Association for Information Systems
AIS Electronic Library (AISeL)

ECIS 2008 Proceedings

European Conference on Information Systems (ECIS)

2008

Application of Social Network Analysis in Knowledge Processes

Claudia Mueller
University of Potsdam, cmueller@wi.uni-potsdam.de

Norbert Gronau
University of Potsdam, ngronau@wi.uni-potsdam.de

Robert Lembcke
University of Potsdam, rlembcke@wi.uni-potsdam.de

APPLICATION OF SOCIAL NETWORK ANALYSIS IN KNOWLEDGE PROCESSES

Mueller, Claudia, University of Potsdam, August-Bebel-Str. 89, 14482 Potsdam, Germany, cmueller@wi.uni-potsdam.de

Gronau, Norbert, University of Potsdam, August-Bebel-Str. 89, 14482 Potsdam, Germany, ngronau@wi.uni-potsdam.de

Lembcke, Robert, University of Potsdam, August-Bebel-Str. 89, 14482 Potsdam, Germany, rlembcke@wi.uni-potsdam.de

Abstract

Social network analysis (SNA) is increasingly applied to analyze existing knowledge exchange processes in companies. In the past, these processes were mainly analyzed by traditional business process modeling methods. Same of these methods are working with a specific knowledge perspective, e.g. KMDL (Knowledge Modeling and Description Language). The knowledge needs and demands are modeled along knowledge-intensive business processes. The SNA however has no direct reference to processes; here persons with specific relationships (e.g. knowledge exchange) are mainly the object of interest. Therefore a combination of both views – the SNA and knowledge-intensive business process modeling – is valuable. In the following contribution two approaches are introduced to combine social network analysis with the modeling of knowledge intensive business processes based on KMDL. In addition to a theoretical introduction to SNA and KMDL there are two examples introduced, which illustrate the practical implementation of these approaches. In this work in progress traditional techniques like interviews are combined with electronically gathered data. Also by using different perspectives a better understanding of all events in knowledge processes can be extracted.

Keywords: Social network analysis, knowledge processes, modeling method, Wiki.

1 INTRODUCTION

Today, formerly centralized hierarchical organizations are increasingly transforming to decentralized and network-based constructs, which promote more openness. In these systems, which are subject to continuous change processes, people and available information are key resources (Tapscott & Williams 2007), (Barabási 2003). Flat hierarchies and flexible structures, team orientation as well as a progressively growing dependency on knowledge are fundamental characteristics (Cross & Parker 2004). These changes in organizational structures have also impacted corporate knowledge management. New technologies are also increasingly used in knowledge management, e.g. social software. This enables new possibilities to gather data electronically. Social network analysis (SNA) is applied to analyze existing knowledge processes which are mainly based on communication in virtual environments such as Wikis. For instance, the results can be used to evaluate the effectiveness of tool applications in terms of knowledge processes (Mueller & Meuthrath 2007).

Furthermore, in order to cope with these changes new modeling techniques are necessary to analyze the existing knowledge and information flows in companies. In the past, existing modeling methods integrated the knowledge perspective in their approaches, e.g. KMDL and EPC (event-driven process chains). But more and more processes in companies possess a volatile processes character. Process models which are focused on instance modeling are mostly obsolete after a short time. Process models on a type stage (e.g. reference processes) are too imprecise to comprehend all specific characteristics after a while. Therefore new methods have to be applied to extend existing methods of knowledge process modeling to improve analysis results. A new approach is introduced by extending an existing modeling method with a social network description. Therefore SNA metrics can be used to analyze knowledge and information flows in companies or in specific communities.

The following research question is answered: Is there an improved understanding of knowledge sharing processes in knowledge management systems based on a combination of process modeling methods and network analysis?

This contribution is structured as follows: the second section gives a short theoretical introduction in SNA based on the graph theory, its scopes of analysis with basic measures and a multi-perspective approach to differentiated network types. The third section introduces briefly the Knowledge Modeling and Description Language (KMDL) and its new communication view, which enables an integration of the SNA with knowledge process management. Finally, this new approach is applied to two real-life examples, which shows how SNA can enhance existing process descriptions.

2 APPLYING NETWORK ANALYSIS IN KNOWLEDGE SHARING PROCESSES

Social network analysis (SNA) is an enabler to systematically describe and investigate network structures and their principles of order in organizations. Basically, a social network is “a specific set of linkages among a defined set of persons, with the additional property that the characteristics of these linkages as a whole may be used to interpret the social behavior of these persons involved” (Mitchell 1969, p.2). The social environment is described by patterns or regularities among interacting persons. Social relationships of persons constitute the social capital of an organization. Social capital is the sum of actual or virtual resources, which an individual or a group possesses. That means the social capital metaphor is that people who do better are somehow better connected. Consequently, if there is an understanding of existing social networks in a company, the processes of knowledge sharing can also be improved. SNA is increasingly applied for analyzing and reorganizing distributed communication and collaboration structures in knowledge management (e.g. (Cross & Parker 2004)). The required data for network analysis can be gathered electronically or via interviews. One approach utilizes user interaction with knowledge management systems (e.g. E-mail, Wikis, Blogs) as applied in this contribution.

