Aalto University

School of Science

Degree Programme in Industrial Engineering and Management

Marina Timmermans

Patent analysis as an input to strategy: case of electric vehicle industry

Master´s Thesis

Espoo, May 28, 2014

Supervisor: Professor Robin Gustafsson

Instructor: Chris Rowell, M.Sc.

AALTO UNIVERSITY SCHOOL OF SCIENCE

ABSTRACT OF THE MASTER'S THESIS

Degree Programme in Industrial Engineering and Management

Author: Marina Timmermans		
Subject of the thesis: Patent analysis as an input to strategy: case of electric vehicle industry		
Number of pages: 74	Date: 28.05.2014	Library location: TU
Professorship: Strategic Management	Code of professorship: TU-91	
Supervisor: Prof. Robin Gustafsson	Instructor: M.Sc. Chris Rowell	

This research work examines technological developments of an emerging field from the perspective of patented innovations and major industry players. The selected domain of the study is the electric vehicle (EV) industry, which represents an emerging technological field driven by innovations. Patent-to-patent citation information and entities associated with each patent (i.e. patent holder, technology field, country) were utilized for the visualization of the relationship between patents in the form of a network. Bibliographic coupling (BC) is the methodology used to establish these relationships. From the viewpoint of companies, this relationship indicates similarities in the technological development direction and areas of R&D activities, suggesting for possible competition or cooperation. From the perspective of patented technology, the association indicates that the technologies or their applications are closely related.

The focus of the study is placed on the possibilities and limitations provided by patent analysis based on BC, so to facilitate further exploration and application of this methodology as a valuable tool for the support of managers' assessment of technological environment in real time and planning of the R&D projects within an emerging field. Managers can use patent maps as an additional source of information and communication support in the strategic decision-making process.

Using the bibliographic information of patents, the technological landscape and recent developments in EV sector during the recent six years were analyzed based on the statistical examination and the graph theory provided by social network analysis. Citation networks were divided into clusters, the patent assignee in each cluster were tracked, and citation networks with characteristic technology field for each cluster were analyzed. Overall structural changes of the EV industry were explored by categorizing patent assignees into four main groups, i.e. automotive OEMs, suppliers, infrastructure providers and other players, and exploring the changes in patenting activities between these groups. Analysis of patent network dynamics reveals the changes in the structure of innovation landscape within an emerging field of EVs. Expert opinion of the Finnish automaker was included in the analysis of this study. Limitations of the methodology and suggestions for further research directions are discussed.

Keywords: patent analysis, patent citation network, patent mapping, bibliographic coupling, Gephi, electric vehicle, emerging field	Published language: English

Acknowledgments

The research work described in this thesis would have not been possible without the help and guidance of a number of people to whom I would like to express my gratitude.

I am truly thankful to my supervisor, Prof. Robin Gustafsson, who accepted me into his group and suggested an interesting research topic. Thank you for the opportunity to become a member of a strategy research team and gain new knowledge and experience, for your guidance of the research work and helpful advice, for your readiness to help, discuss and connect with the experts in the field. I would like to thank my instructor of this work, M.Sc. Chris Rowell, for providing insight into the field of electric vehicles, giving useful ideas during the study and sharing his enthusiasm – an invisible key to success. A positive and friendly working environment that Robin and Chris create in the team encourages progress and desire to excel.

I would like to thank the members of the Department of Industrial Engineering and Management and the Institute of Strategy, former and present, who helped me in one way or another during my studies and thesis research work. I am grateful to Prof. Martin Meyer for his valuable input to the research methodology and the focus of this study. I want to thank Dr. Anssi Smedlund for his assistance with Gephi software at the start of the research project. Many thanks to Maria Tikka for her administrative assistance, and the researchers working at the Institute of Strategy for a friendly working environment.

This research work would have not been possible without a fruitful collaboration with the Finnish automaker. I am thankful to the company for providing an expert opinion for the results analysis and making a significant input to the research outcome. I would also like to thank the patent attorney Jonna Sahlin for her consultation on patenting-related issues.

Finally, I would like to express my deep gratitude to my family. My special heartfelt thanks to my father Yuri for his encouragement of my study progress; my dear husband Colin for his unmeasurable help and support, and our daughter Evy for making me a quick learner of certain general management skills and simply a happy mother.

Leuven, Belgium, May 2014

Marina Y. Timmermans

Table of contents

List of abbreviations

BC	Bibliographic coupling
CC	Co-citation
EPA	Environmental Protection Agency
EPO	European Patent Office
EV	Electric vehicle
ICE	Internal combustion engine
IPC	International Patent Classification
OEM	Original equipment manufacturer
PCN	Patent citation network
R&D	Research and development
USPTO	United States Patent and Trademark Office

1 Introduction and research problem

Understanding the technology landscapes and trends in innovation advancement is strategically important for the success of most companies operating in a highly competitive and constantly evolving business environment of emerging technologies. Managers need to recognize the real competition and collaboration possibilities, and formulate effective technological strategies in order to achieve commercial success. This study uses patent citation information to examine the possibilities and limitations of its use as a method to gain insight into the development of emerging technologies in support of strategic decision-making. A dynamic network view of patent citations provides a large-scale understanding of the relationships between patented innovations, revealing changes in patenting activities between technology fields and organizations. Here, the observed trends are more generally referred to as technological developments.

Bibliographic coupling (BC) is a long-established method in science mapping for the analysis of technological interconnectedness between the patents based on their citations. Two patents are bibliographically coupled if they share one or more citations, which represent their common intellectual background [1, 2]. In recent years this technique has received a growing interest due to the ability of the method to group very recent patents and obtain snapshots of early stages of the evolution of the industry [3]. In contrast, patent analysis based on co-citations tends to cluster older patent documents, since time is required for new research topics to receive enough citations [3, 4]. Therefore, BC is the method of choice in this research work. Notable advancement have been made in the BC methodology as a tool for analyzing the cognitive structure of various research fields and producing maps of research fronts and knowledge bases of research fields [1, 5]. It was shown, however, that accurate clustering of patents within the network and the interpretation of the clusters are especially important to decision-makers, in particular when used for real-world research planning and industry evaluation [4, 6]. With a significant advancement in the methodology of patent citation analysis using BC approach, very few studies focused their research results on practical applicability of such analysis as a useful input into corporate strategic planning in the field of emerging technologies.

This study examines and applies the methodology of BC for building patent citation networks (PCNs) as a tool to analyze recent developments and current stage (open options) in an emerging technological field [3, 7]. Emerging technological fields are a context in which BC is considered to be especially powerful because it is capable to show the early stages of the industry development. A number of social network analysis tools are readily available these days to assist managers in mapping and analyzing patents in a form of a network. These tools allow refocusing the researchers' efforts on data visualization and interpretation, rather than investing significant time on learning the advanced techniques, which challenges their potential applicability. Such data visualization, performed in a meaningful and relevant way, allows discovering new knowledge about the research topic of interest or verifying trends that the firm already knows or has an intuition about. However, what kind of useful information can be obtained from this type of analysis and what are the existing limitations?

This study makes two contributions. First, this research work explores the methodology of patent citation analysis based on BC approach from the viewpoint of its potential contribution in support of technological strategy formation within the field of emerging technologies. Prior research has been mainly focusing on the methodology of BC for the demonstration of the static relationship between documents at a specific point in time [8]. By looking at the dynamic relationship between patented innovations and patent assignees throughout several recent years, this study attempts to pinpoint the recent developments in the technology focus, position of key players within the field and the structural changes within the industry in general. Combining BC with the analysis of patent assignees allowed to explore the changes of patenting activities between different industry sectors.

Second, a separate attention is paid to summarizing and discussing the limitations of patent citation analysis approach based on BC. In this work I discuss the existing methodological limitations, stemming from the use of patents and patent citations as a data source; intentional limitations, originating from the chosen source of patent data and patent search strategy; and capability limitations of patent network visualization, clustering and mapping interpretation. With the focus on the methodological aspects and opportunities of patent citation analysis using BC approach within the chosen case study of an emerging technology field, this research in part reflects the expert opinion of a firm operating in this field when verifying the visual representation of patent data and pointing out analysis limitations. Practical approach for the applicability of the methodology in focus provides a valuable contribution for the future study of patent citation analysis and strategy planning. The research question of this thesis is:

> *What can a firm, operating in an emerging technological field, infer from patent citation analysis using bibliographic coupling approach and what are the limitations?*

The research question is investigated on the field of electric vehicle (EV) technology. EV is an emerging technology capable to disrupt the established technology environment of automotive manufacturers and provide opportunities for new entrants into the sector [9]. This study provides a dynamic network view of the field during the recent years through PCNs. The analysis gives insight into the technological environment and development trends in the area. The results of this work facilitate further exploration and application of patent mapping based on BC for technology analysis and strategic planning.

