

## Analyzing industrial clusters using measures of structural complexity management

Danilo Marcello Schmidt<sup>1</sup>, Marc Haas<sup>1</sup>, Daniel Kammerl<sup>1</sup>, Julian Wilberg<sup>1</sup>, Maximilian Philipp Kisse<sup>2</sup>, Udo Lindemann<sup>1</sup>

<sup>1</sup>Institute of Product Development, Technische Universität München, Germany

<sup>2</sup>Soley GbR, Munich, Germany

**Abstract:** Companies organize in industrial clusters to exchange knowledge, to identify new options for cooperation and to improve the regional competences for a special industry sector. For optimizing industrial clusters, it is necessary to assess factors influencing performance or effectivity of industrial clusters. This evaluation or analysis of cluster's performance can reveal strengths and weaknesses of the cluster. Interpreting the weaknesses might detect activities for improving the performance of the industrial cluster. For this performance analysis, we use measures and metrics of structural complexity management to investigate the cluster's inner structure, e.g. the cooperation and linkage between employees of companies, which are in the same cluster. We applied the measures at the MAI Carbon cluster and interpreted the results of the performance analysis. The user data of the cluster's online platform serve as the basis for this analysis.

*Keywords:* DSM, DMM, Structural Complexity, Industrial Cluster, Cluster Structure

### 1 Introduction

An industrial cluster is a pool of several companies, which agreed to cooperate and to inter-change knowledge. Porter (1990) is known as the founder of the cluster theory, the majority of today's research refers to Porters work for a definition of industrial clusters (Feser and Bergman, 2000; Morkovina et al., 2014; Padmore and Gibson, 1998). Reasons for organizations and companies to organize themselves in an industrial cluster are the globalization and its requirement for companies to compete with other companies for product knowledge, learning ability and flexibility (Porter, 1990): The needed knowledge and abilities are too complex for companies to fulfill them by themselves across all stages of production. The division of labor increases as well as the link between companies. Not the company itself, also the company's surroundings determine the company's success. This makes the regional factors of a company essential for the company's competitiveness. According to Porter (1990), industrial clusters are always of regional nature, examples for this are the semiconductor industries in Silicon Valley or the automotive industries in Detroit. To support industrial and regional clusters, the program "go-cluster" of the German Federal Ministry for Economic Affairs and Energy promotes industrial clusters like the MAI Carbon cluster in the region of Munich, Augsburg and Ingolstadt for carbon composites.

In this context, an evaluation of the efficiency and effectivity of industrial clusters is needed to identify the strengths and weaknesses of the industrial cluster. The result of such

an evaluation can reveal relevant points for optimizing and improving the cluster. Furthermore, a suitable level of efforts and costs to build industrial clusters might be derived from the results of such an evaluation. This helps to implement an industrial cluster in a cost-efficient way.

In this paper, we define methods for evaluating industrial clusters and apply them at the MAI Carbon cluster. For this, we use data of the cluster's online platform and analyze them from different perspectives. We first explain the state of research of evaluation methods for industrial clusters. Then it comes to the methods we used to analyze the cluster and apply them on the online platform of MAI Carbon. After interpreting the analyses' results, we proved a short conclusion and outlook.