A short introduction in basics of graph theory is given before an introduction in the multi-perspective approach and scopes of network analysis is given. An isomorphic mapping of a network as graph enables an analysis of its topological structure and existing interdependencies. A graph G(N,L) consists of a finite number of g-nodes $N=\{n_1, n_2, ..., n_g\}$ and L-edges $L=\{l_1, l_2, ..., l_L\}$. Two vertices in a graph are adjacent, if the edge $l_k=(n_i, n_j)$ is in the set of the edges L. In a graph there are g vertices and L edges (Wasserman & Faust 1997). There are undirected and directed graphs. Whereas undirected graphs possess no direction edges (arcs) of a directed graph cross in one direction only, e.g. hyperlinks in an HTML document. Each arc is an ordered pair of distinct vertices. In weighted graphs edges carry additional information like strength or intensity, e.g. frequency of interaction, rating of friendship. A graph can be represented as adjacency matrix or visualized as network. In an adjacency matrix there is a row and a column for each vertex. Entries of a matrix indicate if two vertices are adjacent. A graph G(N,L) can be presented in a diagram or a matrix. In diagrams points depict vertices and a line between two points is an edge, if there is an edge between these vertices in the set L. The simplest form of such a diagram is a sociogram (Moreno 1954). The location of the points and the length of the lines depend on the used algorithm.

2.1 Multi-perspective approach

A specification of networks is possible in various contexts. Knowledge management, operations research and social networks techniques are combined together to define a meta-matrix. In this meta-matrix multiple representation of the entities and the connections among them are specified (Carley 2002). Three (person, information, event/task) of the defined four stages (person, information, event/task, organization) are introduced which are summarized in six perspectives. Each perspective covers specific network definitions. Depending on the field under investigation, this meta-matrix can be applied to derive specific networks. Relationships in one network usually imply relationships in another (Carley 2003). Therefore, to understand all knowledge processes in virtual information spaces, an analysis of these different network types is necessary. In this contribution, the social perspective is used to illustrate the specific approach.

Social networks comprise of interaction networks, like communication, discussion, and collaboration networks. Answering the question “who interacts with whom?” and structural measurements are central. In social networks vertices are persons with their corresponding relationships. Two examples are introduced: Wiki-based collaboration networks and e-mail-based communication networks.

Basically, a Wiki-based collaboration network is used to investigate the nature and extend of collaboration between authors of Wiki articles and it enables the analysis of the information exchange between communities (Mueller & Meuthrath 2007). Collaboration is based on mutually referred asynchronous generated contributions in a knowledge process. In a collaboration network vertices are authors, whereas the edges are constructed on common edits of an article. Whenever two or multiple authors worked on the same article in a Wiki, they are connected to each other. The central assumption is that changes of different authors on an article imply collaboration. A strong tie exists, if two authors have contributed comparatively more than other authors (in terms of content) or they work collaboratively very often on one or more articles (in terms of frequency).

An e-mail based communication network is used to investigate interactions in organizations based on e-mail traffic. An e-mail network is a directed network. In an e-mail network, the authors of e-mails are represented vertices. Every primary e-mail between two persons is represented as an edge. Is there more then one e-mail per direction, the number of e-mails is described with the help of the edge weight. So every new e-mail per direction increments the edge weight between the sender and the receiver of the e-mail. The assumption is, that persons with a high edge weight in both direction between each other work often and effective and highly collaborative together. Another assumption is that persons with an extreme high outgoing edge weight might be sharing their knowledge and become the star in the e-mail network.

The knowledge perspective comprises of knowledge networks. Cultural aspects are mainly investigated and the central question is “who knows what?”. The knowledge networks are so-called 2-mode networks means that in these networks two different types of vertices exist. One vertices type is persons and the other is an information object. For instance Wiki specific realizations are competence networks. This type of network shows existing thematic competences of authors. The edge between two vertices (article, author) represents the editing activity of a certain author for a given article.

The need perspective is dealing with persons and their specific task or events they are part of. The network is a social network but the relation based on a specific point in a time, e.g. flight, meeting, project. The central question is “who needs what”. Mostly, it is a sub-network of the already known communication or collaboration network. It is also possible to define it as a bimodal network, with persons and specific events or tasks as vertices. In terms of tasks processes are visible and in terms of events the greatest common subset can be identified.

In the information perspective the question „which information are available?“ are considered. Information objects form a network. Applying the Wiki information space again, a Wiki-link network can be defined. A Wiki-link network consists of articles as vertices and Wiki-links as edges. It is a directed network since Wiki-links are not reciprocal.

The attendance perspective should answer the question “What information is used in which task/event?”. Whereas the information perspective relates information by semantic or syntactic linkages the attendance perspective relates information by specific task or events. As in the need perspectives it is possible to define sub-networks or bimodal networks.

In the temporal perspective, the temporal development of knowledge topics and their dependencies of specific events or tasks are central. The question “what happened after what?” should be answered. In the Wiki context, it may be interesting to gather information about how the Wiki is used not only by its contributors but also by its readers. The visiting-flow network sequences Wiki pages by its temporal order by means of the referrer web server log.