The remainder of the thesis is organized in the following way. Chapter 2 reviews the literature on methodology of patent mapping with the focus on BC approach and its potential applications. Chapter 3 presents the details of the research design and methodology. Results of the thesis on potential applicability of the methodology and its limitations are described and discussed in Chapter 4, followed by the conclusions and avenues for future research in Chapter 5.

2 Literature review

2.1 Patents and patent citations as a data source

Patents have long been used as one of the sources of information for the analysis of the technological development of a particular industry [10-13]. A patent is “a temporary legal monopoly granted to inventors for the commercial use of an invention” [14]. For companies patents are vital to protect proprietary technologies and core business concepts [15]. Therefore, often patents represent the main outputs of research and development (R&D) activities [16], and allow to measure the technology assets of a firm for the analysis of its strength and weakness [17] as well as gain information about competition trends [18]. Patent records contain broad information on each patented invention, including technological area of the invention, the inventor, the organization to which the inventor assigns the patent property right (typically an employer), and their geographical location. That is why, patent information is a very useful tool to understand industrial developments, revealing all about the most recent and commercially feasible technologies [19].

Patents contain citations to previous patents which allow to establish relations between patented innovations [20]. Patent citations identify prior art or the body of knowledge publicly known before the filing date of a patent application [21]. This may include other existing patents, scientific journals, conference papers, books or other reference works. They help to define the property right granted by the patent, since a patent cannot have a claim on the previously existing knowledge. The patent applicant of the patent is required to disclose any knowledge of the prior art. In the US non-compliance with the so-called “duty of disclosure” requirement can lead to patent invalidation [22]. The inventor typically identifies the cited references with the help of the inventor’s patent attorney, but the final decision on patents to be cited lays on the patent examiner, who is expected to be an expert in the area to determine the relevant prior art and make the citations complete.

Patents are considered to be “a proxy for inventive output”, and patent citations are “a proxy for knowledge flows or knowledge impacts” [14]. Patents can be linked together based on citations and displayed in the form of a network, reflecting related innovative ideas. Therefore, patents and patent citations are important indicators of technology [12, 23, 24]. For a firm patents are an essential source of technical and commercial knowledge, which is accentuated in emerging technological fields [2, 25]. Patent data may reveal existing collaborative or competitive network of relations among companies within a certain field, which represents a valuable information when developing strategies for intellectual property, R&D and marketing activities [15]. Besides patents and patent citation analysis provide a useful tool to observe innovation advancement and identify the emergence of novel technological fields [3, 7].

2.2 Patent citation analysis

The influence of firm’s context on its behavior and performance has long been acknowledged in strategy and organizational theory research works [26]. When a firm is planning to enter a new market or is exploring the direction of its R&D activities, knowing the key players in the area, most prominent inventions, trends and changes in technology developments, as well as emergent technological trajectories within the field is strategically important. R&D investments are often risky because innovation is unpredictable [27]. The biggest challenge for a company is to understand threats and anticipate opportunities from the emergence of new technologies in order to avoid false investment decisions and reduce risks. It is also difficult for companies to recognize real competitor or cooperators, since the firms of different industries can enter the market becoming competitors [6].

Patents represent publicly available source of data on industrial R&D and inventive activities. Patent documents also contain citations to prior art in the respective technological field, displayed on its front page as 'references cited' [2, 22]. After the first findings that put forward patent citations as indicators of the importance or value of innovations [17, 28], significant research work on the usefulness of patent citation analysis from a technological and an economical point of view was further undertaken [2, 14, 29-31]. In analogy to scientific citations, an innovation may be partly based on an earlier patented innovation, in which case an inventor has the duty to disclose all the known 'prior art' related to the invention. Therefore, the underlying idea of patent citation research is that the more citations a patent receives, the more significant it is in terms of technological impact and innovative development [2]. The inventor benefits from the work that was achieved before, and at the same time contributes to the base of knowledge for building future inventions [20]. Historically, the most common methodology of patent analysis was based on simply counting and statistically comparing the amount of patents belonging to a technological field or a firm [32]. Patent citations have been widely used in the innovation literature for measuring patent quality [28, 30, 33], analyzing knowledge flows and spillovers [34-38], and exploring strategic behavior of a firm [39, 40]. There are very few studies that discuss direct applicability of patent citation analysis to support business strategy formation of a firm within an emerging industry. Implication of the strategic perspective is often mentioned in the conclusions of these studies, suggesting the applicability of patent analysis for strategic decision-making, i.e. for the analyzes of changes in technology and prediction of emerging technologies [16, 25, 27] and establishing technological position and relationship of a firm so to recognize competitors and identify potential partnership [6, 41, 42].

Patent citation analysis has taken advantage of the recent developments of social network analysis tools. Social network analysis is a quantitative technique based on graph theory which views social relationships and explores the interactions among the actors within the social network [43]. Construction of PCNs using tools from social network analysis allows to obtain a global view on technological innovation, revealing the relationships between patents as separate pieces of technical knowledge, or connections among technology field players as patent holders. PCN is a collection of nodes connected by edges, where the nodes represent the patents and the edges represent links between them based on either citations from one patent to another or common citations. While patent portfolio of a firm reveals its capabilities, a map of patent citations have been used to identify the important players in technology fields and understand the relationship of firms, including competition and cooperation [44, 45], or map technological trajectories and patterns of technology diffusion [41, 46, 47]. The development of the techniques for the visualization and analysis of social networks opened up wide possibilities for the interpretation of patent citation data and improvement of current visualization techniques in patent analysis in order to gain a better understanding of innovation and knowledge flow processes on the basis of patent information [45].

2.3 Bibliographic coupling method

PCNs as a tool to support strategic decision-making is gaining recent research attention. Co-citation (CC) and BC have been the most widely used techniques to map scientific documents based on their citations in order to analyze technological trends.

BC method or reference co-occurance was proposed by Kessler in the 60s as a method to group technical and scientific documents, according to which "a number of scientific papers bear a meaningful relation to each other (they are coupled) when they have one or more references in common" [48-50]. In his further publications, Kessler tested the method's generalizability to larger populations and for different disciplines, demonstrating the existence of subject relatedness between bibliographically coupled papers [49]. Moreover, Kessler demonstrated that the BC method of grouping documents "operates both on the past and future literature, as measured from time of document P_0" [51]. Therefore, this method is useful to observe the life span of a given domain.

In earlier works BC arose interest as a method allowing to group documents on the basis of use rather than content, showing that "it is possible to group papers into sub-groups mechanically, implying no knowledge of science or judgment of content" [48]. Instead of counting the amount of common references, In 1966 Cleverdon suggested the improvement to the BC method by using another form of coupling measure between the documents, i.e. the proportion of the coupled references over the total number of references in the document [52]. Further theoretical studies on BC introduced a measure of coupling strength between documents described by the angle between vectors representing the documents (0 < ϴ < 90°) [53], which is the graphical interpretation of the so-called Salton's cosine measure [54, 55]. However, after its first introduction, the validity of BC method was questioned. For instance, Weinberg published a review on BC method in 1974, where she underlined a current challenge of the method's applicability for a more complex interdisciplinary environment [56]. For some time the preference of the scientific community was shifted to other citation-based science mapping approaches largely due to the lack of advanced computational resources to test larger populations so to demonstrate the validity of BC [54]. In 1984, after 20 years from the first Kessler's report, more studies on heterogeneous and larger scale samples were performed once again revealing a strong subject relatedness of BC groups [57].

