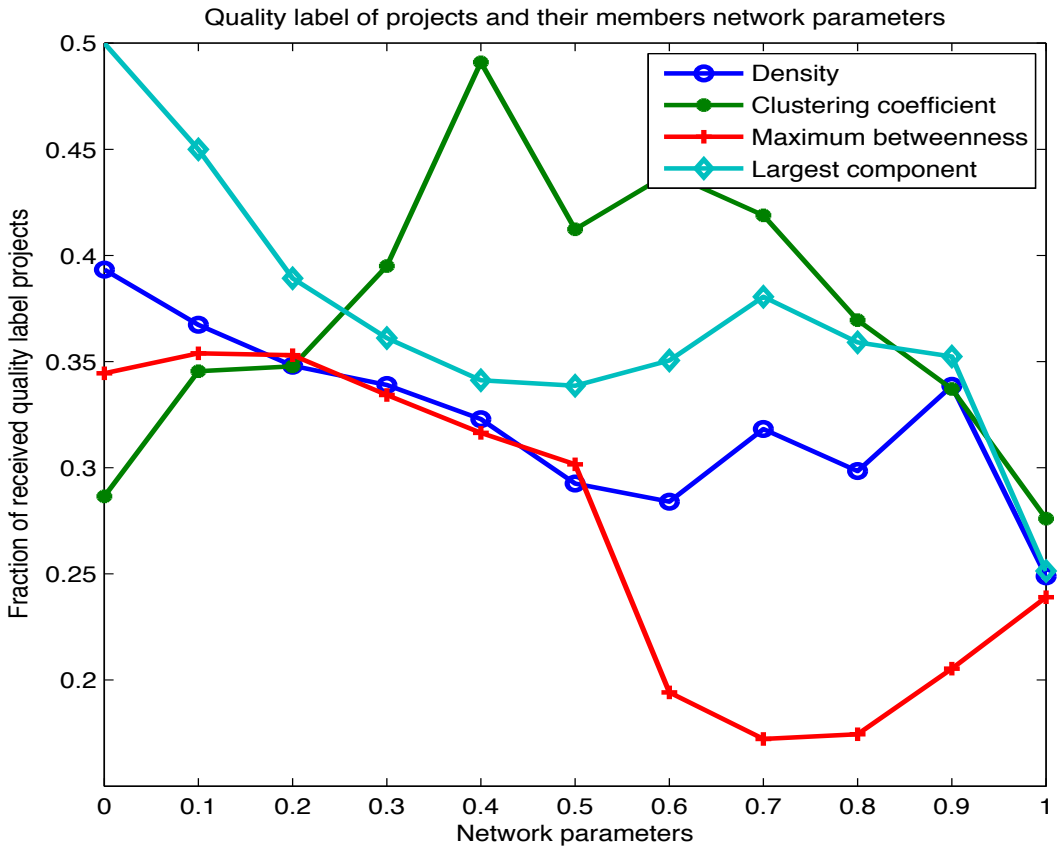


Figure 11: Project achievement and member network structure

5.4. Summary

In this section, we study social capital as a property of individuals (i.e. teachers) and as a property of groups (i.e. projects) in eTwinning. Four networks in our study (the project, contact, project diary, and journal networks) are complex networks (scale-free networks) with a power-law degree distribution, indicating the existence of super connectors (hubs) in the networks. A further analysis of local structure of the hubs shows the co-existence of two forms of social capital in eTwinning: the structural hole and closure. Connecting the local structure of teachers and their achievement (manifested by Quality Label), we find that the structural hole is more effective in terms of informational advantage and diversity of choices in selecting partners.

This form of social capital, structural holes, is also proved for the projects. Projects that connect several groups together in their membership achieve higher quality in terms of Quality Labels, again due to the diversity of the membership as well as the trust relationship in each group.

6. Tracking development pattern of teachers

The last issue we investigate in this deliverable is the development pattern of teachers in term of network structure. In Section 5, we have discussed about the current stage of teachers in the networks as well as the forms of social capital that they tend to follow. Here we address the following question: how do teachers develop to this stage? (Question 6) Specifically, what is the strategy in making contacts with other teachers and participating in projects? What can teachers learn from other teachers?