2.2 Scopes of network analysis

A network is a complex system. Based on changing external influences, these complex systems have a certain structural variability. This structural variability comes along with different phases of a network. Based on specific parameters transitions from one phase to another are possible based on the internal dynamic of the network. There are different structures in different phases. Each phase is related to new macroscopic characteristics of a network. However, it is very important to differentiate between random fluctuations and phase transitions which strongly influence the users’ behavior. Therefore, an analysis of different orientations of networks is necessary. Four different analysis methods are available: attribute, position, structure, and dynamics analysis. Depending on the objective of analysis the specific method is applied. Figure 1 shows this classification.

A typical network analysis focuses on analyzing relational data (Scott 2000, p. 2f.). Nevertheless, a network offers in addition to relational data also attribute data. In an attribute analysis, the network data are combined and evaluated in relation to these conventional data types (Jansen 2003, p.51). The attribute analysis as analysis of characteristics enables capturing the properties of vertices and edges and to relate them to the structure of the network. These data are investigated by applying the exploratory statistics (Tukey 1977). There are only minor information about the relationships of these data known. One method is data mining where specific pattern are extracted from data (Fayyad 1996, p. 39). A combination of relational and attribute data takes place in network data mining. Conventional explicit data types from data mining are augmented by relational data from network analysis to implicit relations between data sets (Galloway 2006).

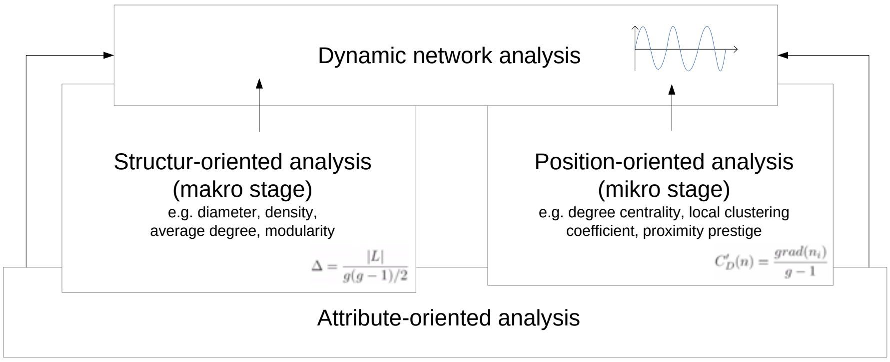

Figure 1. Scopes of network analysis with selected metrics (Mueller& Meuthrath & Baumgraß 2008)

A position analysis investigates the micro stage of a network. The analysis aims to investigate the single nodes of a network, their characteristics and their position in the network. One simple measure is the degree of a node. The degree of a vertex $d(n_i)$ is the number of vertices that are incident with it and reflects the activity of a person in a social network (Wasserman & Faust 1997, p.100). In the basic centrality concept three approaches can be differentiated: the degree centrality, the closeness centrality and the betweenness centrality. The activity of a person in a social network can be investigated with the metric degree centrality. This metric helps to identify the “most important” actor, meaning those that are extensively involved in relationships with other actors in the social network. A vertex is central in the sense of degree centrality (C_D), if a vertex has many relations to adjacent vertices. The closeness centrality (C_C) focuses how close one person to all other persons in the network is and is a function of the geodesics distance. The direct and indirect relations are considered. The betweenness centrality (C_B) based on the idea that one person is then important if this person lies often on a path between two persons. In this case this specific person can “control” interactions between two nonadjacent persons. In directed graph, the measure prestige can be applied. Concepts of prestige operationalize the prominence of a network position. For instance it can be measured, how often messages are transferred over a specific node in a communication network (PD) or how often one article is referred to in other articles in an information network. The asymmetry of relation is important that means, if a node is a source or drain of a relation (Jansen 2003, p. 142ff.).

With the structural analysis the macroscopic characteristics of a network can be studied. Measures to investigate the structure of the network are for instance density, average distance, and average degree. These metrics are used here to describe the network structure on a macroscopic level. Density of a graph is the ratio of relations over number of maximal possible relations (Wasserman & Faust 1997). The highest possible number of edges is reached, if each vertex of the graph is connected with each other vertex in the graph. The (geodesic) distance is defined as the length of a geodesic between two vertices. The average distance (average shortest path between all vertices) is calculated for the whole network. This metric is equal to the degree of separation which connects all pairs of nodes in a group. Short distances transmit information accurately and in a timely way. A long distance degree implies slower even distorted information transmission (Cross & Parker 2004). The degree of a node is the number of edges that are connected to this node (adjacent nodes). This measure helps to evaluate the “attractiveness” of a node. The average degree is calculated for all nodes in the network.

All measures of the introduced analysis can be utilized to investigate dynamic characteristics of networks as well. The dynamic perspective of network analysis deals with the progression of state changes over time and enables the observation of events and its analysis. It is necessary to compare at least two different network states at different times to determine changes of any form in the network. Dynamic network analysis (DNA) enables analyzing network states, the vertices and their edges as well as changes in structure and configuration of the network (Carley 2003). The cumulative analysis enables the aggregation of a long time period of the network, including all changes and elements, which are or were part of the network during the period. The calculated measures and their changes are visualized in a diagram. Basically, the network dynamic and the network evolution can be differentiated (Stockman & Doreian 1997, p. 234). The network dynamic is a general concept in which the change in time is described. Opposite to network evolution, where the change processes itself which are the source of a changing network structure are being investigated.