Since then BC was demonstrated as a powerful method and information seeking tool in many contexts, allowing to gain insights into the structure of research fronts[1] [54, 58, 59], identify potential collaboration partnership [60], or discover the intellectual structure of a certain discipline [8]. For example, Morris *et al.* showed that clustering documents into research fronts using BC is a valuable tool to help experts in a technical field to identify specific research areas and follow their activities over time [59]. Later, comparing different citation-based approaches, i.e. CC analysis, BC and direct citation, on a large scale (data set of 2,153,769 recent articles from the biomedical literature from 2004 to 2008), Boyack and Klavans concluded that BC most accurately represents the research front on a large scale [4]. More advanced visualization techniques of BC publications were proposed by Schiebel in his work on representing research fronts and knowledge bases with two or three dimensional graphics [5]. BC method was also shown by Nicolaisen *et al.* to be a promising technique to measure the level of consensus in science, basing the assumption

[1] Research fronts represent documents that consistently cite a fixed, time invariant set of base documents [59].

that if a field has a high level of intellectual agreement, it is characterized by a high number of shared references [61]. A possible refinement of BC by incorporating the age of common references into the assessment of document relatedness was recently proposed by Soós [62]. A significant advantage of BC method is its capability to capture the most recent trends and early developmental stages as discussed in the works on BC by Glänzel [63, 64]. This is enabled by a capability to establish BC between two documents immediately after the publication of the second document of a pair [59]. Another interesting approach demonstrated with BC is to construct time lines to display relations among research topics over time [59]. Few recent studies proposed a BC-based methodology to track the technology front evolution through the consequent time windows and map out the technology developmental paths over time, thus revealing the dynamic nature of real-world environment [65]. Drawing on the advantages of BC methodology to map the early stages of a speciality's evolution, Kuusi *et al.* have successfully applied this method in the foresight context, i.e. for exposing expectation and anticipations about progress and breakthroughs in emerging fields of technology [3]. They demonstrated that BC can be a useful technique to identify technological guiding images of the emerging field that provide a basis for different professionals and disciplines to work in the same direction [3].

2.4 Bibliographic coupling vs. co-citation approach

BC and CC analysis are both citation-based science mapping approaches that occur when two documents make a reference to a third document. The two methods are often confused or wrongly considered to be equivalent, even though from the technical point of view the methods are very different [63]. CC method for document coupling is defined as the frequency with which two documents are cited together [66]. BC links documents that cite the same references, while CC links documents that are cited together as references [59]. A schematic illustration comparing the two methods is shown in Figure 1. Both methods are assumed to produce groups of related documents that represent the same or similar research topics [64]. In a longitudinal dataset BC is able to group very recent papers but fewer of the older ones, while CC, on the opposite, clusters older papers but is not able to cluster the very recent papers that have not been cited yet [4]. CC relationship between a pair of documents varies with time, approaching a stable value over a certain period of time [59]. Therefore, accurate CC mapping contains a time delay which is needed for a highly cited publication on a new research topic to gain a critical mass of citations [3]. BC, on the other hand, is accessible as soon as the other document of the pair is published and this relation does not change over time [59]. That is why BC is typically a method of choice in the context of recent developments and future trends due to its capability to capture the early stages of technology development and provide more current information about the published invention.

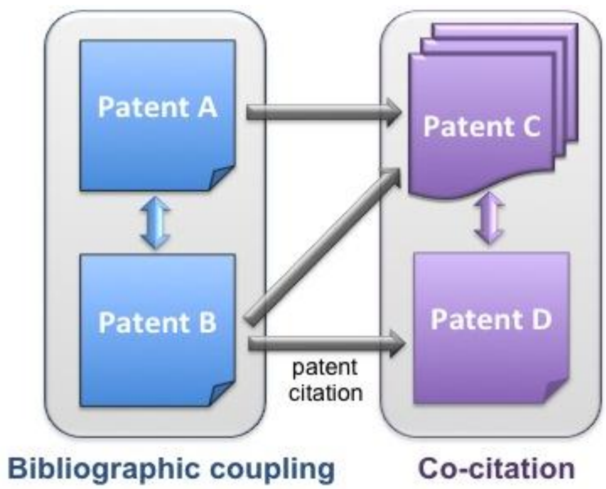


Figure 1: Bibliographic coupling vs. co-citation patent mapping approaches.

BC and CC analysis could be applied together in a hybrid approach. Von Wartburg *et al.* were the first to demonstrate a multi-sage citation analysis approach for understanding technological developments [2]. They showed the usefulness of BC approach specifically for measuring ‘shared specialization’ between patents (or how much technologies they have in common) and observing the fields of technology that show a pattern of cumulative innovation (when building on prior knowledge is important for the success of new inventions) [2]. BC is a preferred additional analysis technique to assess the degree of relevance of citation pairs, filter irrelevant citation pairs and identify missing relevant patent links in the citation network in order to provide a more comprehensive view on the relationship among patents [67-69]. After mapping of document networks based on their BC, the next essential analysis step is identification of document clusters that share common technological features or deal with a similar research line. A number of studies proposed the use of cluster analytical methods combined with BC to ensure the coherent grouping of related documents [54, 70].

Even though the potential value of PCNs as a tool for strategic decision-making by company CEOs, CTOs, R&D or IP managers was widely recognized in scientific literature, and BC as an analysis technique was demonstrated to be an advantageous method for the construction of the most recent maps based on citations, not much focus in literature was directed to the practical applicability of patent citation mapping using BC approach to support effective communication of technology planning and decision making regarding firm strategy in the context of an emerging technology field. Application of social analysis tools allows visual representation of patent data and its further analysis. However, such data presentation in a visual way needs to be meaningful, relevant and accurate for it to be useful for the firm. This study attempts to make a contribution to this line of research efforts of exploring the BC analysis approach.

3 Research methodology

3.1 The field of electric vehicle technology as a case study

The emerging electric vehicle (EV) technology was chosen as the case study of this research work. EV technology is undergoing a rapid growth worldwide, representing a major pathway towards the reduction of petroleum dependence and environmental footprint of transportation, while at the same time providing a positive effect on the economic growth and technological competitiveness of the vehicle market [71]. EVs in this work include battery electric vehicles, hybrid gasoline electric vehicles and plug-in hybrid electric vehicles which are expected to replace conventional internal combustion engine (ICE) vehicles in the future [72]. According to IDTechEx, consulting and research organization, the total market for hybrid and pure electric vehicles is expected to grow from 38.8 million vehicles in 2013 to 116 million vehicles in 2023 [73]. Even though the cost of EVs is higher in comparison with ICE vehicles of a similar functionality due to the battery cost, the purchase price of EVs is expected to be compensated by much lower consumption costs, i.e. gasoline costs vs electricity costs, and incentivized by government regulations related to avoiding city congestion or emission charges [74]. The broad-scale introduction and adoption of EVs in the coming future will bring significant changes to the existing transportation sector and related technological and infrastructure systems. Emergent pure battery EVs are relatively simple in their vehicle structure and driving capability as compared to conventional ICE vehicles, and as a result, less adjustments or collaboration are required between components and vehicles in the phase of R&D process [75]. This could bring crucial structural changes in the automotive industry from largely vertical, when key automakers manage their products across the value chain including engineering, raw material procurement, assembly and distribution, to predominantly horizontal with independent industry players, as demonstrated in Figure 2 [75]. Therefore, further development of EV technology will lower the barriers for new entrants to the automotive industry and close cooperation with firms and inventors beyond the existing product development networks of the firms becomes essential [9].

In the view of these potential substantial changes, it becomes strategically important for firms, operating in this environment, to understand the current technological development and anticipate future events. Here I attempt to examine the possibility for analyzing the technology development map by means of PCNs based on BC approach.

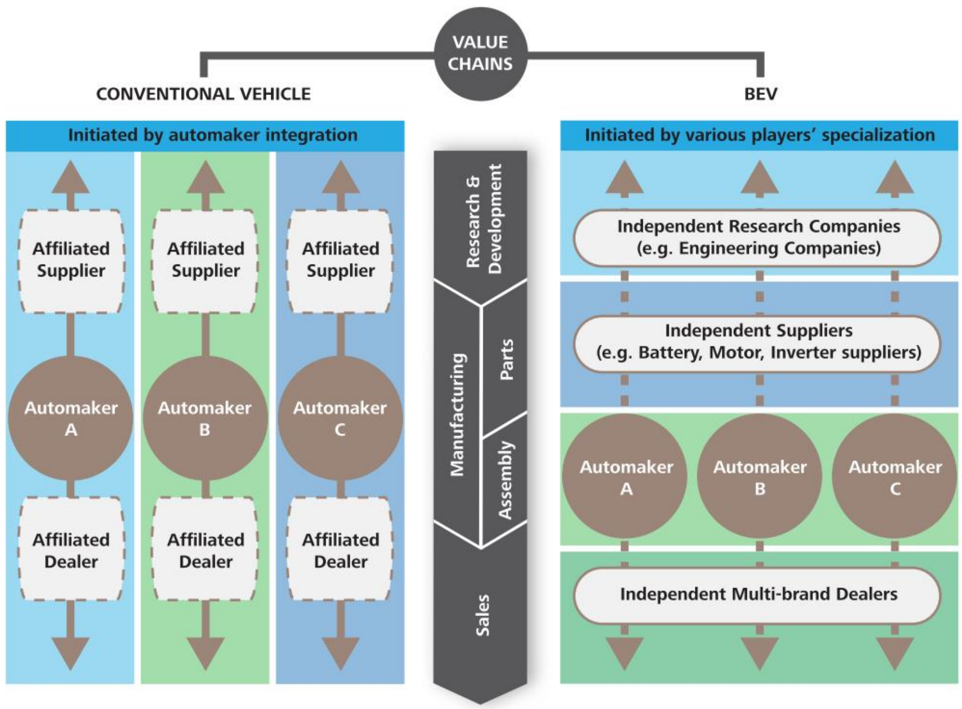


Figure 2. Structural change in value chain caused by battery EV spread [75].