The answers for these questions can have an important application for teachers. Currently within the eTwinning network, teachers have no information about the whole network as well as their own positions there. They even have little information about their local virtual network within eTwinning and the location of other teachers. Therefore, it could be helpful to provide teachers with this information to align themselves in the right community and to learn from successful teachers who could be considered as role models. Tools are needed to give hints/suggestions to teachers based on this information. A question teachers may ask themselves: given my current stage, what can I do to improve or to maintain my stage and therefore also enhance my professional development?

A possible application is a recommendation tool that helps teachers choose projects to participate in or suggest the contacts to make. The recommendation tool should be based on the form of social capital in eTwinning that has been proved in Section 5, i.e. the structural hole, and the network parameters (i.e. betweenness and clustering coefficient) to make recommendations.

In this section, we discuss about the identification of the role models in eTwinning networks, i.e. what is the development pattern of "star teachers" recognized by the Quality Label? "Star teacher" as such does not exist in eTwinning, and the term is only used in this context to demonstrate the case.

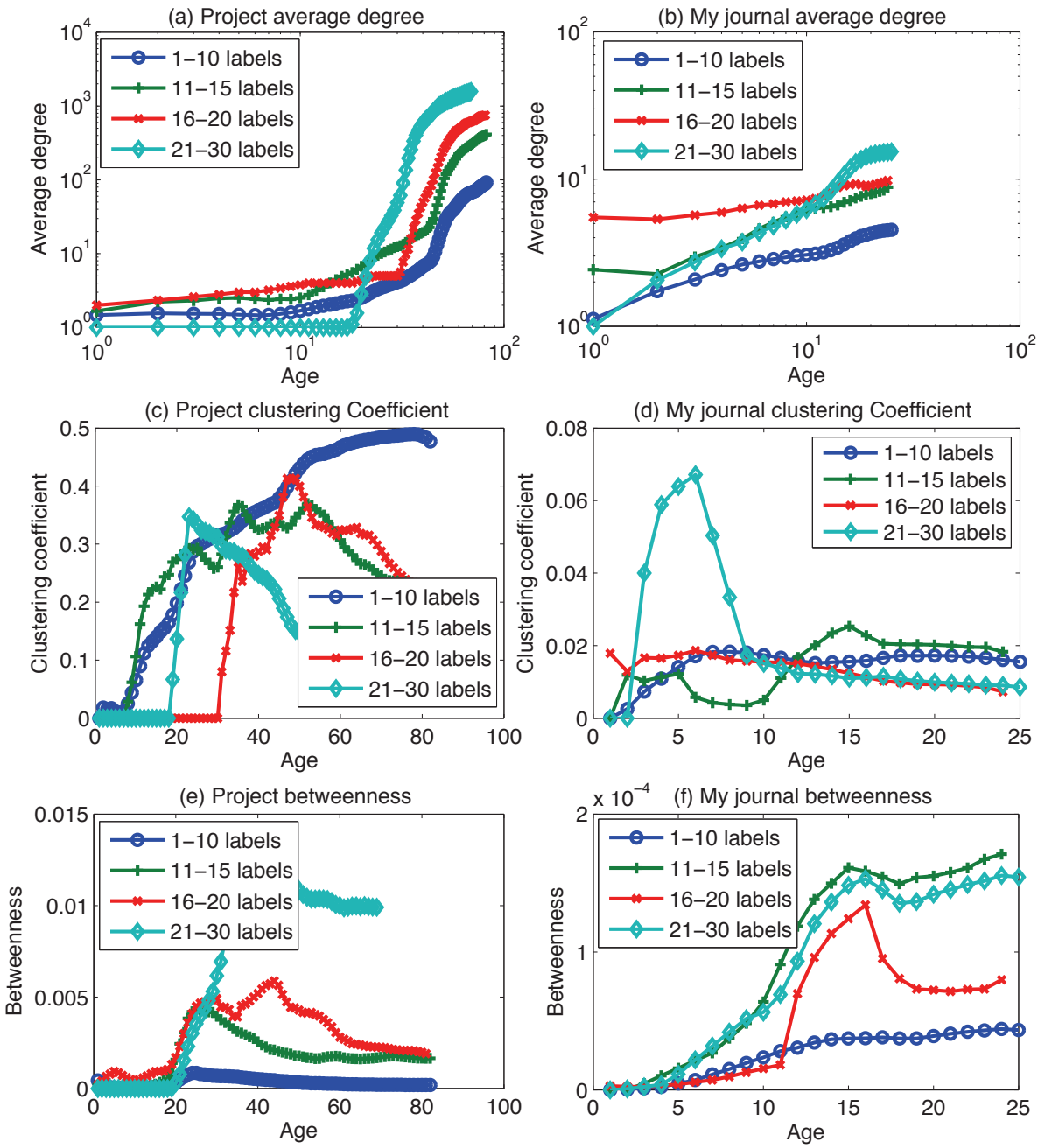


Figure 12: Local network structure over time and Quality Label

Figure 12 depicts the local network parameters (degree, local clustering coefficient and betweenness) of **four classes** of teachers who have received Quality Labels in project and my journal networks over time.

