Visual Navigation of Distributed Knowledge Structures in Groupware-based Organizational Memories

Stefan Smolnik – Ingo Erdmann
Department of Business Computing 2, University of Paderborn, Germany
{Stefan.Smolnik|Ingo.Erdmann}@notes.upb.de

Abstract

Many of today's organizations already have a strong integration of groupware systems in their IT-infrastructure. The shared databases of these groupware systems form organizational memories, which comprise the complete knowledge of an organization collected over the time of its existence. One key problem is how to find relevant knowledge in continuously growing and distributed organizational memories. The basic functionalities and mechanisms in groupware systems are not sufficient to support users in finding required knowledge. Topic maps provide strong paradigms and concepts for the semantic structuring of link networks and therefore, they are a considerable solution for organizing and navigating large and continuously growing organizational memories. The K-Discovery project suggests applying topic maps to groupware-based organizational memories to create knowledge structures and address the mentioned challenges. Visual navigation capabilities to exploit the created knowledge structures are based on hyperbolic geometry concepts and provide users with intuitive access mechanisms to the required knowledge.

1. Introduction and overview

Within many organizations, groupware-based environments are the basis of communication and information management. With an increasing integration of these environments in internal business processes, organizational memories grow fast. These shared databases, often existing in groupware systems, allow transformation of individual knowledge into common knowledge. Looking at the technical perspective of knowledge management, groupware seems to be a suitable platform for managing organizational memories (cf. [1]). The growing flood of available information requires powerful concepts and mechanisms to support users, who search for relevant knowledge objects. Mechanisms of navigation and linking as well as functionalities for extensive searches and investigations are needed to explore and use complex offers of knowledge. These functionalities are a necessary prerequisite for the core processes of knowledge identification and knowledge use [2]. Therefore, effective search and navigation mechanisms, which improve organizational use of existing individual and common knowledge objects, contribute to the process of knowledge generation, i.e. the development or collection of new knowledge. Particularly, visual navigation techniques add an additional efficiency benefit to this process of knowledge acquisition.

Topic maps – as defined in ISO/IEC 13250 – used on information sets create knowledge structures and form a structured semantic link network upon a large set of information [3]. Thus, topic maps are a perfect basis to realize the mentioned mechanisms and functionalities for the identification of relevant knowledge objects. In this contribution, we describe the fundamental idea of the research project K-Discovery: Using topic maps to identify distributed knowledge structures in groupware-based organizational memories. In addition to the existing techniques of search and navigation, such as hierarchical navigation in categories of views or full text search, users can be supported by the strong concept of associative navigation in and visualization of semantic networks (provided by topic maps) (cf. [4]). Further, a groupware-based implementation of topic maps benefits from several aspects of groupware paradigms and technologies, for example distributed database architectures and replication, workflow management, or security and access control mechanisms.

The purpose of this paper is to support users in finding relevant knowledge in distributed organizational memories efficiently by combining groupware paradigms, topic maps, and visualization techniques. This contribution is structured as followed: In section 2, we briefly describe a specific case study of a groupware-based organizational memory and we give a motivation for this work. The basic concepts of topic maps are shortly...
introduced in section 3. In section 4, we present the conceptual approach and primary idea of the K-Discovery project: Applying topic maps to groupware-based organizational memories to identify distributed knowledge structures. In section 5, we discuss a concept for visualization and navigation of groupware-based topic maps based on hyperbolic geometry. Firstly, requirements of an efficient navigation tool for structured data and knowledge spaces are defined. Secondly, we suggest a concept of simplifying the relation structure of knowledge objects to enhance navigation efficiency of groupware-based topic maps. Thirdly, we introduce and discuss the HyperbolicTree concept for exploration of groupware-based organizational memories. The following short case illustrates the need for these mechanisms:

Within the Groupware Competence Center (GCC) of the University of Paderborn, Pavone Enterprise Office, an integrated correspondence, office, document, workflow, knowledge and archive management system, is used in almost every part of operative work. There are several core applications in different contexts, such as daily office work, web publishing, teaching and examinations, projects and cooperation, research, and literature maintenance. These organizational memories consist of more than 30 databases with more than 60,000 documents. Even though there are a lot of semantic relations between the information in these databases, it is not possible to navigate between them or to identify knowledge structures. The capabilities for accessing the knowledge objects are restricted to full text searching and navigating through context sensitive views and category structures. Full text indexes are not sufficient for searching information, and structures like document types or taxonomies are sometimes too constraining to qualify or categorize information [6]. Furthermore, usage and scope of these techniques is limited to a single database.