2.3 Analysis social networks in virtual information spaces

An implementation of these concepts has taken place in SONIVIS:Project. This project aims at analyzing virtual information spaces. The primary objective is to investigate the different development stages of these stages and therefore to recommend depending on each stage appropriate knowledge management activities. This way existing knowledge management is enhanced by analyzing virtual knowledge and information flows between departments, projects, or persons in a company. Central questions are: What are the “hot” topics in my company?, Who collaborates with whom?, To what extend the information space is used? To conduct these analyses, a special tool has been developed. The SONIVIS:Tool provides statistical analysis of virtual information spaces. It is based on the eclipse Rich Client Platform (RCP). Predefined graphical analyses are used which offer a quick overview on recent developments. Additionally, various SNA metrics are provided on a microscopic and macroscopic level. At the moment, the analysis of Wiki-based and e-mail based networks is supported.

3 MODELLING AND ANALYSIS OF KNOWLEDGE SHARING PROCESSES

The following section introduces shortly the applied modeling method KMDL (Knowledge Modeling and Description Language), which is used to model knowledge intensive processes and its activities. A new extension of this modeling method is described with enables a connection of business process analysis with network analysis. Basically, two approaches that combine business process management with network analysis are compared and explained shortly.

3.1 Basic modeling method

The knowledge modeling and description language KMDL[1] is a semi formal modeling language to identify and evaluate knowledge and information flows in knowledge intensive processes. At the moment KMDL compromises two modeling views - the process view and the activity view. The process view describes a work flow from the view of the specific process steps. It consists of tasks and sub-process steps. In the process view, it becomes clear which task is handled in which order and shows possible alternatives for each task of the process. Also in this view resources which are necessary to fulfill the task are assigned to it. Different parts of processes can be connected via process interfaces. Therefore, even complex processes can be modeled with the help of the KMDL process view (Froeming & Gronau & Schmid 2006).

In figure 2 a very simple process is modeled. The central object in this view is a specific task “use case modeling”, which is specified in the activity view. The modeled tasks are surrounded by process interfaces. These interfaces are keeping the connections to the whole process. The task itself is connected with further information such as the IT systems which are necessary to fulfill this task and a role which executes this task.

The activity view forms the core of KMDL. It allows a detail description of every knowledge conversion (knowledge and information flows) which appears in specific knowledge intensive tasks. The activity view describes the tasks in detail.

---

[1] A detailed description and definition of KMDL and its objects is available at http://www.kmdl.de/.

The aactivity view is focused on information and knowledge flows. It decomposes every task into different activities. The differentiation between explicit and tacit knowledge by Polanyi (1958) is utilized: Tacit knowledge is personal, context specific and difficult to communicate. In contrast, it is possible to hand over explicit knowledge in a formal and systematic language (Nonaka & Takeuchi 1995). Along with this, used and generated information within the knowledge activities are educible. Information is a flow of messages, while knowledge is generally created by the very flow of information, through anchoring it in the beliefs and commitments of its holder. This understanding emphasizes that knowledge is essentially related to human action (Nonaka & Takeuchi 1995). In KMDL, information and explicit knowledge are represented as information objects. Tacit knowledge used within activities is represented by knowledge objects. Both types should be examined together, because their interactions create new or extend existing knowledge. This interdependency is referred by Nonaka und Takeuchi as knowledge conversion (Nonaka & Takeuchi 1995). Accordingly, the modelling of the used and generated information and knowledge objects and the knowledge conversions enriches the sequential description of the process view. Hence, there are different conversion types specified: socialization, externalization, combination, and internalization (Nonaka & Takeuchi 1995).

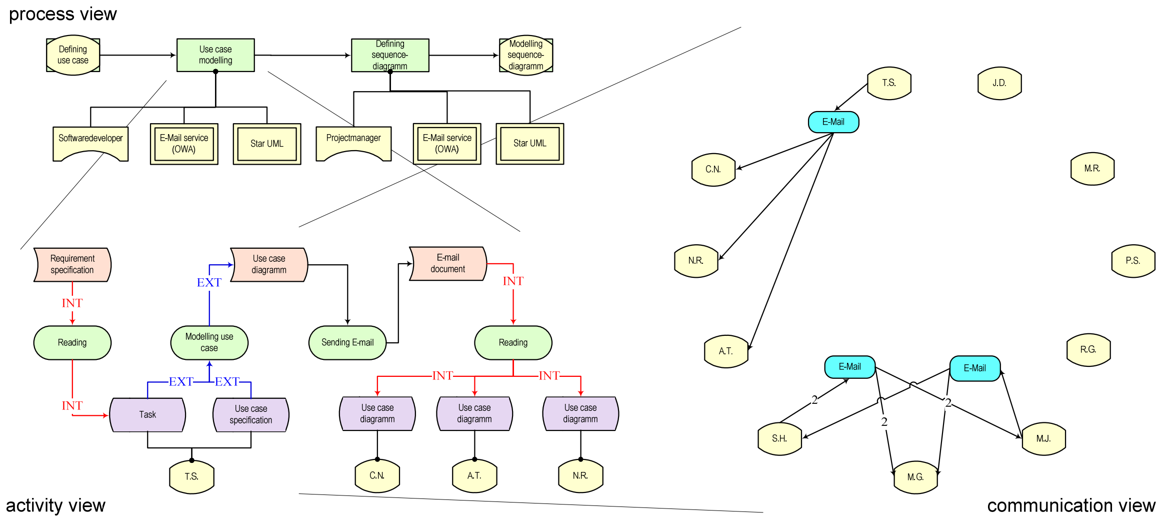

Figure 2. Example of the process view, activity view, and communication view in KMDL