- First class contains teachers who have received 1 to 10 Quality Labels.
- The second class has received 11 to 15 Quality Labels,
- The third classes has received 16 to 20 Quality Labels,
- The fourth class has received 21 to 30 Quality Labels.

Since the contact network has no time information, we are not able to analyze it. Project diary network is mainly about the exchange of information between project members and project reports which are already in the project network. Therefore, we do not analyze it.

From Figure 12, we can clearly see the development trend of "star teachers". The basic strategy is to join first in and build up a strong and densely connected community. It is shown by the early increasing of local clustering coefficient of teachers (Figure 12c at month 10, 11, 34 and 21, and Figure 12d at month 4, 3, 4, and 4 for the first, second, third and fourth classes, respectively). Then "star teachers" start to collaborate or communicate across the communities. This is shown by the increasing of betweenness (Figure 12e at month 20 and Figure 12f at month 5 for all four classes) in both networks.

A closer look at Figure 12 reveals also the difference in the development of teachers in these four above-described classes. The more Quality Labels teachers have, the earlier their clustering coefficient shrinks and the higher their betweenness is before it shrinks. That means teachers who have less Quality Labels stay longer with their community than teachers who have more Quality Labels. There is also difference in the values of these parameters. Teachers who have less Quality Labels seem to stay closer with their local community, while teachers who have more Quality Labels have a good balance of communication with many communities. This is shown by the descending order of betweenness values and ascending order of clustering coefficient values over time of the first, second, third and fourth classes. These observations suggest that "star teachers" can quickly align themselves in their local community and then quickly connect to other communities.

Further exploration and visualization of these findings need to be investigated, which we have planned to the next deliverable. Tools are also needed to help teachers monitor their development in term of networking. Dynamic visualization techniques to visualize teachers' local networks and their network parameters would be helpful. Furthermore, tools to help teachers learn from other teachers are also needed. For example, teachers can compare their development pattern with other teachers' pattern, especially with star teachers, within or outside their community. Finally, as mentioned earlier, a recommender will be very useful for teachers to define their own strategy in finding projects and making contacts.

7. Conclusions and future work

Our research results reveal two significant aspects of social network analysis for professional development networks in eTwinning. First, that the degree distributions follow the power law, which indicates the existence of complex network and its underlying community structures in the eTwinning professional development network. In this regard, the project cooperation network is more interesting, because it has a better connectivity than project diary, contact, and my journal networks. The existence of complex networks assures the effective employment of social network analysis methods to assess roles and positions of individual nodes.

The position of teacher in the network can be an indicator or a predictor for their performance in eTwinning projects (a manifest of which is the Quality Label). Lying on the border of different communities has an advantage: teachers in this position have more control and power over the network as well as more sources for new information. This form of social

capital is called structural holes, and is approved through calculation of node and network properties. Those teachers who are more active in project cooperation, blogging, and other activities in eTwinning Portal have more Quality Labels than the other teachers who are not.

Second, a dynamic analysis of teachers' development path, especially those called "star teachers", reveals an important pattern. The basic strategy is to engage and build a strong community first and then start connecting to other communities. This result implies two important applications. On the one hand, teachers need a tool to monitor their current stage and to learn from other teachers. On the other hand, a recommender would help teachers in building their local community and define their strategy by suggesting projects and contacts.

Finally, besides the subjective factors such as languages, subject, country and institution, the quality of a project depends much on the previous collaboration between the members of the projects. SNA methods applied on the member network of projects give us an insight into the strategy to form the membership. The most significant result is that the diversity in the membership of a project has great impact on its quality. Projects that are able to bring and combine members from different communities have a big advantage regarding to the diversity of information sources and ideas, and the trust relationship in each community. This result again proves the structural hole – one form of social capital in eTwinning.