## 2 State of Research

In literature, many approaches exist to evaluate clusters' efficiency or performance (Bergman et al., 1996; Czamanski and Ablas, 1979; Feser and Bergman, 2000; Kozhinova and Ferova, 2012; Morkovina et al., 2014; Mukhiddini and Bobojonov, 2014; Padmore and Gibson, 1998; Ribeiro Carpinetti et al., 2008). There are several approaches, whose focus lies in algorithms and methods to cluster or to identify clusters, or to evaluate the suitability of companies for the cluster (Bergman et al., 1996; Czamanski and Ablas, 1979; Feser and Bergman, 2000; Kozhinova and Ferova, 2012). Those algorithms describe ways to find clusters in a set of relevant companies. Some of them refer to Porter's view on industrial clusters (Kozhinova and Ferova, 2012) and base their evaluations on evaluating competitive advantages (e.g. technological development, system of support for new firms establishment) Those approaches use criteria like "assistance to total employment" or "sustainability of economy" (Kozhinova and Ferova, 2012) for cluster evaluation. Those evaluation methods are qualitative and they do not consider the actual state of an industrial cluster, but the cluster mechanisms. Feser and Bergman (2000) provide a framework for applied regional cluster analysis to assess the fitness of regional industry sectors for an industrial cluster. For this, they quantify criteria like number of employees or companies (Feser and Bergman, 2000). However, their focus is not on the actual state of the cluster but only on the identification of industrial clusters. Furthermore, they do not analyze the inner linkages between the companies, which is relevant for a clusters' performance (Bergman et al., 1996; Czamanski and Ablas, 1979; Feser and Bergman, 2000; Kozhinova and Ferova, 2012). For this reasons, those measures are not relevant for our considerations, as we want to analyze industrial clusters by identifying their strengths and weaknesses.

Other authors focus on the actual state of industrial clusters to evaluate their current level of performance (Morkovina et al., 2014; Mukhiddini and Bobojonov, 2014; Padmore and Gibson, 1998; Ribeiro Carpinetti et al., 2008). Mukhiddini and Bobojonov (2014) use criteria like "number of workers and the general assets of the company" or "labor productivity" to indicate the development stage of an industrial cluster. This might be helpful for decisions concerning activities to configure or organize the cluster but not for identifying weaknesses or need for optimization (Mukhiddini and Bobojonov, 2014). Padmore and Gibson (1998) refer to Porter's diamond model (Porter, 1990) and use six criteria to evaluate clusters: Resources, infrastructure, supplier and related industries, firm structures, strategies and rivalry, local markets and access to external markets (Padmore

and Gibson, 1998). They score clusters concerning those six determinants by comparing several industrial clusters and ranking them. This is a qualitative way of assessing clusters, which can barely help in identifying new weaknesses, if they exist in all regarded clusters. However, it is a debatable point whether different clusters are comparable on this level. Furthermore, this methodology lacks in analyzing the interactions between different companies. Ribeiro Carpinetti et al. (2008) took ideas from the balanced scorecard (Kaplan and Norton, 1992) to evaluate clusters performance and used metrics like “labor productivity”, “total labor force” or “percentage of companies involved with cooperation” (Padmore and Gibson, 1998). They consider more abstract measures to assess the performance. To quantify the cooperation between companies, they took criteria like “collective purchasing of raw material” or “establishing a cooperation program” into account. As they did not look at the actual cooperation of employees from different companies, their statements regarding the cooperation between companies are too abstract to obtain detailed results for improving the cluster concerning the cooperation of employees from different organizations. Morkovina et al. (2014) consider the relations between clusters’ companies by asking employees for their cooperation with suppliers or competitors. Furthermore, they use simple measures like “fraction of the volume of production of small enterprises in the volume of production of all cluster members” (Morkovina et al., 2014). We claim that we can reach more meaningful results by applying other measures of structural complexity management.

The considered approaches from literature do not consider the inner structure of industrial clusters or the communication or cooperation of employees working at integrated companies. This aspect is quite important for evaluating clusters’ performances, as abstract measures considered by the considered approaches might not include the characteristics of industrial clusters. For this reason, we claim that the usage of measures and metrics from structural complexity management can help in identifying weaknesses and strengths of cooperation between companies belonging to the same industrial cluster. We apply those measures on data of the online Web 2.0-platform of MAI Carbon, which is a networking platform for employers working in MAI Carbon. Then, we interpret how results of applied measures help in analyzing cluster’s performance.

### **3 Methods of Structural Complexity Management for Analyzing Industrial Clusters**

In order to analyze the structure of the mentioned cluster it is necessary to have a structured approach for this problem. One solution is to manage the complexity of the system by structuring the huge amount of data in matrices and graphs by using methods of structural complexity management.

Using a Design Structure Matrix (DSM) allows analyzing the structure of systems. A DSM is a matrix with an equal number of columns and rows (square matrix). It allows to map elements of the system and their relationship (Maurer, 2007; Steward, 1981).