A socialisation (SOZ) is a directed graph from one or more knowledge objects of one person to one or more knowledge objects of another person. The externalisation (EXT) is modelled by the connection of at least one knowledge object with an information object. The association of at least two information objects and the generation of a new information object are named as combination (COM). The internalisation (INT) is started by an information object and ends with a knowledge object. However, there are exists some relations, which can not be assigned to one of the mentioned relation types. These relations are named as undefined relations.

Figure 2 shows on the left side a simple example of an activity view with the basic objects: information object, activity, knowledge object and person. Reading an information object, "requirement specification" creates certain tacit knowledge about this information. If the person fulfills an activity like modeling a use case, s/he uses knowledge objects about the task and use case specification to fulfill the activity. If the activity contains the creation of a new information object such as in figure 2, the person externalizes (EXT) his or her knowledge into the information object.

3.2 Extension of modelling method

In the existing modeling approach, information and knowledge flows are only analyzed in relation to knowledge activities. One challenge was to identify main knowledge sources and drains based on selected knowledge activities. The KMDL communication view enables a new perspective on the process, because instead of knowledge activities people and their relationships are of the object of interest. The communication view describes communications in a knowledge-intensive task or a specific organization. The knowledge flow which is described in the KMDL activity view causes interaction between the involved persons. This interaction might be based on e-mails, in person conversations, telephone calls, instant messaging, and so on. For analyzing flows of knowledge it is necessary to know the information flow. The communication view of KMDL connects the view of the knowledge intensive processes with the view of networks. The modeled communication network in KMDL represents at least one part of the social network. The main idea is to model communication in relation to knowledge processes and not to model the knowledge process or social networks itself.

Defined objects of the communication view are based on the object definition in KMDL. There are some new modeling objects specified. People who are involved in the knowledge process are modeled as person-objects. The definition of persons-objects is based on the KMDL activity view. One object with high importance is “CommunicationWay” (CW). This object describes explicitly the medium used for communication (e.g. e-mail, telephone). The input relation of this object comes from a person but the output relation can reach either a person or a group of persons. An important fact is that the input relation and the output relation of the CW object have to be the same relationship type.

In figure 2 on the right side, basic objects of the communications view are shown. These objects are CWs and persons. The defined communication view is based on various activity views whereas in figure 2 only one of them is illustrated. The main interest of this view is the relationships between persons. On the one hand it is possible to differentiate relationships in formal and informal communications. And on the other hand, they are classified based on characteristics of communication. In KMDL communication is classified in four groups (Reichwald et al. 1998, p. 5). The first group consists of communication which happens at same place and time, the second group consists of communication which happens at the same time but at different places, the third group consists of communication which happens at the same place but not at the same time and the last group consists of all communication which happens at different time and places.

The communication view in figure 2 shows all people who are involved in the whole process. The communication which takes place in the activity view in figure 2 is modeled as a CW-object between senders and receivers of information. The relationships between senders and receivers via the CW-object are classified into the fourth group of communication. The frequency of communication is modeled as the edge weight of the relations.

3.3 Transformation of process model in network model

There are two approaches to integrate the process-oriented modeling method KMDL with the network analysis. In figure 3, these approaches are visualized. Based on interviews, the process view is captured and specific knowledge activities are selected. Then based on further interviews the KMDL activity model is specified. To get the communication view (CV) out of the activity view, every communication must be identified. Firstly, all participating persons are modeled in the CV. After that, the communication medium is modeled as a CW object. A CW object should be named with a short and clear description of the applied medium (e.g. “sending an e-mail”). Senders of information are the start of the relation between it and the CW object. The relationship between the CW object and the receiver starts at the CW object. The type of the relationship is based on the classification of the communication.