Reference

Barrat, A., Barthelemy, M., Vespignani, A. (2008). Dynamic Processes on Complex Networks. Cambridge University Press, 2008

Borgatti, S. (1998). Network measures of social capital. *Wire*, *21*(2), 27-36.

Breuer, R., Klamma, R, Cao, Y. & Vuorikari, R. (2009). Social Network Analysis of 45,000 Schools: A Case Study of Technology Enhanced Learning in Europe. Cress, U., Dimitrova, V. and Specht, M. (Eds.) Learning in the Synergy of Multiple Disciplines, Proceedings of 4th European Conference on Technology Enhanced Learning, EC-TEL 2009, Nice, France, Sept./Oct. 2009, LNCS 5794, Springer, pp. 166-180.

Burt, R. S. (2001). Structural Holes versus Network Closure as Social Capital. In N. Lin, K. Cook and R. S. Burt: *Social Capital: Theory and Research.* Sociology and Economics: Controversy and Integration series. New York: Aldine de Gruyter, 2001, pp. 31-56

Carley, K. M. (2003). "Dynamic Network Analysis" in Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers, Ronald Breiger, Kathleen Carley, and Philippa Pattison, (Eds.). Committee on Human Factors, National Research Council, National Research Council, pp. 133–145, Washington, DC.

Cheetham, G. & Chivers, G. (2005). Professions, Competence and Informal Learning. Northampton: Edward Elgar Publishing.

Coleman, James S. (1988). Social Capital in the Creation of Human Capital. American Journal of Sociology Vol. 94, pp. 95-125, the University of Chicago Press.

Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440-442. Nature Publishing Group.

Fetter, S., Rajagopal, K., Berlanga, A.J. & Sloep, P.B. (2011). Ad hoc transient groups: Instruments for Awareness in Learning Networks. In Workshop on Awareness and Reflection in Learning Networks. European Conference on Technology Enhanced Learning, (ECTEL, 2011), Palermo, Italy, September, 2011 (in press).

Gladwell, M. (2000). The tipping point: how little things can make a big difference, Little, Brown.

Hoadley, M. C. (2005). The shape of the elephant: scope and membership of the CSCL community. In Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years! (CSCL '05). International Society of the Learning Sciences 205-210.

Johnson, L., Smith, R., Willis, H., Levine, A. & Haywood, K. (2011). *The 2011 Horizon Report*. Austin, Texas: The New Media Consortium.

Travers, J. & Milgram, S. (1969). An Experimental Study of the Small World Problem. *Sociometry*, *32*(4), 425-443.

Kienle, A. & Wessner, M. (2005). Our way to Taipei: an analysis of the first ten years of the CSCL community. In Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years! (CSCL '05). International Society of the Learning Sciences 262-271.

Kienle, A. & Wessner, M. (2006). Analysing and cultivating scientific communities of practice. Int J Web Based Communities 2:377–393.

Leskovec, J. (2008). Dynamics of Large Networks. Ph.D. Dissertation. Carnegie Mellon Univ., Pittsburgh, PA, USA. Advisor(s) Christos Faloutsos. AAI3340652.

Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. Physics, 69(2), 1-5. APS. Retrieved from http://arxiv.org/abs/cond-mat/0309508

Newman, M., Barabasi, A.-L. & Watts, D. J. (2006). The structure and dynamics of networks. Princeton University Press, Princeton, NJ, USA.

Pham, M. C. & Klamma, R. (2010). The structure of the computer science knowledge network. Proceedings of the 2010 IEEE International Conference on Advances in Social Network Analysis and Mining (ASONAM 2010), Aug. 9-11, 2010, Odense, Denmark.

Pham, M., Klamma, R. & Jarke, M. (2011). Development of computer science disciplines: a social network analysis approach. Social Network Analysis and Mining, Springer Wien, ISSN: 1869-5450
Siemens, G. (2011). EDUCAUSE: Learning Analytics. Nov. 3, 2011, http://www.learninganalytics.net/?p=146.

Rogers, E. M. (2003). Diffusion of innovation. Free Press.

Song, E., Petrushyna, Z., Cao, Y. & Klamma, R. (2011). Learning Analytics at Large: The Lifelong Learning Network of 160,000 European Teachers, Towards Ubiquitous Learning - Proceedings of 6th European Conference on Technology Enhanced Learning, EC-TEL 2011, Palermo, Italy, September 20-13, 2011, LNCS Vol. 6964, pp. 398-411, Springer.