Analogous to this practical case, many organizations have a strong integration of groupware systems, both in their IT-infrastructure and in their working environment. (To name only two examples, look at Deutsche Bank and PriceWaterhouseCoopers. Both of them use an extensive IT-infrastructure based on Lotus Notes/Domino.) With growing organizational memories on the one hand, and the lack of powerful, effective mechanisms and functionalities for navigating, linking, searching and investigating on the other side, the need for enhanced access mechanisms to explore and use the complex offer of knowledge is getting more and more evident. To address these shortcomings the ISO standard ISO/IEC 13250 Topic Maps defines a model and architecture for the semantic structuring of link networks. By applying topic maps to large sets of heterogeneous information resources, reusable structured semantic link networks are created above those resources. Furthermore, the standard provides navigation paradigms to enable users to search knowledge structures [3].

The key concepts of topic maps are topics, occurrences of topics, and relationships between topics (topic associations). Firstly, a topic is a construct that represents a real world subject and in this sense a topic can be almost everything: a theme, a concept, a subject, a person, an entity, etc. (v. fig. 1). A concrete topic is an instance of a topic type. Therefore, a topic and a topic type form a class-instance relationship. At the same time a topic type is also a topic.

Topics have three kinds of characteristics: names, occurrences and roles in associations. Occurrences and roles in associations are described in detail later on. The base name of a topic is required. In addition, topics can have a display name and a sort name. These concepts apply to multilingual scenarios or to the use in different geographical regions.

A topic may be linked to one or more real knowledge objects, like a report, a comment, a video or a picture, that are considered to be relevant to the topic in some way. Such knowledge objects are called occurrences of the topic (v. fig. 1). Generally, an occurrence is not part of a topic map. The link mechanism itself depends upon the
underlying system. For example, HyTime addressing, XPointer, or document links in a groupware system can be used as a link mechanism. To express different kinds of occurrences, the standard provides the concept of occurrence roles that are topics as well.

**Topic associations** describe the relationships between topics (v. fig. 1). They are completely independent of the real knowledge object and represent the essential value- add of the topic map. This concept leads to some conclusions: A concrete topic map can be applied to different information repositories. Seen from the opposite side, different topic maps can be applied to one information repository and therefore, they can provide different views e.g. for the different needs of different users. Furthermore, topic maps are interchangeable and can be merged.

![Figure 1. Key concepts of topic maps](image)

Generally, topic associations are not one-way relationships. They are symmetric as well as transitive and thus, they have no direction. The construct of association types can be used to group topic associations and the involved topics. An association role describes the role of a topic in a topic association. Again, both the association types and the association roles are topics as well.

Additionally, the topic map standard provides the extended concepts of scope, public subject, and facets, which are not described here. For a comprehensive reference refer to [3] and [8].

### 4. Using topic maps in groupware-based environments

By applying topic maps – as defined in ISO/IEC 13250 – to groupware-based organizational memories, it is possible to close a gap between knowledge and information. So far, views and categories in groupware systems have enabled the creation of little knowledge by conveying documents into different contexts. But this concept is limited to a weak identification of relations between them. Topic maps provide much more appropriate and sophisticated concepts to overcome this limitation and to build a semantic link network from the documents of a database. The basic design elements of a groupware-based application, especially forms, can be used to identify the main subjects (topic types) and relationships between them (association types). (The description of design elements of groupware-based applications is beyond the scope of this paper. For a comprehensive overview, refer to [7].)

While text analyzing methods with the support of artificial intelligence concepts could provide more or less reasonable results in identifying the main topics of a document, we have chosen another approach in the Knowledge Discovery project. To identify the relevant subjects of documents, the fields of the forms must be considered. As mentioned before in the short case, in many groupware systems the knowledge objects, particularly documents, contain many fields of various kinds and purposes: obvious information, like the parts of an address (e.g. city), or the category of a report (e.g. balance sheets 2000), or keywords and linking information (e.g. persons, locations, time, free or predefined keywords). These fields form a set of potential candidates for topic types. Moreover, by looking at the forms as a whole and at the views, a lot of basic associations can be found between topic types. For example, in the context of an address form, the associations “person lives in city” or “person works for company” can be identified. Similarly, a report form contains associations like “report is composed by person” or “report belongs to category”. When considering views, the content of fields is displayed in the columns of a view. This mapping between fields and columns creates another context and thus forms a base for additional associations, like “report is in work by person” in a view, that displays the state of documents in a workflow. Obviously, the different forms, i.e. the different document types, can be considered as occurrence roles, while documents are occurrences in a concrete topic map.