3.2 Overview of the data and methodology

In this exploratory patent data analysis study the tools of patent analysis and social network analysis we combined in order to track the development of EV technology. The research methodology of this study consists of three steps: data acquisition, network construction, and network analysis. The research framework is shown in Figure 3.

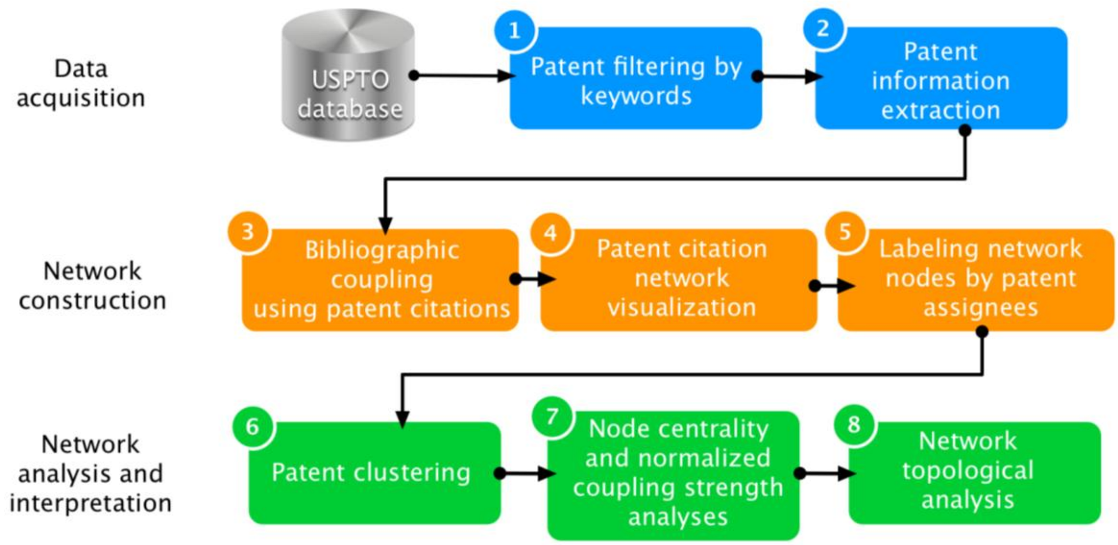


Figure 3. The research framework of this study.

Using the bibliographic information of patents I analyzed the recent technological developments in EV sector based on statistical examination and the graph theory provided by social network analysis. Thus, analyzing the changes in the topological structure of the patent network over time, i.e. arrangement of various nodes of the networks and links between them, I speculate on the changes in the innovation landscape of the industry. This study aims to contribute to the discussion on the methodology and applicability of PCNs for mapping the technological development within the emerging field. The technology development over time during the recent years are analyzed by identifying the major technology clusters and research fronts, determine the main players within the field and follow industry's overall structural changes. Limitations of the methodology and suggestions for further development directions are also discussed. Expert opinion of a Finnish automaker was included in the analysis. In this work we focus on the last six years of patent filing date as the time window of our longitudinal dataset, as discussed further in this chapter. Examination of earlier patents within the EV field by using the same network visualization and analysis methodology could provide interesting insights on the history of the EV sector development.

The method of PCN analysis combines both objective and subjective elements. Linking patents using BC approach represent objective quantitative research methodology. Even though clustering method is based on a mathematical algorithm, interpretation of clusters is more or less subjective. Here the input of technology experts and their opinion to complement the empirical findings is essential, as pointed out in the study by Porter and Newman [76]. The limitations of the analysis are discussed in detail in a separate section (section 4.2).

3.2.1 Data acquisition

In this section I will discuss the patent search strategy used in this study with a detailed explanation for the reasons behind. The original broad data set, extracted at the beginning of the study, provided a general overview of the EV industry development in terms of classification of patents into specific technological fields, distribution of patents by country and the change of patenting activity over the years. This analysis allowed us to choose a reasonable time frame of the data sample to narrow it down for the further construction of PCNs.

Patent data for the case study of this research was extracted from the United States Patent and Trademark Office (USPTO) online database (PatFT) using the keywords search in patent titles or abstracts. USPTO PatFT contains full-text for patents granted from 1976 to the present day. Keyword search in title and abstract is a widely used tool for patent extraction, e.g. [3, 16]. A total of 3683 patents granted between 1976 and 2013 that contain the terms “electric vehicle(s)” or “hybrid electric vehicle(s)” in the title or in the abstract were originally collected[2]. Only utility patents also known as ‘patents for inventions’ have been used for our analysis since these types of patents are perceived to have a direct relationship with the technological activity [77].

[2] The patent data download was carried out on 23.10.2013, which corresponds to the most recent granted patent in the database.

The following information was extracted for each patent document: patent number, patent title and abstract, application year and patent grant year, inventor(s) and assignee(s) of the patent and their geographic location, international patent classification (IPC) and US classification codes of the patent as well as patent references. All of this information was useful for the patent analysis of this study. For the cases where the information on the patent assignee was absent (7% of the whole data set), the patent was assumed to be owned by the individual inventors themselves [78]. For the BC analysis I used the references to ‘other U.S. Patent documents’, referred to as citations. Patents might also contain references to other publications, including mainly technical journals as well as textbooks, handbooks and other reference works [22]. Analyzing these non-patent references is an interesting study on its own which allows examining the relationship between science and technological development [79]. However, a number of studies discuss that the cited scientific literature rarely represents the source of knowledge leading to the invention, and therefore, there is often no direct relationship between the citing patent and the cited paper [22, 80]. In this study I will focus on constructing the citation networks only between patents.

USPTO database was chosen as a patent data source in this research work. The data source was chosen for a number of reasons. First, the USA is the largest market for high-tech products also for most Asian and other manufacturing companies. Even though often the majority of technology production remains close to the home base [81], the amount of international companies patenting their inventions in the USA is larger than that of other nations. USPTO patents are considered to be the most valuable and reliable due to the competitiveness of the US market and desire of firms to secure their intellectual property rights in this largest technological market [82]. Statistics from the USPTO database for patents within the EV field shows that more than half of all US patents are issued to foreign entities (Figure 4).

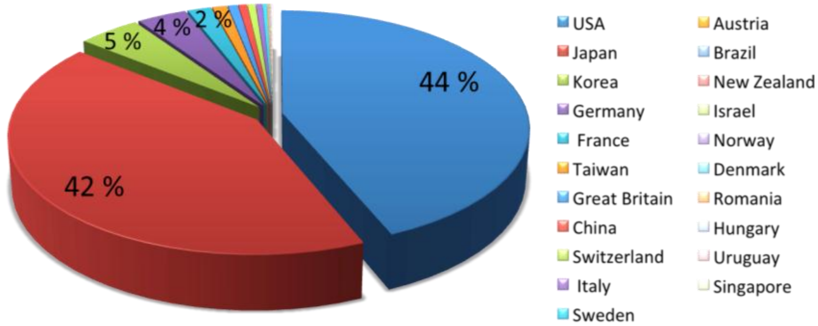


Figure 4. Distribution of patent assignees by country within the EV industry based on the USPTO database extracted using keywords search (patent grant years from 1976 to 2013).

Second, USPTO provides a freely accessible database for patents and their citations, which is the main tool for the analysis in this study. European Patent Office (EPO) online database (Espacenet) was also considered as an alternative source of patent data. This is also a widely used source of patents and patent applications, providing access to national and worldwide database. However, the free version of this database restricts the data retrieval to only the first 500 search results and has limitations for obtaining patent citations. An interesting future study could involve worldwide data analysis in the field of EV technology and examining the patenting activity of the major players in this area worldwide.