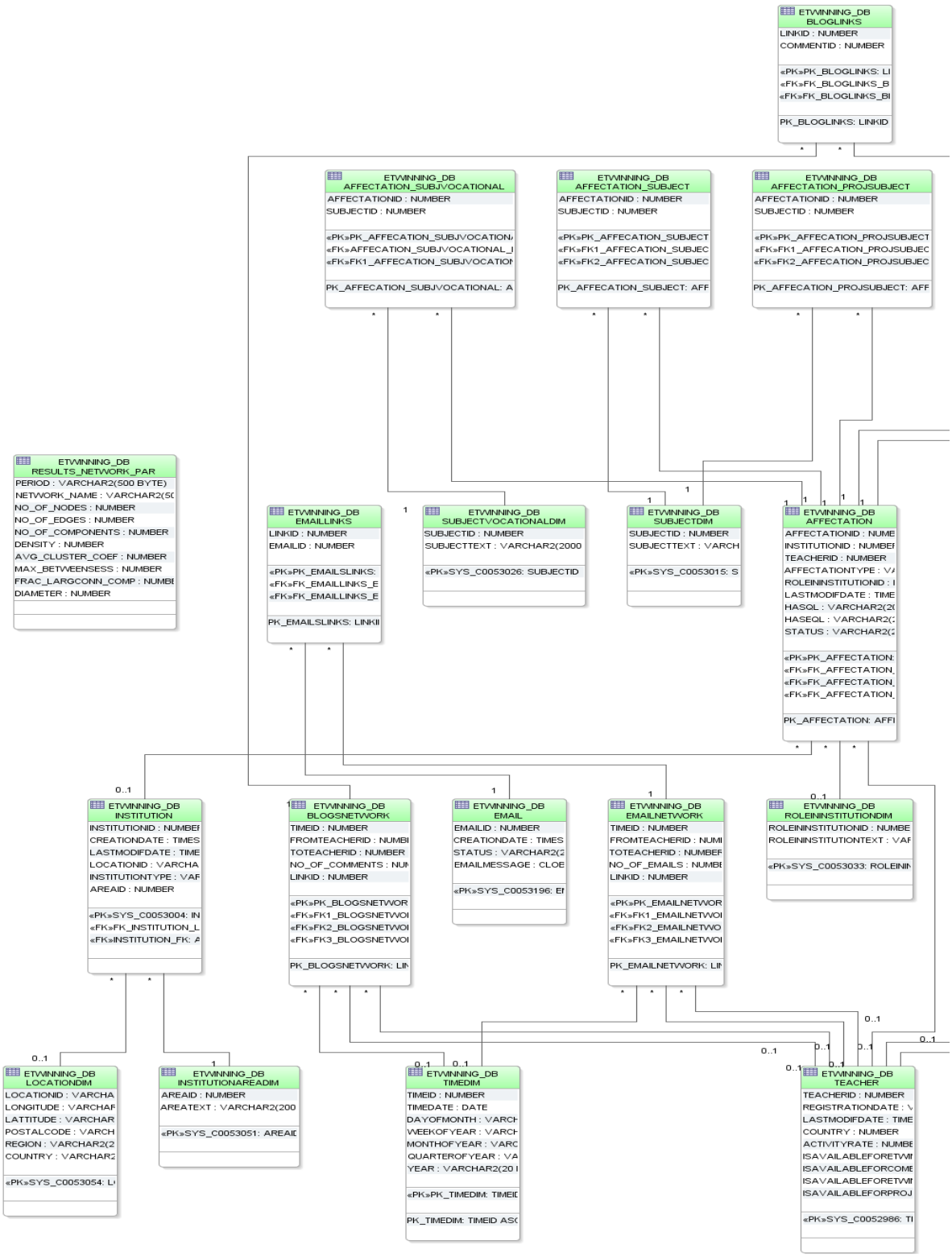
Vuorikari, R. (2010). Teachers' professional development - An overview of current practice. Central Support Service for eTwinning (CSS), December, 2010.