Later, the DSM was extended to Domain Mapping Matrices (DMM) (Danilovic and Browning, 2007). The difference to a DSM is that a DMM includes not just one domain of data. A DMM can handle two domains (Yassine et al., 2003).

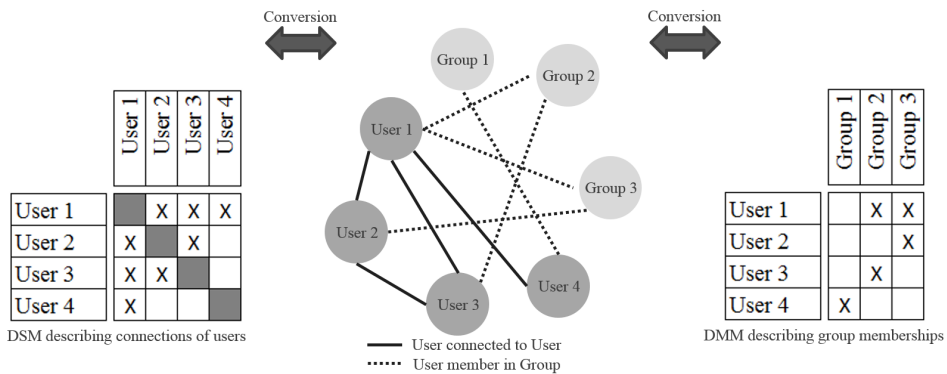


Figure 1. Example of a binary DSM and DMM and the conversion from or to a graph

Figure 1 shows an example of the coherence of binary DSMs and DMMs and graphs that contain nodes and edges. The DSM shows if there is a connection between two elements of the same type (here: Users of a system). The DMM shows which user is member in which user-group. It is possible to convert the matrices into a graph by defining the elements of the matrices as nodes and the relationship between the elements (“X” in the matrix) as edges and vice versa. This allows to structure and depict a problem of great complexity for example when handling a huge amount of data in network analysis.

The general approach to the analysis of the cluster is using methods of structural complexity management and graph theory to evaluate key measures of the network.

The key measures that will be used to analyze the cluster will be the following (see also (Maurer, 2007) and (Kreimeyer, 2009)):

- **Number of nodes:** Significant for the size of the network.
- **Number of edges:** Determinates the level of interactions within a domain.
- **Number of edges per node:** Describes the level of cross-linking within the network.
- **Number of unconnected nodes:** Describes the amount of nodes disconnected of the network and thus independent or not taking part at all.
- **Number of leafs:** Number of nodes only connected to one other node and therefore possess a low level of interaction within the network.
- **Number of interfaces between domains:** Determinates the level interactions between domains.
- **Active and passive sum:** Degree on how impactful a node is to others.
- **Number of domains:** Evaluation of the diversity of domains in the network.

This methods will be applied in order to manage the complexity of the following case study and quantify the measures mentioned above. The data will be analyzed using the software tool “Soley Studio”. This tool allows to generate graphs and evaluations out of unstructured data while handling the relationships and dependencies of system-elements within several domains as well as multiple attributes per element. The software works using the theory of matrix analysis methods and illustrates the results in data graphs.

## 4 Case Study

This case study considers a social network for sharing knowledge and further cooperation between people and organizations in the carbon composite industry. The network is implemented as an online platform where interested people can register to get in contact with like-minded people from different companies and organizations to work together on projects, share and create knowledge and therefore boost the innovation process in the carbon composite field.

Registered users can send contact requests to other users, send private messages, get engaged in a theme or project group and upload documents and comments. The activity of the actors of the online platform is captured and the gathered data is stored. This data will be used to examine processes on the platform and the connections and relations between elements of the network (e.g. users, groups, companies, etc.).

The online platform started in 2013 into the beta phase. The official go-live was in November 2014. Until April 2015 there are more than 780 registered users, e.g. from automotive companies, mechanical engineering companies or related research institutes.