However, one needs to consider certain limitations stemming from the difference in patenting practices between different territories worldwide which might have an effect on macro-interpretation of patent citations [83]. US-originated patents are known to have a higher citation frequencies, and therefore a higher amount of linkages to other patents, in contrast to European-originated ones due to the differences in the examination of patent applications [22]. A law in the USA requires the applicant of the patent to make a full and fair disclosure of the prior art, and if this were not the case, the validity of the patent is challenged [9]. This duty of disclosure by law imposes a pressure on the patent applicant to make a larger amount of references, providing a detailed reasoning about patentability of each claim [22]. Due to that some of the citations might appear to have a less strong relevance to the core of the patented invention in the attempt to ensure that all angles are covered, and certain bias exists due to the significant contribution of the applicants themselves to generating the list of references [9, 22]. These limitations need to be taken into account when interpreting the results. Hence, owing to the greater number of references USPTO data is especially resourceful for building a wide cognitive web of connections between fields and analyze technological spillovers [22]. Considering the differences in patenting systems across nations, all of the sample's data was extracted from USPTO to ensure consistency and reliability of the results.

All patents have patent classification codes manually-assigned by the examiners of the national patent office for ordering patents based on their technical content. Patent search within certain classification codes is an alternative strategy for searching patent documents relevant to the field of interest or reinforcing the search together with the key words, e.g. [72, 84]. For instance, International Patent Classification (IPC) is utilized to order patents worldwide and classifies each technological field into five hierarchical levels: section, class, subclass, main group and subgroup, with about 70,000 categories [85].

Figure 5 shows that patents within the EV field, retrieved via the key word search, mainly belong to section B (52 %) which represents 'performing operations and transporting' technological area and section H (27%), which stands for topics related to 'electricity'. Among other significant categories are section F (13 %), 'mechanical engineering, lighting, heating', and section G (7%), 'physics'. Each section is further divided into classes, classes into sub-classes, sub-classes into groups etc., as shown in Figure 5 for section B as an example (see Appendix A1 for a similar analysis of second main section, H). Figure 5 and Figure A1 demonstrate that patent classification codes may vary significantly within a field, especially when dealing with emerging technologies. Accurate choice of the level of analysis within the hierarchy of the patent classification system (e.g. within a section, class, sub-class, group etc.) is crucial for the validity of studies that use classification-based data search for the analysis of certain technology categories. Therefore, the chosen patent search strategy based on the general keywords of the case technology allows to obtain a broad data sample and to avoid a possible bias when restricting the search within specific patent classification codes. However, there may exista possible bias around patent families or very closely related inventions, which has to be considered in the network analysis.

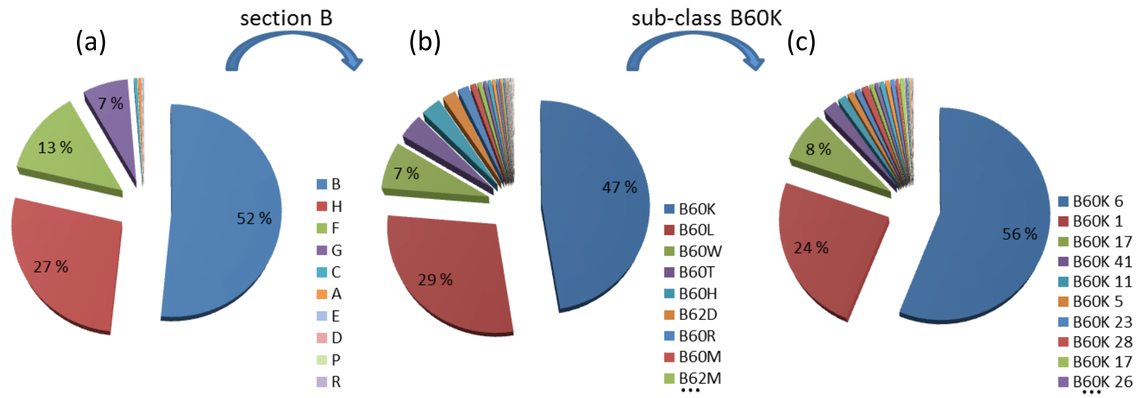


Figure 5. The percentage share of (a) patents in EV technology; (b) patent subclasses within section B; (c) patent groups within sub-class B60K. The main subclasses of section B ("Operations and Transporting") embrace arrangement or mounting of propulsion units or transmissions, prime movers, auxiliary drives, instrumentation or dashboards for vehicles (sub-class B60K, 47%), and the leading group of sub-class B60K belongs to B60K6 (56%)-arrangement or mounting of plural diverse prime-movers for mutual or common propulsion, e.g. hybrid propulsion systems comprising electric motors and internal combustion engines.

Expert judgment within the automotive sector was employed as an additional data source to address the research problem and strengthen the interpretation of the analysis. Several representatives from both strategic and engineering sides of a Finnish small-sized firm, operating in the EV field, were interviewed for verification of technology cluster subject coherence and cluster labeling. Further in this report I refer to them as EV field experts. Information related to grouping patent assignees into different EV industry players was gathered using secondary data from web-based resources.

3.2.2 Dating of patents and application-grant lag

Each patent document contains information about the patent application year (filing year) and the year the patent was granted. Figure 6 demonstrates the annual number of patents granted by filing year (Figure 6 (a)) and by grant year (Figure 6 (b)).

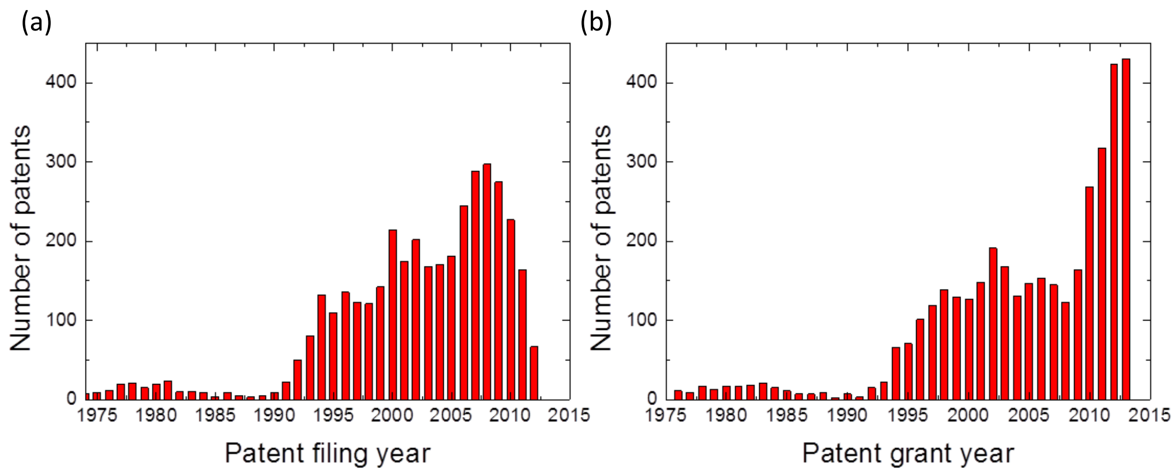


Figure 6. Number of patents by (a) filing year and (b) grant year.

There is typically a lag between the time the patent was filed and granted, which usually depends on the review process at the patent office and takes on average about two-three years [20, 21]. As shown in Table 1 application-grant lag for patents within the EV industry is about 2.4 years with significant variance (for the patents applied since the year 1989). From this sample the percentage of patents granted after 2 years since the filing date is 37 %, after 3 years is 69 % and after 4 years is 88%. Changes in the mode of operation of the patent office during the past decade introduce the variation of the average application-grant lag which in fact is not related to the timing of the invention itself [20]. Therefore, in order to avoid the possible randomness of data analysis using patents dated by grant year, I use patent filing year when dating patents. Filing date of a patent application is closer to the actual time of the invention, since the inventors have a strong motivation to file a patent application as soon as possible to be entitled for the patented innovation in accordance with “first-inventor-to-file” system [21]. There always exists a lag in time between the patent filing date and the time when the invention was made. However, typically patents are granted well before the product introduction to the market and thus patent filing date is a good indicator for the timing of the invention [86].

Table 1: Application-grant lag distribution by 3-year periods in the field of EV.