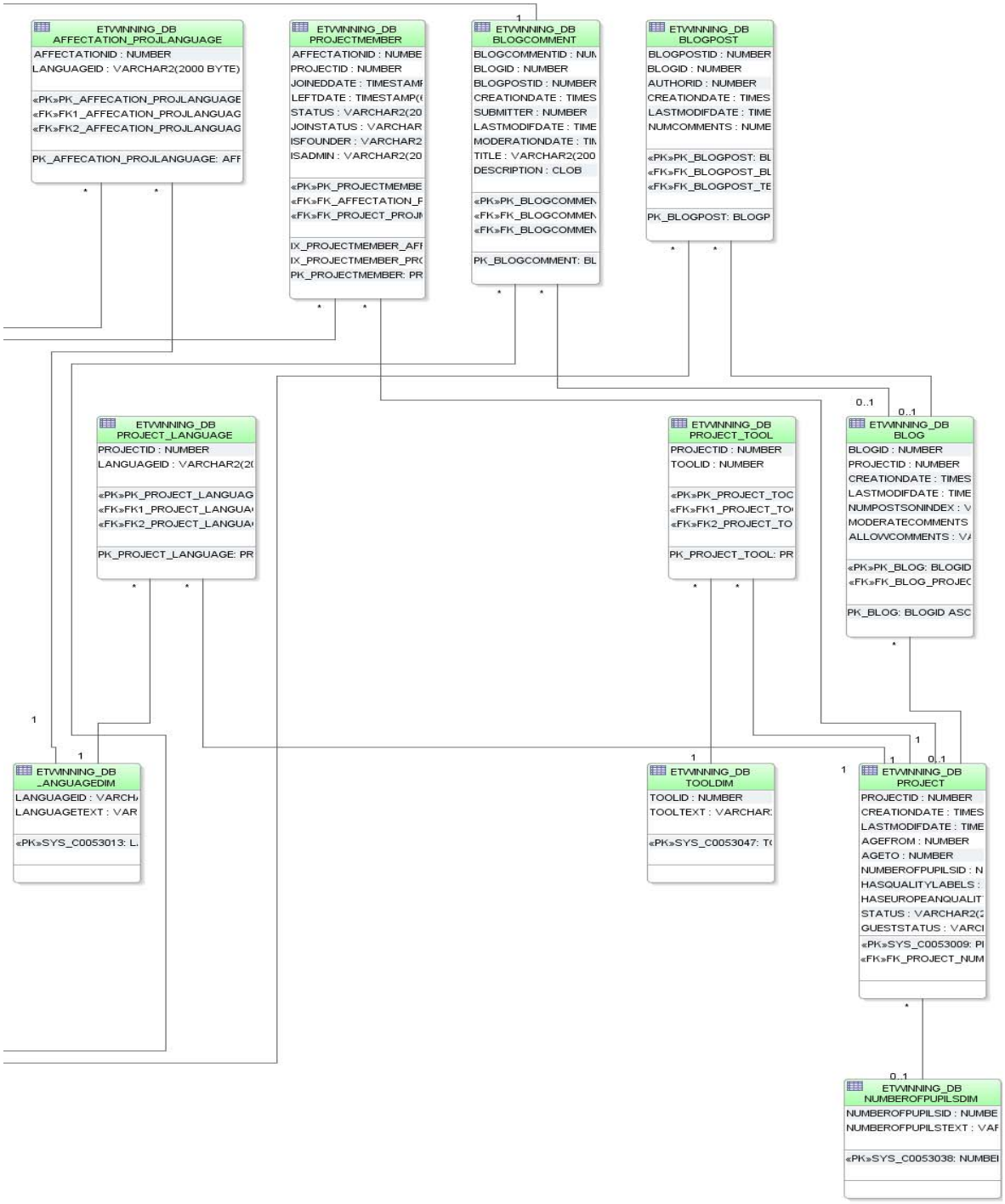
Vuorikari, R. et al. (2011). ICT-based School Collaboration, Teachers' Networks and their Opportunities for Teachers' Professional Development - a Case Study on eTwinning. Advanced Web-based Learning - Proceedings of 10th International Conference on Web-based Learning, ICWL 2011, Hong Kong, December 8-9, 2011, LNCS Vol. 7848, pp.111-120, Springer.

Wasserman, S. & Faust, K. (1994). Social Network Analysis: Methods and Applications. Cambridge University Press, New York, NY, USA.