These considerations about fields, forms, and views of a groupware-based application, and their connections and relationships, result in a template for a topic map. By applying this template to groupware-based organizational memories, a concrete topic map can be built up. In particular – following the definitions of topic types, association types, and occurrence roles made in the template – topics, topic associations, and occurrences are derived from all documents (for architectural and implementation approaches refer to [9]) and form a topic map. This topic map provides a structured semantic link network and therefore allows easy and selective
navigation to the required knowledge objects, independent of their origin (cf. [3]). In contrast to full text search and hierarchical navigation, searching in knowledge structures is enabled, i.e. search takes place in a meta level above the knowledge objects, and transparent access to these objects is provided. Further, knowledge objects can be found in new contexts, regardless of potential manually assigned categories in groupware-based organizational memories, and therefore, new contextual knowledge can be gathered (cf. [10]).

To sum up, applying topic maps to groupware-based organizational memories provides a huge potential and a lot of advantages for organizational knowledge management. Seen from the perspective of groupware, groupware systems leverage the benefits of topic maps through their basic concepts and paradigms, like distributed database architectures and replication, workflow management, or security and access control mechanisms. A groupware-based environment enables the definition, generation, and maintenance of a topic map in a consistent manner, on both a stationary and a mobile workstation. Thus, the distributed management of a topic map is independent of time and location. Differentiated access rights can be defined for a certain set of knowledge objects, single knowledge objects, and any part of a knowledge object, as well as for functionality and presentation elements and design elements (like forms and views). These mechanisms enable the realistic adaptation of access structures against the background of complex and real organization models. As an example, some organizational members form an editorial team, which ensures that all published components of a topic map adhere to organizational standards. Another key concept of groupware systems, the workflow management, can be used to describe and support the whole process chain of topic map publishing. A model of an organizational integration of topic maps and supporting workflows is defined in [9].

The described concepts, paradigms, and functionalities of groupware systems facilitate the team-based and organizational wide use of topic maps. In combination, both the groupware systems and the concept of topic maps leverage their mutual potentials.

Based on these conceptual approaches, the K-Discovery project defines a framework and architecture for a groupware-based knowledge management system using topic maps. This architecture integrates repositories as well as interfaces to these repositories, and tools, that access the interfaces of the infrastructure model, the process model, and the topic map model of an organization. By applying a topic map to the set of completed workflows, it is possible to identify relationships between them as well as tacit and inherent process knowledge. The resulting knowledge can be used in several ways: For example, to group similar workflows and to create predefined workflow. Furthermore, a suitable framework for a new kind of knowledge management applications – based on the underlying concepts and paradigms of groupware and topic maps – is provided. An example is a repository of workflows, or the support of the analysis of the process structure as well as the organizational structure. A concrete application for distributing workflow process knowledge in organizations as well as for the semantic associative navigation in process networks is described in detail in [11].

Additionally, the K-Discovery project introduces an implementation approach that provides a prototypic application and tool framework that implements some of the functionalities of the described architecture and thus, allows a generic and almost automatic generation of topic maps upon groupware-based organizational memories (by using various technologies like groupware technologies, relational database management systems, Java, XML). The K-Discovery architecture as well as the implementation approach are described in detail in [9].

5. Visual navigation of groupware-based topic maps

So far, we have motivated the usage of topic maps to create knowledge structures in groupware-based organizational memories. In the following sections, we introduce a visual approach based on hyperbolic geometry to efficiently exploit these knowledge structures. The original K-Discovery topic map navigator only offered a textual user interface, which was less intuitive to use than a visual browser, because users had to be familiar with basic topic map terms and concepts, and could not overlook larger parts of the structure. Therefore, we will present an intuitive navigation approach to offer users a visual interface to use groupware-based topic maps in a more explorative way. Using the so called HyperbolicTree concept, this approach makes it possible to visualize the complete knowledge structure on the screen. Moreover, it adds means of recognizing specific parts of knowledge structures by visually highlighting areas that are important to users. These highlights are referred to as landmarks.

5.1. Requirements of an efficient navigation tool for structured data

The use of landmarks [12] is a very important factor when considering a navigation space for users, as they support user orientation in complex hierarchical structures. Landmarks have a strong positive effect on the
work efficiency of people that often navigate the structure and thus are familiar with it, as well as on people that navigate the structure for the first time. Additionally, landmarks make users feel well acquainted with the structure and as a result, amend the user acceptance of the system as well as user satisfaction.

Therefore, a navigation interface should support landmarks. It can be thought of different types of landmarks, such as color coded objects, bold text, and particularly thumbnails, graphics, and icons. The navigation of the structure should be as simple and intuitive as possible, and the interface should allow direct manipulation and re-adjustment of the area that is focused by the user, to effectively perform searches in adjacent and non adjacent areas of the current focus.