### 4.1 Data Acquisition

The gathered data consists of real user data of the online platform. The researchers have direct access to all necessary data that is required to analyze the work of the online platform. Approximately every 5 weeks a new data report is generated and implemented into the dynamic analysis system. The first report dates back to January, 2014 and the latest report was released in February, 2015. So the analyzed data covers a time span of over a year. These data reports contain:

- User information: UserID, email address, company, location, branch, position, etc.
- Group information: GroupID, group name, group type, number of members
- User contact information: Sent, received, accepted, declined and ignored contact requests from user to user.
- Group member information: number of groups per user, number of members per group.

In the following chapter the results of the analysis will be presented. First of all there will be a few facts about the size and the growth of the network. Afterwards, an in depth analysis of the user contacts and the group memberships will be made, followed by an examination of the different branches that cooperate by using the network and connecting to people from other branches.

### 4.2 Results and Interpretation

The amount of registered users (and therefore the number of nodes) rose by 116% from 360 users in January 2014 to 778 users in February 2015. The average number of contacts per user declined from 4.5 to 3.4 contacts per user during that year. This is because the number of accepted contact requests (number of edges) has not grown as much (60%) as

the mentioned number of users. One possible explanation is that newer users are not as likely to connect to other users.

Analyzing the number of group nodes shows also a rise by 134% to a total of 257 groups in February 2015. On average there is one new group created every four days. This information gives an overall picture of the current state of the network. It is growing in size, but there are a few hints that the activity of the users has declined since the start.

The following Figure 2 shows a force based graph of the user connections, created with Soley Studio.

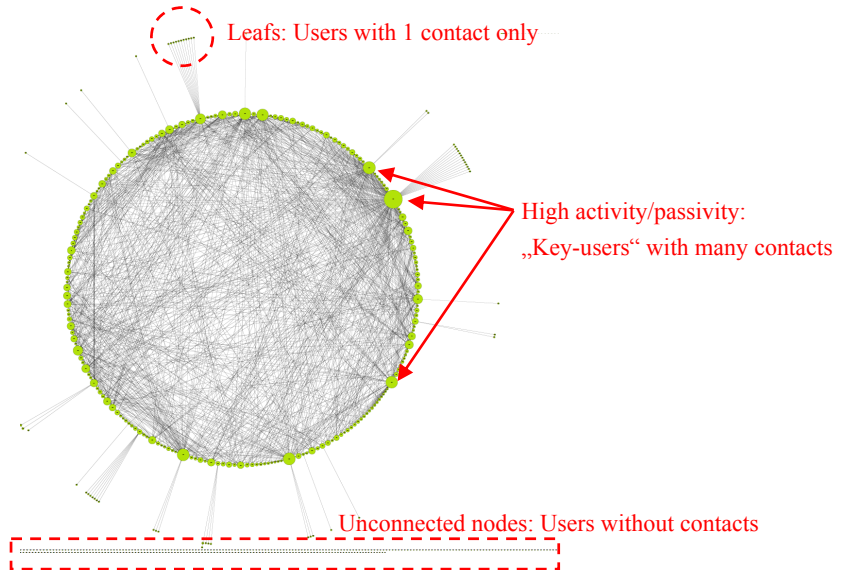


Figure 2. Graph showing connections of users

The nodes (green) represent users of the network. The diameter of the node is mapped to the number of contacts of that user activity + passivity. The bigger the diameter, the more contacts does the user have. Since the direction of the user connections is not considered, the node size is mapped to the sum of its activity and passivity. The edges of the graph are undirected and only represent that there is a connection between two users which means that one of them sent a contact request and the other accepted. It is easy to see, that there are very few key-users with many contacts i.e. users with high activity/passivity. The vast majority of users has only one or no contacts at all. The users with only one contact (leafs) can be found outside of the “inner circle”. Users without contacts (unconnected nodes) are placed at the bottom of the figure. In total, 61 % of the users have no contacts. This means, that one of the key goals of the network, to encourage people to connect with others, has not been reached.