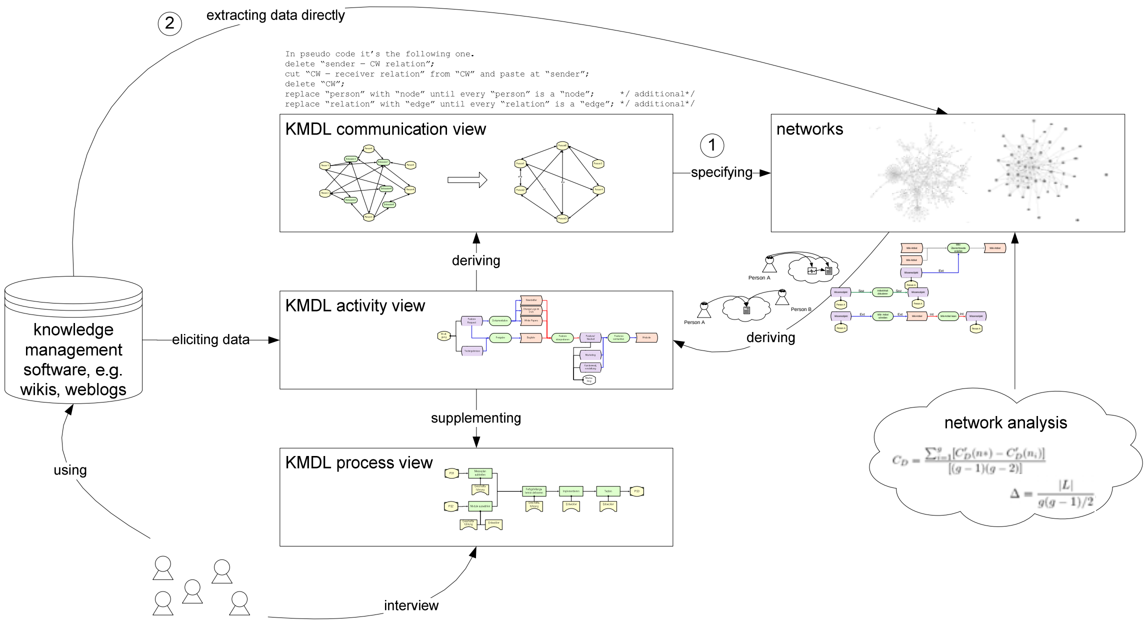

Figure 3. *Two approaches of integration knowledge process modeling with social network analysis*

It is easy to generate a network view out of the KMDL communication view. First, the relationship from the sender to the CW object is deleted. The start of the relationship from the CW object to the receiver of information is modeled. Then the CW object is deleted. Every additional object is also deleted. A network with persons and relationships is derived from a communication view. Every person object can be interpreted as node in a network. The person objects can be replaced by nodes and every relationship can be interpreted as an edge. Because of that relationships can be replaced as network edges.

4 PRACTICAL IMPLEMENTATIONS OF THE APPROACHES

In this section the theoretically defined approaches of combining business process management and network analysis are applied in a practical setting. Two case studies are used to clarify the two proposed integration approaches. The first one is located in a software company which uses a Wiki as an intranet solution. The second one is a university project in which students jointly developed one software solution but they were distributed in different teams.

4.1 Approach 1: Wiki-based knowledge sharing

Wikis are increasingly applied in knowledge management. There are various areas of applications such as a documentation tool for requirements specification, a collaborative writing tool, or intranet solution. The last example is used here to demonstrate the first approach to integrate knowledge process modeling with network analysis.

The data set used is based on a corporate Wiki that was launched in summer of 2005. It serves as a knowledge exchange platform for about 1,000 employees in a software platform. The objective is enhanced knowledge transfers between all business units. Previously existing solutions, e.g. Intranet or file systems were too static or only locally accessible. In April 2007 the corporate Wiki had 185 users and about 8,200 articles. The executive board of the company commissioned the analysis of the real knowledge activities in the Wiki and to evaluate the amount of knowledge activities usage.

Based on available data for this information space, a collaboration network was specified. The analysis of this network enabled the evaluation of existing common interactions in the information space. The number of relationships, that means the amount of collaboration increases in the investigated period but decreases in relation to the number of vertices (table 1).

		24.07.05	**31.12.05**	**23.07.06**	**30.12.06**	**24.07.07**
Collaboration Network	Vertices	8	49	92	176	261
	Edges	7	152	452	878	2

Table 1. Number of vertices and edges in collaboration network

There is no correlation between the growth of the author count and the increasing editing of different articles by different authors. That means in most cases articles continue to be changed by just one or two authors. This is caused by the current development stage of this Wiki, because mostly new articles are created. The Wiki is still in the expansion phase.