Annex I: Data model

Data modeling and populating phase of the project includes building the database tables according to the star schema model and populating the tables with data provided in flat files. The solution consists of two schemas: ETWINNING_DATA and ETWINNING. The ETWINNING_DATA schema contains tables that have the same names and fields of the provided data files and their sole purpose is to hold the raw data for further processing. The ETWINNING contains tables that are designed according to the star schema model with integrity constrains. This schema serves the purpose of modeling the data in a way suitable for querying and analysis. It also finds and stores the links according to the requirements of social network analysis.


ETWINNING_DB BLOGLINKS
LINKID : NUMBER
COMMENTID : NUMBER
PK_BLOGLINKS: LINKID
ETWINNING_DB AFFECTATION_SUBJVOCATIONAL
AFFECTATIONID : NUMBER
SUBJECTID : NUMBER
PK_AFFECATION_SUBJVOCATIONAL: A
ETWINNING_DB AFFECTATION_SUBJECT
AFFECTATIONID : NUMBER
SUBJECTID : NUMBER
PK_AFFECATION_SUBJECT: AFF
ETWINNING_DB AFFECTATION_PROJSUBJECT
AFFECTATIONID : NUMBER
SUBJECTID : NUMBER
PK_AFFECATION_PROJSUBJECT: AFF
ETWINNING_DB RESULTS_NETWORK_PAR
PERIOD : VARCHAR2(500 BYTE)
NO_OF_NODES : NUMBER
NO_OF_EDGES : NUMBER
NO_OF_COMPONENTS : NUMBER
DENSITY : NUMBER
AVG_CLUSTER_COEF : NUMBER
MAX_BETWEENSESS : NUMBER
DIAMETER : NUMBER
ETWINNING_DB EMAILLINKS
LINKID : NUMBER
EMAILID : NUMBER
PK_EMAILSLINKS: LINKII
ETWINNING_DB SUBJECTVOCATIONALDIM
SUBJECTID : NUMBER
SUBJECTTEXT : VARCHAR2(2000
ETWINNING_DB SUBJECTDIM
SUBJECTID : NUMBER
SUBJECTTEXT : VARCH
ETWINNING_DB AFFECTATION
AFFECTATIONID : NUME
INSTITUTIONID : NUMBE
TEACHERID : NUMBER
LASTMODIFDATE : TIME
HASQL : VARCHAR2(2
HASEQL : VARCHAR2(
STATUS : VARCHAR2(
PK_AFFECTATION: AFFI
ETWINNING_DB INSTITUTION
INSTITUTIONID : NUMBE
CREATIONDATE : TIMES
LASTMODIFDATE : TIME
LOCATIONID : VARCHA
INSTITUTIONTYPE : VAF
AREAID : NUMBER
ETWINNING_DB BLOGSNETWORK
TIMEID : NUMBER
FROMTEACHERID : NUMBI
TOTEACHERID : NUMBER
NO_OF_COMMENTS : NUM
LINKID : NUMBER
PK_BLOGSNETWORK: LIN
ETWINNING_DB EMAIL
EMAILID : NUMBER
CREATIONDATE : TIMES
STATUS : VARCHAR2(2
EMAILMESSAGE : CLOE
ETWINNING_DB EMAILNETWORK
TIMEID : NUMBER
FROMTEACHERID : NUMI
TOTEACHERID : NUMBEF
NO_OF_EMAILS : NUMBE
LINKID : NUMBER
PK_EMAILNETWORK: LIN
ETWINNING_DB ROLEININSTITUTIONDIM
ROLEININSTITUTIONID : NUMBE
ROLEININSTITUTIONTEXT : VAF
ETWINNING_DB LOCATIONDIM
LOCATIONID : VARCHA
LONGITUDE : VARCHAF
LATTITUDE : VARCHAR
POSTALCODE : VARCH
REGION : VARCHAR2(2
COUNTRY : VARCHAR2
ETWINNING_DB INSTITUTIONAREADIM
AREAID : NUMBER
AREATEXT : VARCHAR2(200
ETWINNING_DB TIMEDIM
TIMEID : NUMBER
TIMEDATE : DATE
DAYOFMONTH : VARCH
WEEKOFYEAR : VARCH
MONTHOFYEAR : VARC
QUARTEROFYEAR : VA
YEAR : VARCHAR2(20
PK_TIMEDIM: TIMEID ASC
ETWINNING_DB TEACHER
TEACHERID : NUMBER
REGISTRATIONDATE : V
LASTMODIFDATE : TIME
COUNTRY : NUMBER
ACTIVITYRATE : NUMBE
ISAVAILABLEFORETWI
ISAVAILABLEFORCOME
ISAVAILABLEFORETWI
ISAVAILABLEFORPROJ