Explanations to this facts can be found in Figure 3. These graphs were created out of User-DSMs like the one in the example in Figure 1. Analyzing the active sum of the users gives us the number of contacts per user. The left graph shows the number of contacts per user by order of registration date. The closer the data point on the horizontal axis is to the

origin, the earlier this user has registered. The right graph arranges the users by order of number of the user’s contacts.

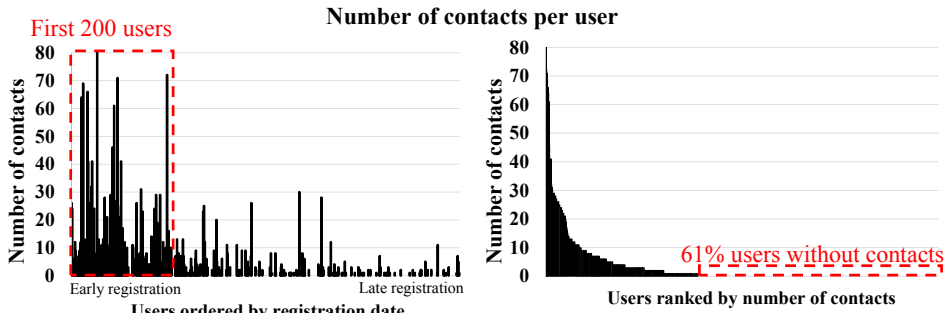


Figure 3. Analysis of number of contacts

This graph confirms the impression of the existence of only a few key users with a lot of contacts and therefore a high activity/passivity. It also shows the great number of users without contacts. The left graph verifies the assumption that newer users have fewer contacts. In fact, the first 200 users combined have 1963 contacts while the other 578 users only have 663 contacts. That means that 75% of the user connections are held by the first 200 users. One possible reason is that new users have trouble in finding interesting contacts or other users they know. Another possibility is that users need some time to get to know the functions of the network in order to use it efficiently. It is also possible that new users are simply not interested in the platform or not convinced that the use is beneficial to them.

The DMM of user and group data allows to analyze the usage and efficiency of sharing knowledge, using the measure “number of interfaces between domains”. The DMM looks like the example in Figure 1 with users placed down on the side of the matrix and groups on the top. This allows to get a first impression on how well the groups are integrated into user’s usage behavior. The number of groups per user and the number of members per group can be found by building the active and passive sums of the matrix. Because of its size, this DMM is not shown but used for analysis. Figure 4 shows the results.