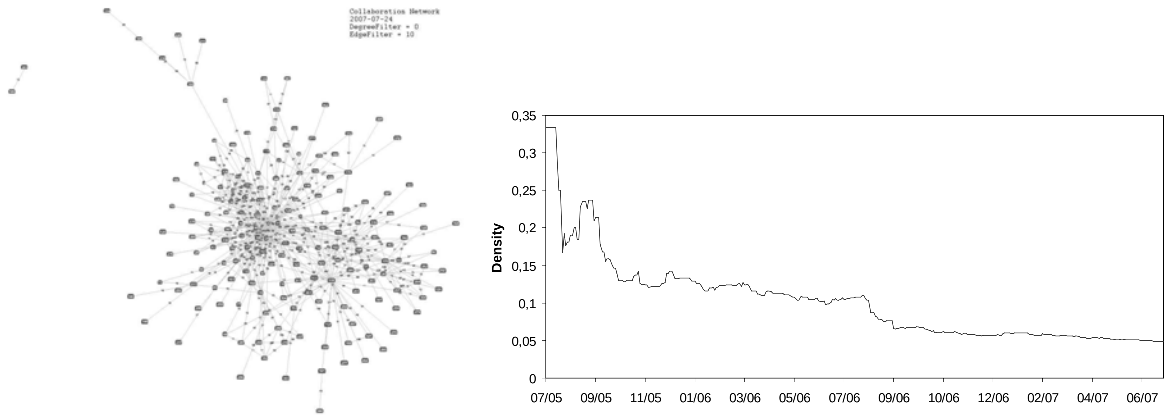

Figure 4. Visualization of collaboration network (06/07) and its density in time

From the Wiki database collected data can be used to define a KMDL activity model. In figure 5 some selected knowledge activities in a Wiki are visualized on the right side and they are transferred in the KMDL activity view syntax. Each Wiki activitiey can be identified by a specific Wiki metric. Retaining knowledge based on the metric article count (AC), changing knowledge on an edit count (EC) and exchanging knowledge on timely time wise closely related edits of two authors on the same article. The KMDL snippets are predefined model schemes. A knowledge object refers to a specific topic (mostly article name) and an information project is the physical equivalent of it. Therefore based on an attribute-oriented analysis (section 2.2) of the Wiki information space KMDL, models can be created automatically. With these models further investigations are possible, e.g. application of KMDL pattern approach (Bahrs & Bogen & Schmid 2005).

In this short example, one approach to derive networks from Wiki-based information spaces is illustrated. Using one network definition – collaboration network – the development of a collaboration network in terms of existing interaction was introduced. As a result, the overall development of this Wiki is positive but it is still in an expansion stage. It was also shown that knowledge activities can be identified based on collected data and that KMDL activities models can be specified.

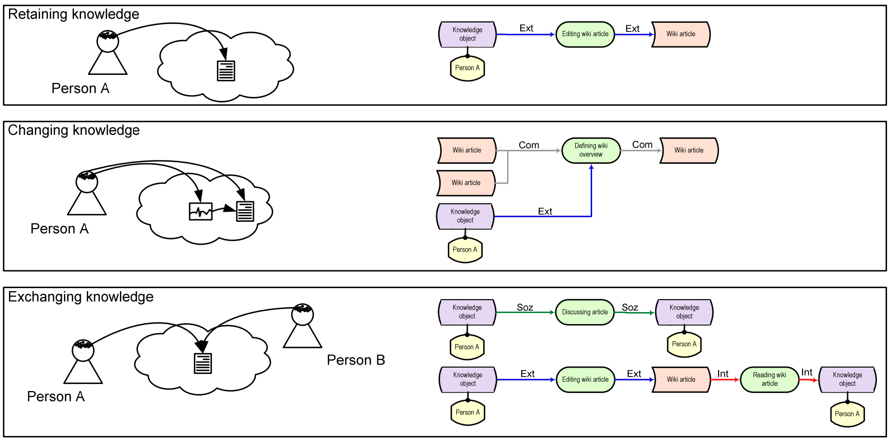

Figure 5. Transfer of Wiki-based knowledge activities in KMDL schemes

4.2 Approach 2: E-mail based communication in software development

One important necessity for knowledge intensive processes is communication. The development of software is a process. Communication processes can be used to identify networks with cliques and different network positions like star, gatekeeper and bridge (Schenk 1984, p. 250). Analyzing communication networks makes it possible to investigate informal communication, to find key connectors (=knowledge carriers), to evaluate existing competency, or to assess connectivity across groups. The question is how and in what way the work in a software development process is executed. Therefore the process will be transformed into a communication network as it was described in section 3.2.

A group of eleven IT System Engineering students got the task to design software for a car rental service. They had to model all relevant use cases, entity-relationship-diagrams, and UML class and sequence diagrams. They worked in three teams and once a week they had a status meeting with their academic advisor. Their primary communication medium was e-mail. Specifically their internal communication was e-mail based. The e-mail based communication refers to different work locations.

Person	Team	$d_{in}(n_i)$	$d_{out}(n_i)$	$C_D(n_i)$	$C_C(n_i)$	$C_B(n_i)$	$P_D(n_i)$
T.S.	1	5	13	0,109	0,275	0,25	0,042
C.N.	1	7	0	0	0	0	0,059
N.R.	1	7	0	0	0	0	0,059
A.T.	1	7	0	0	0	0	0,059
R.G.	2	7	2	0,017	0,119	0,083	0,059
P.S.	2	8	0	0	0	0	0,067
M.R.	2	6	0	0	0	0	0,05
J.D.	2	7	11	0,092	0,325	0,458	0,059
S.H.	3	25	20	0,168	0,179	0,125	0,21
M.G.	3	24	17	0,143	0,037	0	0,202
M.J.	3	16	56	0,47	0,064	0,083	0,134

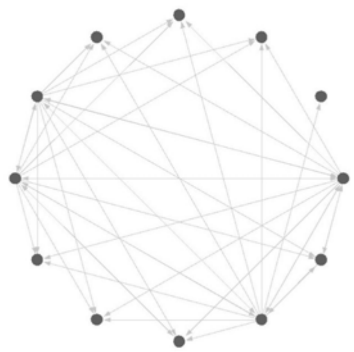

Table 2. Analysis results of e-mail communication (during one week) and visualization of the communication network

Table 2 shows, that the person M.J. got the highest out-degree. M.J. sent most of the emails. Maybe M.J. did most of the work, because this person's degree centrality value is higher than that of the others. This leads to the assumption that this person stayed in close contact with all persons.