	Application years							
	1998-1991	**1992-1994**	**1995-1997**	**1998-2000**	**2001-2003**	**2004-2006**	**2007-2009**	**2010-2012**
Lag	Distribution of lags (in %)							
Up to 1 year	8,6	8,2	2,5	6,2	6,3	2,9	1,3	14,9
From 1 to 2 years	54,3	50,8	49,6	53,4	40,7	14,8	8,6	34,0
From 2 to 3 years	31,4	34,4	38,3	27,9	31,7	30,6	28,5	43,3
From 3 to 4 years	5,7	5,5	6,9	8,3	13,8	25,8	35,4	7,8
From 4 to 5 years		0,8	1,9	2,1	4,7	15,5	21,9	
From 5 to 6 years		0,4	0,3	1,1	1,9	6,1	3,7	
From 6 to 7 years			0,6	0,6	0,7	4,0	0,6	
From 7 to 8 years				0,4	0,2	0,3		
Total	100,	100,	100,	100,	100,	100,	100,	100,
Number of patents	36	257	264	471	537	595	846	398
Mean and standard deviation of the lag (years)								
Mean	**1,91**	**1,98**	**2,15**	**2,07**	**2,37**	**3,27**	**3,36**	**1,99**
S.d.	8,22	8,38	9,35	11,71	13,48	16,60	12,19	9,55

In our study only those patents that have been granted are considered. Patent applications that are still in the review process by the patent office is a possible data source to assess the recent innovative output within a certain technology area being the first published detailed information about the invention [72]. There exists a separate USPTO online database which contains information on patent applications (AppFT) starting from the year 2001. Since the end of 2000 a new US patent mandates the publication of all patent application 18 months after it is filed [87]. However, even though patent applications provide access to the most recent data on technological inventions, one should be mindful about certain limitations of this data. First of all, not all patents that are in the application process are eventually going to be granted by the patent offices. Besides, firms may try to file patent applications without necessarily intending to get it granted with the strategic objective to artificially generate prior art and block other patent applications, so to compromise the patentability of inventions made by rivals and impose barriers to entry into markets and technologies [88]. Taking these limitations into account, analyses of patent applications could be a subject of our further study especially when focusing on the examination of emerging technological trends and future developments.

It can be observed from Figure 6 and Table 1 that the number of patents within the EV field has increased rapidly, showing a drastic leap from around 36 patents filed in between 1989 and 1991 to 846 patents filed between 2007 and 2009. Here it is important to consider the truncation problem, which affects the patent data of recent years. This is associated with the fact that the data include patents granted up to the date the patent search was conducted (here, October 2013), and, therefore, contain only those patents that were granted relatively fast. A number of patents filed in that period that will be granted later is missing from the dataset. That is why we observe a significant drop in the amount of patents filed in recent years as a consequence of the truncation problem. This artifact imposes a limitation on the data analysis of the most recent patents dated by filing year.

Earlier research proposed that patents carry the most value within around 5 years of their issue date which varies with respect to the technology area [89]. Since in this study I am interested to provide a glimpse on the direction of EV technology development in recent years, I focus on the last 6 years of patent filing date, i.e. 2007-2012, for building PCNs. As shown in Appendix 2, the average time gap between the application date and the grant date of patents filed within that period is around 3 years with about 99,6 % patents granted 6 years after filing. Therefore, the time frame of the data sample is limited to the years 2007-2012[3]. This is the period of a significant growth in the patenting activity within the EV field as demonstrated in Figure 6. Examining earlier patents using the same methodology could further provide interesting insights on the historic developments within the field of EV technology. From now on patent filing year is used as the time placer for patents.

3.2.3 Patent citation network construction and analysis

In this study networks are visualized and analyzed through the use of the social network analysis software Gephi [90]. Gephi is an open source software for network

[3] The most recent patent filing date in the database is 19.12.2012, which corresponds to the patent granted in 2013.

exploration and manipulation [91, 92]. BC analysis is employed to plot PCNs within the EV technology field. This method allows to graphically display technology clusters on the level of individual patents or their assignees and the association among them. Two or more patents are bibliographically coupled if they cite the same reference. The more references the patents share, the more common technical background for the development these patents are based on [49]. In other words, the greater the BC strength of patents, the more relevant they are expected to be [19]. In this study I neglect any relationship between patents containing more than one stage of BC, due to the complexity of such relationships and a possibility of obtaining misleading information from linking patents that might be only loosely related to the core of the patented invention, as discussed in the earlier study by Wartburg et al. [2]. Similarly, coupling patents based on the shared classification codes to which they are assigned by patent examiners (so called, co-classification analysis) is another useful approach to be considered for the future study when analyzing technological relationships between patents within a broader technological area [7, 93].

PCN is a two-dimensional graph that visually displays the relationship between individual patents formed via BC [19]. It consists of nodes that represent patents and edges that link patents based on their BC. The distance between the nodes on the map represents the association between the patents expressed through their BC, with shorter distances indicating a closer relationships and thus a closer technological field, and vice versa. The patent map essentially illustrates the cognitive structure of the field [7]. The thickness of the edges indicates the strength of the BC between the patents. The thicker the connecting link between patents within the network, the higher is the relatedness of technological features among them. BC links patents with common references and does not reflect the direction of relationship between the patents; therefore from a graph theoretical viewpoint the edges or links between patents are undirected.

However, BC strength is considered to be not an optimal measure of document similarity [54]. It is important to consider the coupling strength between two documents as well as the combined length of the reference lists of both documents [57]. Therefore, the coupling strength of the document pairs needs to be normalized. I use Jaccard index as a similarity measure in PCNs [94], which is applied and defined as:

$$NCS = \frac{n(A \cap B)}{n\ (A \cup B)}\ ,$$

where NCS is the normalized coupling strength between two patents *a* and *b*, A is the number of references in patent *a*, and B is the number of references in patent *b*. The Jaccard index is calculated as the size of the intersection (or the number of common references to both patents *a* and *b*) divided by the size of the union of the two patents [94]. The value of NCS ranges from 0 to 1. The coupling strength of the patent documents can also be normalized based on Salton's cosine measure, which was shown to be monotonic to Jaccard index [95, 96] and yield a numerical value that is twice Jaccard index [97]. Table 2 demonstrates the average Jaccard index values per PCN of this work (*NCS*), as well as the average BC strength ($n(A \cap B)$) and average size of the union ($n(A \cup B)$).

Table 2: Average values of bibliographic coupling strength, size of the union and normalized bibliographic coupling strength of patent pairs for each network within the last six years of patent filing years in EV field.

	Application years		
	2007-2008	**2009-2010**	**2011-2012**
Average BC strength, *n(A∩B)*	2,60	1,97	4,63
S.d.	11,67	4,55	18,01
Average size of the union, *n(A∪B)*	92,38	78,12	112,55
S.d.	112,81	80,86	90,81
Average Jaccard index (NCS)	**0,045**	**0,063**	**0,079**
S.d.	0,101	0,139	0,205

Social network analysis provides a number of approaches in order to evaluate the position of individual actors within a network [98]. Centrality is a fundamental concept in network analysis [99]. In our study eigenvector centrality measure was used for ranking the size of each network node. This measure indicates the importance of the patent within the network and allows to pinpoint the main actors within patent clusters. The largest nodes ranked by eigenvector centrality measure represent patents that have the most connections to the other patents which in turn have the most connections the other nodes within the network [100]. Therefore, eigenvector centrality measures the importance of a node for the total network [101]. In contrast to measuring how well the node in the network is connected, betweenness centrality allows to measure how much a node is positioned between the densely connected clusters [102]. This type of centrality measure was shown to be useful when focusing the study on the analysis of links between different patent technology clusters and identifying patents that play a key role in the technology knowledge transfer (so-called small world phenomenon in the PCN) [103].

With the aim to gather insight on the direction of EV technology development and pinpoint the main technology trends, I grouped the data set of patents filed in the recent six years into three separate networks, each containing patents filed over the period of two adjacent years between 2007 and 2012. This allows to obtain information on the dynamics of technology development within the field as well as reduces the amount of patents per network for a better visibility of main technology clusters at the patent level. Further analysis could widen the sample scope to also account for earlier years within the EV industry. For the purpose of clarity of visualization only patent documents with 10 or more coupling links are presented in PCNs for the years of 2007-2008 and 2009-2010. Additionally a 'Giant Component' topology filter was applied to display only the main largest component of the network, removing the isolates and small disconnected clusters[4] [104].

Direct representation of the data on linking patents as a graph using Gephi presents randomly aligned interconnected nodes in a two-dimensional space, providing no clear idea about the structure of the technology field. Therefore, 'ForceAtlas2' algorithm, derived from force-layout algorithm for graph clustering was applied using Gephi software [105]. This algorithm treats the network nodes as a system of interacting particles, which repel each other like magnets. Network ties or edges attract connected nodes to each other, like springs, taking into account the weight of the edges (in our case edge weight

[4] The percentage of nodes (patents) connected to the main component is 97.3 % in 2007-2008, 90.8 % in 2009-2010 and 91.6 % in 2011-2012.

stands for the NCS of BC pairs). Thus, BC patent pairs have a stronger attraction in case of larger NCS between them [101]. This algorithm produces a more readable representation of the network structure, reducing the distance between connected nodes and maximizing the distance between unconnected nodes [104]. Overall, when using 'ForceAtlas2' algorithm, most connected nodes (or hubs) are pushed away from each other, while the nodes that are connected to the hubs are aligned in clusters around them [106]. Additionally a 'Label Adjust' algorithm was applied where the node size was added to the repulsion to avoid visual overlapping of nodes.