Especially in the field of explorative collaborative systems it is important, that “users become completely absorbed in their task domain.” [13]. “Focus + Context” techniques can help users to become absorbed by the task domain navigation and knowledge browsing. Focus + Context is the name of a class of visualization concepts, that displays structured data in a way that makes it possible to present as many relevant knowledge objects as possible on the screen at the same time. In a Focus + Context user interface, the area that gets the current users attention is called focus, and the surrounding area is called context.

The basic assumption is that users should be able to view focus and context areas at the same time and to present an overview of the whole knowledge structure. That enables them to easily refocus the tree into areas that are close to the current focus and thus, makes navigation more efficient. It can be assumed that the user needs more detailed information in the context area than in the focus area. The reason is that the visual acuity, and thus, the sensual faculty, rapidly decreases outside of the focus area [14]. Therefore, it is not necessary to offer detailed information about knowledge objects, which are currently of minor importance. [15] even states, that unnecessary knowledge objects burden users, because of the limited capacity of the human nervous system.

A perfect user interface would visualize the whole knowledge structure at the same time on the screen, to give users an ideal overview of the structure. As it is not possible to assign each knowledge object the same display space, consequently the display area per knowledge object in the context area is less than in the focus area, while the areas are not strictly divided. Moreover, the used area per knowledge object decreases with increasing distance from the focus, proportional to the sensual faculty of the user. As the available display area per object decreases, the amount of information per object has to be decreased as well. Techniques that can be used to decrease the amount of information are filtering, selective aggregation, highlighting, and distortion [16].

Detailed information and examples of early implementations of Focus + Context technologies can be found in [17]. More current examples are presented in [18] and [19].

To summarize, very important requirements of a navigation tool for visualizations of structured data are that the navigation should be efficient, the usage should be easily to learn and intuitively to use. Focus + Context characteristics help to increase the efficiency of navigation. Additionally, it should support landmarks.

5.2. Projecting semantic link network structures to trees

In the field of visualization of knowledge structures, an important aspect is the visualization of the relations between knowledge objects. Often, a relation or connection between two objects is represented by a line or an arc drawn between them. In network structures, such as the WWW or a complex semantic network like a topic map, navigation of visualized network structures can be difficult and not very intuitive. This is due to the fact, that all visible objects can have several connections between each other, and some of them cannot even be drawn because of several reasons. For example, objects can be located outside of the display area and users cannot see the object at the end of the line, if there is any. Or in a Focus + Context user interface, objects can be located in the distorted area of the structure and thus connections between them cannot be displayed. Additionally, a clear layout, that guarantees a good oversight over the entire network structure is hardly impossible, as the data is not hierarchically structured. Fig. 2 illustrates that even a simple network with only 13 nodes and all connections visible, makes it hard to visually identify which nodes are connected.

Therefore, we suggest projecting the semantic link network structure to a hierarchically organized model such as a tree. In a tree, one object has to be defined as the root object. This could be the best match out of a search using a groupware-based topic map or a defined root object in the companies’ intranet etc. With that root defined, all related objects can be laid out from the root as children and descendants of that object. The network structure becomes directed that way, as the path through the network always starts at the defined root. Fig. 3 shows a simple network with very few connections between the nodes to illustrate the projection mechanism. In the network, the node named “A” is chosen as the root node for a tree structure that results from the projection. Three types of projections can be differentiated. The resulting
tree of a projection is called *strict projection*, if each node in the network, that is directly connected with the root node, is a child node of the root node in the tree. Then, each connected node is recursively examined for connected nodes in the network, and respectively additional nodes are added to the tree.

As connections of objects in an undirected network always represent two directions, each node in a strict projection has one child node that is syntactically the same as its parent node. As this relation is not relevant to a user that navigates a knowledge structure, these relations do not have to be considered to be part of the projected tree. A projection that is reduced in the described way is called *clean projection*. Clean projections can still contain irrelevant connections, because paths in a network often contain loops, which let syntactically equivalent objects appear several times in the projected tree. With loops, the projected tree can grow up to an infinite level. Therefore, loops have to be eliminated while building up a projected tree. Loops in the tree can be easily recognized by looking for equivalent nodes on the path from the root node to the current node. Fig. 4 shows a so called *loop filtered projection* of the network that is presented in Fig. 3.

If the path from the root to the leaf is very long, the structure of the tree is called *deep*. The depth of the resulting tree has to be limited to ensure an acceptable performance during the projection process. If many nodes in a tree have many child nodes, the structure of the tree is called *broad*. Projected trees have the characteristic that they can grow very deep, while they seldom grow very broad. The reason is that knowledge objects are heavily linked between each other if the whole knowledge structure is considered, but each object contains a limited number of links, which are represented by the children of the node.