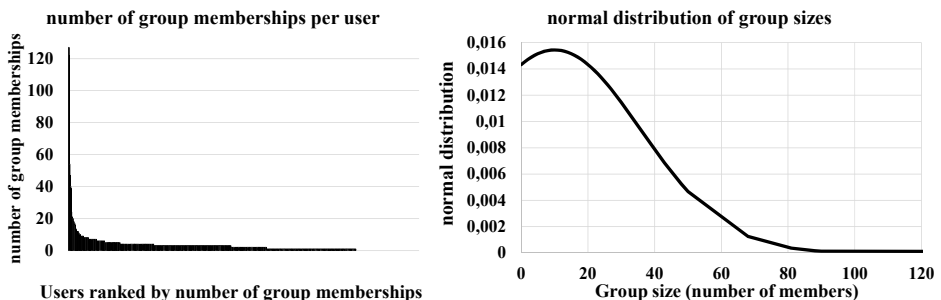


Figure 4. Analysis of group memberships

The left graph in figure 4 shows the number of group memberships per user, sorted from most to least group memberships. It is easy to see that, similar to the user contacts, there are very few key users that are member in a lot of groups. This nodes therefore possess a high amount of inter-domain interfaces while the vast majority of the users do not seem to be interested in group memberships as much. In fact, 18% of users are not members of groups at all. 79% of users take part in 1-10 groups while only 3% are members in more than 10 groups. Three of the users are member in more than 100 groups so that could mean that this users are administrators of the network. The right graph of figure 4 depicts the normal of the group sizes concerning the number of users belonging to the group. The maximum of this distribution is at 10 members per group. This shows that most groups are connecting 10 users, which is an advantageous number of users per group: Too many users in a group might decrease the efficiency of the communication for this group. A number smaller than 10 users belonging to a group might be helpful for the efficiency of the communication, however, a small group connects a too small number of users to each other to be useful for the cluster. Overall, users seem to accept groups better than direct connections to others. Users might consider being members in groups more beneficial to them than actually connecting to other users. The advantages of groups reveal the reasons for this matter. Groups offer possibilities to get quick access to information and other functions such as, uploading and sharing documents, videos and other useful data, planning events using a calendar app and discussing topics with the other group members. Another possible explanation is that users are more likely to connect to people they know. Since the user base is quite small it is possible that users simply can't find people they know and hesitate to contact foreign users. This impacts the cooperation between companies and branches which is also a goal of the network.

The following Figure 5 shows the connections of branches via users. This graph was generated by combining a user-DSM together with a user-branch-DMM. The number of domains is three. Suitable methods for combining DSMs with DMMs are explained by (Maurer, 2007) and (Kreimeyer, 2009). Whenever there is a connection between two users of two different branches, the branches are considered connected. This is shown in the graph by edges with different thickness. The thicker an edge is, the more connections exist between users who work in the respective branches. Similar to Figure 2, the size of the branch nodes refer to the number of people working in this branches. This information allows a detailed picture of the interdisciplinary connections of the network.

Overall, there are 224 different companies of 28 different branches existing in the network. The thick edges shows that the interdisciplinary connection of users is high. Evaluating the data of the matrices results in 70% of the user connections are between users working in different branches which confirms the first impression and is a sign for an effective transfer of knowledge from one branch to another. This contradicts the assumption that users only connect to users they know since it must be doubted that most of the people connected by different branches know each other. An explanation for this contradiction can be that users most likely connect to users they share a group with since the groups are interdisciplinary.



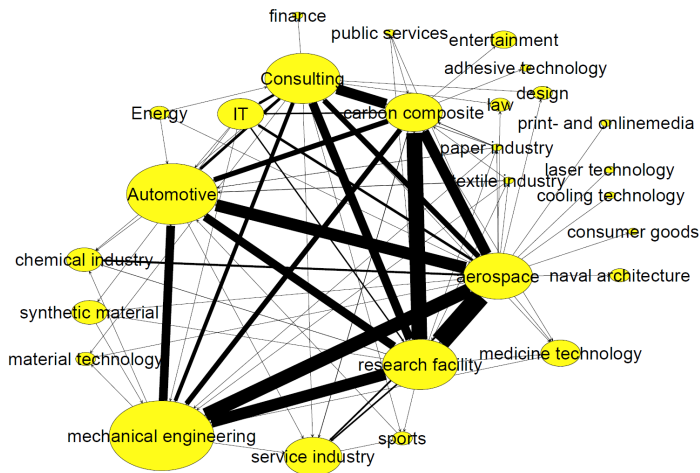


Figure 5. Cooperation and connection of branches

## 5 Conclusion

The results of the analysis of the available data show, that the network is slowly growing but the interaction between users must be improved. It seems as if the network is run by only a few key-users that have connections to a lot of others and are member in many groups while the vast majority does not take part in interacting and using the network as much.

Concerning the data provided it must be doubted that it is sufficient to get one data set per month in order to analyze the network development and track the changes to its user and group base in detail. It has to be examined if a live tracking of the network data will provide better results. But this also comes in hand with a lot higher effort in providing the data and implementing a live data-mining system.

The usage of methods of structural complexity management has provided detailed, quantitative results on the current state of the network. The advantages of using online-data of the platform are that they allow an in depth view into the general statistics of the network development and quantify its key-measures to evaluate its efficiency. However, this quantitative approach does not allow conclusions about the general usability and user-friendliness of the platform which is a major issue concerning how well the platform is accepted by its users and therefore an explanation on why or why not the activity of the users is as it is. It does not show perspectives on how to improve the major disadvantages and provide approaches on how to fix them.