ETWINNING_DB
AFFECTATION_PROJLANGUAGE
AFFECTATIONID : NUMBER
LANGUAGEID : VARCHAR2(2000 BYTE)
«PK»PK_AFFECATION_PROJLANGUAGE
«FK»FK1_AFFECATION_PROJLANGUAG
«FK»FK2_AFFECATION_PROJLANGUAG
PK_AFFECATION_PROJLANGUAGE: AFF
ETWINNING_DB
PROJECTMEMBER
AFFECTATIONID : NUMBE
PROJECTID : NUMBER
JOINEDDATE : TIMESTAMI
LEFTDATE : TIMESTAMP(
STATUS : VARCHAR2(20
JOINSTATUS : VARCHAR
ISFOUNDER : VARCHAR2
ISADMIN : VARCHAR2(20
«PK»PK_PROJECTMEMBE
«FK»FK_AFFECTATION_F
«FK»FK_PROJECT_PROJI
IX_PROJECTMEMBER_AFI
IX_PROJECTMEMBER_PR(
PK_PROJECTMEMBER: PR
ETWINNING_DB
BLOGCOMMENT
BLOGCOMMENTID : NUI
BLOGID : NUMBER
BLOGPOSTID : NUMBER
CREATIONDATE : TIMES
SUBMITTER : NUMBER
LASTMODIFDATE : TIME
MODERATIONDATE : TI
TITLE : VARCHAR2(200
DESCRIPTION : CLOB
«PK»PK_BLOGCOMMEN
«FK»FK_BLOGCOMMEN
«FK»FK_BLOGCOMMEN
PK_BLOGCOMMENT: BL
ETWINNING_DB
BLOGPOST
BLOGPOSTID : NUMBER
BLOGID : NUMBER
AUTHORID : NUMBER
CREATIONDATE : TIMES
LASTMODIFDATE : TIME
NUMCOMMENTS : NUME
«PK»PK_BLOGPOST: BL
«FK»FK_BLOGPOST_BL
«FK»FK_BLOGPOST_TE
PK_BLOGPOST: BLOGP
ETWINNING_DB
PROJECT_LANGUAGE
PROJECTID : NUMBER
LANGUAGEID : VARCHAR2(2(
«PK»PK_PROJECT_LANGUAG
«FK»FK1_PROJECT_LANGUA
«FK»FK2_PROJECT_LANGUA
PK_PROJECT_LANGUAGE: PR
ETWINNING_DB
PROJECT_TOOL
PROJECTID : NUMBER
TOOLID : NUMBER
«PK»PK_PROJECT_TOC
«FK»FK1_PROJECT_TO
«FK»FK2_PROJECT_TO
PK_PROJECT_TOOL: PR
ETWINNING_DB
BLOG
BLOGID : NUMBER
PROJECTID : NUMBER
CREATIONDATE : TIMES
LASTMODIFDATE : TIME
NUMPOSTSONINDEX : V
MODERATECOMMENTS
ALLOWCOMMENTS : VA
«PK»PK_BLOG: BLOGID
«FK»FK_BLOG_PROJEC
PK_BLOG: BLOGID ASC
ETWINNING_DB
_ANGUAGEDIM
LANGUAGEID : VARCH
LANGUAGETEXT : VAR
«PK»SYS_C0053013: L
ETWINNING_DB
TOOLDIM
TOOLID : NUMBER
TOOLTEXT : VARCHAR
«PK»SYS_C0053047: T
ETWINNING_DB
PROJECT
PROJECTID : NUMBER
CREATIONDATE : TIMES
LASTMODIFDATE : TIME
AGEFROM : NUMBER
AGETO : NUMBER
NUMBEROFPUPILSID : N
HASQUALITYLABELS :
HASEUROPEANQUALIT
STATUS : VARCHAR2(
GUESTSTATUS : VARCI
«PK»SYS_C0053009: PI
«FK»FK_PROJECT_NUM
ETWINNING_DB
NUMBEROFPUPILSDIM
NUMBEROFPUPILSID : NUMBE
NUMBEROFPUPILSTEXT : VAF
«PK»SYS_C0053038: NUMBEI