Thus using the concept of projection, the user can easily navigate the knowledge structure using a visual interface that is optimized for navigating hierarchies. Additionally, tree structures are very well suited to be navigated using Focus + Context techniques. But the projection of the network only visualizes that part of the network, that displays the context of the currently defined root node. Therefore, the user interface must feature an option to select every node in the network as the root node of the tree. That way, the user can navigate deeper into the structure while keeping the maximum depth of the tree fixed, by selecting each node to be the new root node and continue the search for relevant knowledge objects from that point. Landmarks could be used to highlight the descendant nodes of the current root, that have been root nodes once before. That way, users can easily see their entire search path in the network.

5.3. Visualizing of groupware-based topic maps with the HyperbolicTree

To create efficient, ergonomic navigational concepts, different teams in the research field of Human-Computer-Interaction (HCI) introduced innovative navigation interface concepts, that should increase efficiency of known browsing concepts such as the Microsoft Windows Explorer. We chose the HyperbolicTree concept as the most suitable approach as a navigation interface for semantic link networks. A HyperbolicTree initially displays a tree with its root node at the center of the display area. The hierarchical structure is then laid out on a plane in hyperbolic space and mapped onto a circular display region. The circumference of a circle on the hyperbolic plane grows exponentially with its radius, which means that exponentially more space is available with increasing distance from the root node. The tree uses distortion in the context area to reduce the information per knowledge object ratio. That way it is possible to display an overview of the entire structure on the available display area.

Projections of networks mostly have a deep rather than a broad structure. [19] have shown that a HyperbolicTree is very well suited to visualize and efficiently navigate this kind of hierarchical structures. It is not suited to visualize...
very broad structures, because the distortion effect in tree levels close to the root is already too strong, which makes the structure inconvenient to navigate. But very broad structures do not occur very often in projected network structures.

Figure 5. HyperbolicTree including landmarks

Fig. 5 shows a HyperbolicTree that uses graphical landmarks such as thumbnails and icons, as well as color coding textual information and node background. For more information about the HyperbolicTree concept see [18] and [19].

For efficient projection and tree creation, the depth of the tree has to be limited in real world environments. This situation is acceptable because the root node is the result of a user performed search, odds are good that the user finds the needed information somewhere near the root area. In case interesting information has been located, the user can select the node of interest as the new root node of the projection. A new projection of the network is then performed, starting at the newly located node. That way, users explore the environment of that node. Due to the nature of the HyperbolicTree concept, the depth limitation of the tree during the projection can be very high which enables users to search a large part of the knowledge structure without having to refocus the tree to a new root. Additionally, the layout technique of the HyperbolicTree offers great flexibility in the use of graphical landmarks. A comprehensive overview of several visualization concepts other than the HyperbolicTree can be found in [16] and [20].

By applying the HyperbolicTree concept to groupware-based topic maps using the projection of semantic link network structures, as described in section 5.2, it is possible to provide users with a visual navigation tool. In this newly created interface, nodes represent topics, occurrences, and topic associations (v. fig. 6). At the beginning of a navigation process, the root node is either manually selected by users or is the result of an initial search in the topic map. All descendants of this root node are presented and each of them can be selected to be the new root node. Sub-trees of a node can be dynamically displayed or hidden.

Figure 6. Visualization of a topic map using the HyperbolicTree concept

6. Conclusions and future areas of research

An approach of discovery and representation of knowledge structures in groupware-based organizational memories based on the ISO standard ISO/IEC 13250 Topic Maps has been presented. The K-Discovery project combines groupware paradigms, topic maps, and visualization concepts, leading to substantial synergies for organizational knowledge management. These synergies leverage their mutual advantages and therefore, provide benefits for the research fields. From our point of view, the K-Discovery project provides a considerable framework to identify, visualize, and navigate knowledge structures in groupware-based environments. Therefore, it supports users in finding required knowledge in organizational memories. In particular, a graphical visualization and navigation tool based on hyperbolic technologies provides users with intuitive access mechanisms to required knowledge, that can alternatively be used instead of the existing topic map navigator, which provides a tabular and textual user interface. In addition, the findings are of high practical relevance as shown in the short case. In fact, there are companies that have shown interest in a solution based on the K-Discovery framework.

Prototypes of the core modules and the graphical visualization and navigation tool have been implemented and can be used for trial purposes. Eventually, further knowledge management applications upon this framework for groupware-based topic maps will be developed. Specifically, taxonomy issues in virtual organizations will be addressed.
References