In order to analyze the structure of the network for the discussion of technology development paths in the field over time, I cluster the PCN applying community detection mechanism employed by Gephi [107, 108]. This method for the community structure extraction is based on modularity optimization, where the nodes more densely connected together belong to the same community [108]. In our study I first allow for randomizing the algorithm for a better network decomposition, resulting in a higher modularity score [109]. Due to a larger size data set, the resolution for community detection was set to 5, providing a more general level for the analysis of major developments in the field [110]. The use of edge weights was set on. The community detection algorithm creates 'Modularity Class' value for each node within the network, which can be used to calorize communities of highly interconnected nodes [109].

The cluster is assumed to contain patents that share similarity of technological features. The main technology topics that each cluster deals with were identified based on patents content, US patent classification and statistical analysis of the most frequent words in patent titles and abstracts. Wordle is utilized as a research tool for the analysis of the most frequently used words in patent titles and abstracts [111]. The word-cloud analysis was shown to be a useful adjunct tool for preliminary analysis and validation of findings to aid educational research [112]. It allows to obtain a quick visualization of certain general patterns in text [112], however since this tool neglects the semantics of the words and also the phrases the words are composed of, more sophisticated text mining techniques are recommended in cases when frequency analysis is used as a standalong research tool, e.g. using TF-IDF (term frequency–inverse document frequency) analysis method [70]. Common words (e.g. the, at, with, to etc.) as well as words 'electric' and 'vehicle' and their variations, used as key search terms for data extraction, were disregarded for word frequency analysis. For labeling technology clusters I have drawn upon the US patent classification system instead of IPC classes. The main reasoning behind is that US patent classification is perceived to be more accurate in USPTO database, since US patent examiners classify patents using US system first, which is often followed by generating IPC codes using automated concordance tool that might not always be optimal [113]. Expert comments in the sense-making of technological clusters of patents and their subject coherence is known to be valuable for the validation of the method's relevance [54]. The expert comments on the validation and improvement of cluster labeling were taken into account in the discussion of results and avenues for future research forthcoming from this thesis work.

Nodes within the network were also partitioned by patent assignees or groups of assignees, as shown further in the Results section of this thesis. Labeling of nodes could be varied depending on the purpose of analysis, e.g. by patent number, patent assignee etc.

4 Results and discussion

4.1 Patent citation network and technology clusters

This study demonstrates an example of patent citation mapping, aggregating patents in the emerging field of EVs by the ‘shared-ness’ of their technological features, as suggested by BC methodology [2]. With the aim to capture the recent developments within the field, I constructed a PCN using BC analysis of the patents that were filed within the last 6 years. This time period was chosen taking into account the patent application-grant time lag as discussed in detail in the Methodology section. From the time period of 6 recent years, i.e. between 2007 to 2012 of patent filing years, I retrieved a total of 1244 granted utility patents. Linking these patents by means of BC resulted in a network that consisted of 1118 coupling patent pairs with 11,411 links between all of them. The PCN containing this large amount of patent data can easily be processed and visualized using Gephi social analysis tool as show in Appendix 4. However, analysis of patent networks with such a resolution at the patent level is quite complex and too broad, therefore less meaningful for the purpose of this work. That is why a more fine grained patent citation analysis was conducted by dividing the data sample for this time window into a series of three consequent snapshots each containing patents filed within two adjacent years. This approach allows to compare how the technology evolved in recent years and at the same time enlarge the resolution of the technology field so to be able to grasp the developments in more details.

4.1.1 Bibliographic coupling of patents filed between 2007-2008

Figure 7 shows the results of the bibliographically coupled patents, where the dots represent the patents and the connections between them represent their BC, as described in detail in the Methodology section. In total there are 582 granted utility patents filed in that period of time. After the application of 'Giant component' and 'Degree range > 10' filters as described in the Methodology section, the simplified patent network displays 226 coupled patents linked together by 2411 edges. Labeling each node of the patent network by patent assignees, as shown in Figure 7, allows displaying a map of patents from the viewpoint of patent holders. This way the relationship between different firms are identified through the examination of the citation relationships existing between their patents. The patent citation map distinguishes four main clusters of patents that share similar technological features according to Gephi's community detection mechanism presented in the Methodology section. This patent map provides an overview of the major technology domains within which the key players perform their patenting activities in the EV field.

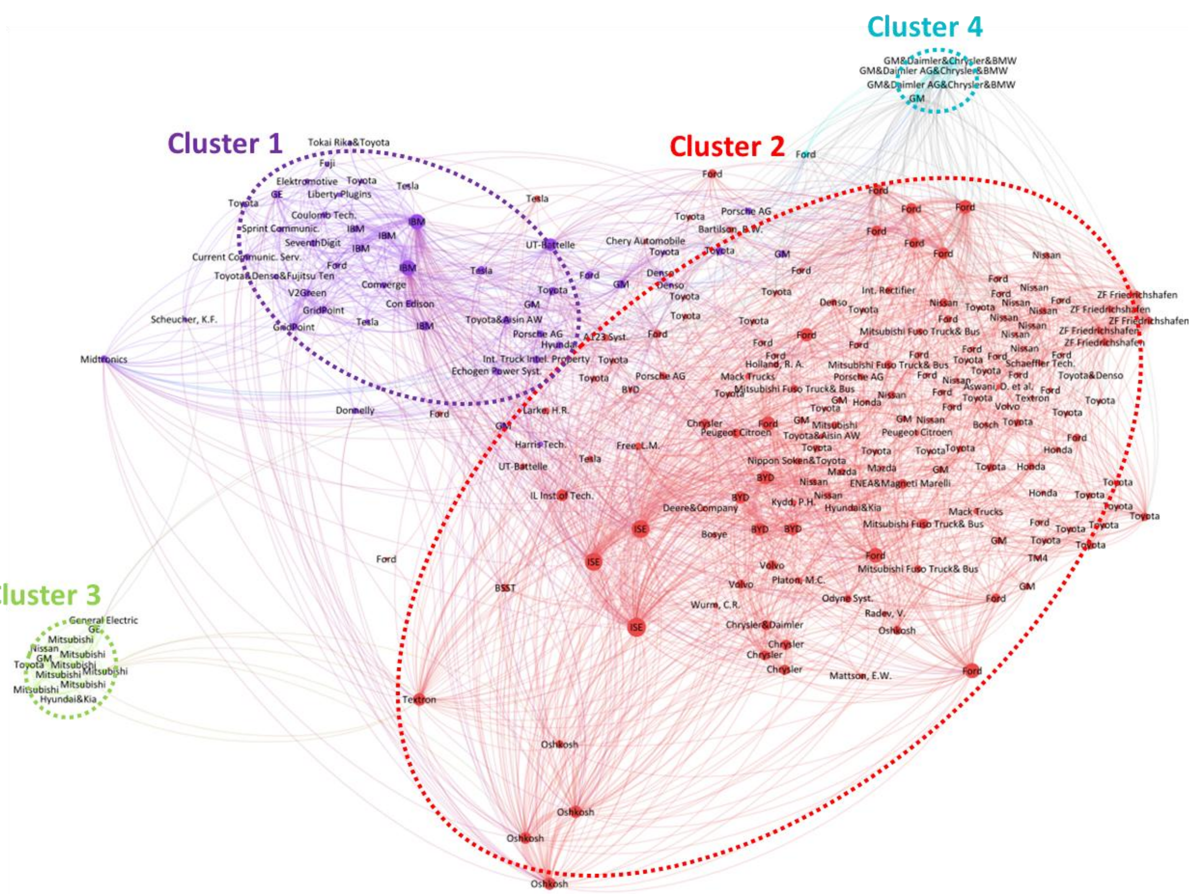
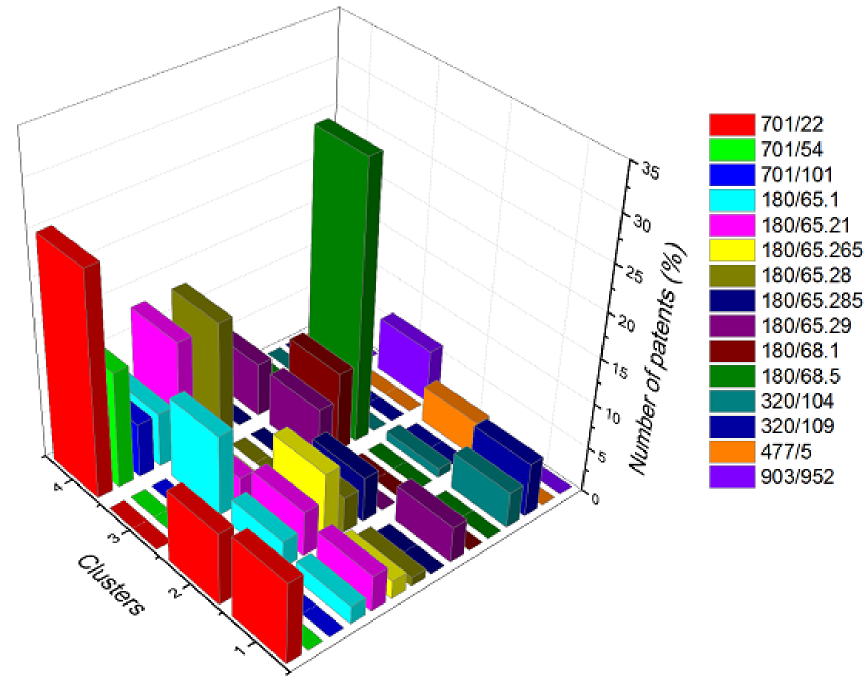