S.H. and M.G. got both very high in-degree values. Additionally, their prestige value is much higher than the prestige value of all other persons. Because prestige values are based upon in-degree values, in this evaluation prestige shows the value of the information a person get from others. A high prestige value means that the person gets a lot of information. Values of members from team one and two are mostly balanced. Referring to the fact, that the data set was captured at the end of the study period, it seems apparent that team 3 had not finished its work at that time. That manifests itself in the high values in all evaluated network dimensions of all team 3 members.

These analysis results enable the evaluation of the activity state of each team. Based on it the work progress can be specified. Furthermore team members who are knowledge drains can be identified. Also the person (star of the network) who has the highest out-degree (= knowledge source) can also be identified. These results can only be derived because of the application of the communication network analysis and not by simply using the activity view.

5 SUMMARY AND OUTLOOK

To understand virtual information spaces a multi-perspective approach is necessary. In this approach the decomposition into different network types is proposed. Each network type is related to a specific view on the space. Each network can be analyzed applying different analysis scopes – attribute, position, structure, and dynamic analysis. Selected metrics are introduced as well as the KMDL modeling method, which focuses on modeling knowledge intensive processes. There are process and activity views in the existing description language. Both views have visualization on different aggregation levels. These views are extended by a new communication view which uses information of both views to describe communication processes in the investigated processes. Additionally, two approaches are explained to transform process models in network models or visa versa. Both approaches are investigated in more detail based on two case studies. These examples have shown that an improved understanding of knowledge sharing processes based on a combination of process modeling methods with network analysis is possible. The network view is orthogonal to existing process description principles. With the help of well-applied business modeling methods the network analysis can be introduced in a corporate context more easily. At the moment an integration of the already existing modeling tools is planned and further research is conducted to specify the benefits more accurately.

References

Barabasi, A.-L. (2003). Linked. Plume Printing, New York.

Bahrs, J., Bogen, J., Schmid, S. (2005). Pattern based Analysis and Redesign of knowledge intensive Business Processes, In: Jantke, K.P., Fähnrich K.-P., Wittig, W.S.: Marktplatz Internet: Von e-Learning bis e-Payment (in German), GI Lecture Notes in Informatics, P-72, p. 124-130, 2005.

Carley, K.M. (2002). Smart Agents and Organizations of the Future. In Lievrouw, L. and Livingstone, S.: The Handbook of New Media, p. 206-220, Sage, Thousand Oaks.

Carley, K.M. (2003). Dynamic Network Analysis. In Brelger, R., Carley, K.M. and Pattison, P.: Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers, p. 133-145, National Academy Press, Washington, D.C..

Cross, R.L. and Parker, A. (2004). The Hidden Power of Social Networks: Understanding how work really gets done in organizations. Harvard Business School Press, Cambridge.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. In American Association for Artificial Intelligence Magazine, Fall, p. 37-54.

Froeming, J. and Gronau, N. and Schmid, S. (2006). Improvement of software engineering by modeling of knowledge intensive business processes, IJKM, International Journal of Knowledge Management, Volume 2, 4/2006.

Galloway, J. and Simoff, S.J. (2006). Network data mining: methods and techniques for discovering deep linkage between attributes. In: Stumptner, M., Hartmann, S. and Kiyoki, Y.: APCCM, 53, CRPIT, p. 21-32, Australian Computer Society.

Jansen, D. (2003). Einführung in die Netzwerkanalyse (in German). 2nd Edition. Leske + Budrich, Opladen.

Mitchell, C. (1969). Social Networks in urban situations: Analysis of personal relationships in Central African towns. University Press, Manchester.

Moreno, J.L. (1954). Die Grundlagen der Soziometrie (in German). 1st Edition. Westdeutscher Verlag, Köln, Opladen.

Müller, C. and Meuthrath, B. (2007). Analyzing Wiki-based networks to improve knowledge processes in organizations. Proceedings of I-Know'07, p. 103-110.

Mueller , C., Meuthrath, B., Baumgraß, A.: Analyzing Wiki-based networks to improve knowledge processes in organizations. Accepted contribution in J.UCS Journal of Universal Computer Science, 2008.

Nonaka, I. and Takeuchi, H. (1995). The knowledge-creating company: how Japanese companies create the dynamics of innovation. Oxford University Press, New York, Oxford.

Polanyi, M. (1958). Personal Knowledge: Towards a Post-Critical Philosophy. The University of Chicago Press, Chicago.

Reichwald, R., Möslein, K., Sachenbacher, H., Englberger, H. (1998) Telekooperation. Verteilte Arbeits- und Organisationsformen (in German). Springer Verlag, Berlin Heidelberg New York.

Schenk, M. (1984). Soziale Netzwerke und Kommunikation (in German). J.C.B. Mohr, Tübingen.

Scott, J. (2000). Social Network Analysis. 2nd Edition. SAGE Publications, London.

Stockman, F.N. and Doreian, P. (1997). Evolution of Social Networks: Processes and Principles. In: Evolution of Social Networks, p. 233-250. Gordon & Breach, Amsterdam.

Tapscott, D. and Williams, A.D. (2007). Wikinomics. Portfolio, London.

Tukey, J.W. (1977). Exploratory Data Analysis. Addison-Wesley, Massachusetts.

Wasserman, S. and Faust, K. (1997). Social network analysis: methods and applications. 1st Edition. Cambridge Univ. Press, Cambridge.