A qualitative approach on examining the functions of the platform might be needed, however, structural complexity management does not offer methods for qualitative network analysis. Other methods must be used like user surveys or interviews to examine the key advantages and especially disadvantages the platform has. This allows a user-based feedback approach and provides useful information on how to improve the platform and its functions.

In further research, additional measures could be analyzed to give a more in depth view into the network using the provided data. This measures can be the clustering coefficient, path length, density degree or path centrality. Furthermore, the measures could be structured in goal trees or other methods, for example a balanced score card, to allow a better understanding of the measures and their correlations. This requires a stronger focus on the strategic goals of the platform rather than only looking at its key measures and deriving conclusions from that point. Combining quantitative and qualitative approaches as mentioned above could provide a better understanding and more detailed results on how well the platform performs and provides to reaching the strategic goals. This could also happen by analyzing strengths and weaknesses using a SWOT analysis.

## References

- Bergman, E.M., Feser, E.J., Sweeney, S., 1996. Targeting North Carolina Manufacturing: Understanding the State's Economy Through Industrial Cluster Analysis. UNC Institute for Economic Development, Chapel Hill, North Carolina.
- Czamanski, S., Ablas, L.A., 1979. Identification of Industrial Clusters and Complexes: a Comparison of Methods and Findings. *Urban Studies* 16, 61-80.
- Danilovic, M., Browning, T., 2007. Managing complex product development projects with design structure matrices and domain mapping matrices. *International Journal of Project Management* 25.
- Feser, E.J., Bergman, E.M., 2000. National Industry Cluster Templates: A Framework for Applied Regional Cluster Analysis. *Regional Studies* 34, 1-19.
- Kaplan, R.S., Norton, D.P., 1992. The Balanced Scorecard – Measures That Drive Performance. *Harvard Business Review* 70, 71-79.
- Kozhina, T.V., Ferova, I.S., 2012. Evaluation of Industrial Clusters Efficiency in the Context of Innovation Policy (on the Example of the Krasnoyarsk Territory). *Humanities and Social Sciences* 5, 153-160.
- Kreimeyer, M., 2009. A Structural Measurement System for Engineering Design Processes. Technische Universität München.
- Maurer, M., 2007. Structural Awareness in Complex Product Design, Institute of Product Development. Technische Universität München, München.
- Morkovina, S.S., Popkova, E.G., Santalova, M.S., Konstantinov, A.V., 2014. Development of Methodological Approaches to the Efficiency Analysis of Territorial-Industry Cluster Formation in the Forest Sector. *Asian Social Science* 10, 85-94.
- Mukhiddini, M.P., Bobojonov, R., 2014. Innovative methods for evaluating the performance and development of business clusters. *The Macrotheme Review* 3, 49-57.
- Padmore, T., Gibson, H., 1998. Modelling systems of innovation:: II. A framework for industrial cluster analysis in regions. *Research Policy* 26, 625-641.
- Porter, M.E., 1990. The Competitive Advantage of Nations, Harvard Business Review. Harvard College, pp. 73-93.
- Ribeiro Carpinetti, L.C., Cardoza Galdámez, E.V., Cecilio Gerolamo, M., 2008. A measurement system for managing performance of industrial clusters. *International Journal of Productivity and Performance Management* 57, 405-419.
- Steward, D., 1981. Design structure system: A method for managing the design of complex systems. *IEEE Transactions on Engineering Management* 28, 71-74.
- Yassine, A., Whitney, D., Daleiden, S., Lavine, J., 2003. Connectivity maps: modeling and analysing relationships in product development processes. *Journal of Engineering Design* 14, 377-394.

## **Acknowledgments**

We thank the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung) for funding this project as a part of the MAI Carbon cluster.

**Contact: D. M. Schmidt**, Institute of Product Development, Technische Universität München, Boltzmannstraße 15, 85748 Garching, Germany, +49 89 289 151 35, +49 89 289 151 44, [schmidt@pe.mw.tum.de](mailto:schmidt@pe.mw.tum.de), <http://www.pe.mw.tum.de>