Figure 7: Patent citation map based on BC of patents filed between 2007 and 2008. Patents are clustered using community detection mechanism employed by Gephi. Each cluster contains patents that share similar technological features.

One of the most challenging issues of patent network analysis is identification and interpretation of clusters within the network of patented technological innovations. A hybrid approach was used for the analysis of patent clusters. As a starting point I draw on the US patent classification system according to which all patents are assigned by patent examiners to a certain technology field with respect to the subject they deal with. US patent examiners are perceived to classify patents especially accurate with US marks, which are frequently revised and improved. Analyzing the most frequently observed patent classes per cluster provides insight into the main subject that the cluster deals with. Figure 8 displays the distribution of such technology fields per cluster within the patent map of 2007-2008.

As the next step I analyze the patent titles and abstracts per each cluster. Here I also look at the most frequent key words in titles and abstracts of patens, as shown in Figure 9. The word-cloud analysis was employed as an additional research tool to visualize the most frequently used words in patent titles and abstracts. These results served as a background support when labeling the main technology domains represented by patent network clusters, but did not define clusters per se due to the limitations of the method as described in detail in the Methodology section. PCNs of this study were presented to field experts, who were asked to comment on labeling the technology clusters, validate the clusters' subject coherence and evaluate in general the usefulness of such patent mapping as an input to their strategic decision-making.

(a)



(b)

US class	Class definition
701/22	Vehicle control, guidance, operation, or indication/Electric vehicle
701/54	Vehicle control, guidance, operation, or indication/Transmission control and other vehicle control: Engine output control
701/101	Vehicle control, guidance, operation, or indication/With indication or control of power plant: Internal-combustion engine
180/65.1	Motor vehicles/Power: Electric
180/65.21	Motor vehicles/Power: Electric: Hybrid vehicle
180/65.265	Motor vehicles/Power: Electric: Hybrid vehicle: Control of multiple systems specific to hybrid operation
180/65.28	Motor vehicles/Power: Electric: Hybrid vehicle: Control of multiple systems specific to hybrid operation: Control of engine specific to hybrid operation
180/65.285	Motor vehicles/Power: Electric: Hybrid vehicle: Control of individual subunit specific to hybrid operation: Control of motor or generator specific to hybrid operation
180/65.29	Motor vehicles/Power: Electric: Hybrid vehicle: Control of individual subunit specific to hybrid operation: Control of battery specific to hybrid operation
180/68.1	Motor vehicles/Power: With means to guide and/or control air for power plant cooling
180/68.5	Motor vehicles/Power: Battery mountings and holders
320/104	Electricity: battery or capacitor charging or discharging/One cell or battery charges another: Vehicle battery charging
320/109	Electricity: battery or capacitor charging or discharging/Cell or battery charger structure: Charging station for electrically powered vehicle
477/5	Interrelated power delivery controls, incl. engine control/Plural engines: Electric engine: With clutch control
903/952	Hybrid electric vehicles/ Prime movers comprising electrical and internal combustion motors: Having energy storing means(e.g. battery): Housing details

Figure 8: (a) Distribution of technology fields per cluster. United States Patent Classification (USPC) categories represent patent technology fields. For comparison the amount of patents that are assigned to a certain technology field are shown in % of the total amount of patents per cluster. (b) The table clarifies the definition of each category as described by USPTO.

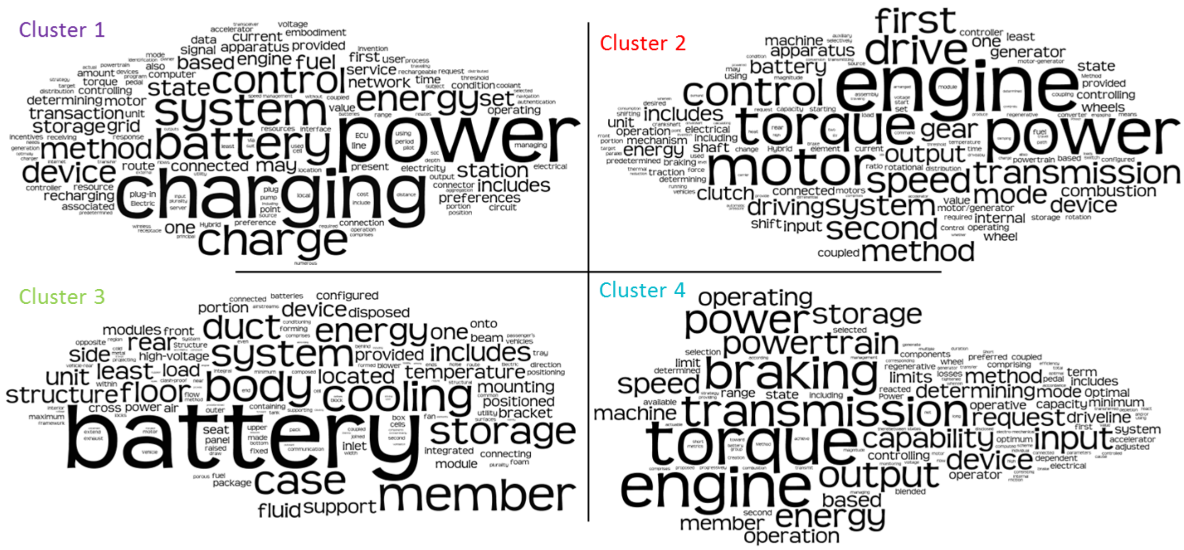


Figure 9: The word-cloud analysis of the most frequently used words in patent titles and abstracts per cluster. The patent search words “electric”, “vehicle/s”, “hybrid” were manually removed from the word cloud representation.

As a combined outcome of both quantitative and qualitative analysis, patent clusters were roughly categorized by their shared technological characteristics. This patent map presents the cognitive structure of the EV field in the period between 2007 and 2008 at the macro level. *Cluster 1* represents a mixed group of patents that deal with charging and power control in EV/Hybrid vehicles. The central player of this cluster is International Business Machines Corporation (IBM) that holds patents dealing with the technology for managing EV charging transactions. *Cluster 2* contains patents that mainly deal with power control in hybrid drivetrain and vehicle control in general. Cluster 1 and Cluster 2 are the largest clusters of the patent map and are closely interrelated with ‘power control in vehicles’ being a common subject area. According to expert opinion, these two clusters could be further broken down for a more fine grained analysis. There are two distinct smaller clusters, Cluster 3 and Cluster 4, which are positioned quite separate from the main component of the map. *Cluster 3* consists of patents that deal with cooling and mounting of batteries. The key patent holders of this cluster are, namely, Mitsubishi Jidosha Kogyo Kabushiki Kaisha, GM Global Technology Operations LLC and General Electric Company. *Cluster 4* forms a unique group of patents that share very similar technological features. Majority of the patents that form this cluster belong to the alliance consisting of General Motors Technology Operations LLC, Daimler AG, Chrysler LLC and Bayerische Motoren Werke Aktiengesellschaft (BMW). This 'Global Hybrid Cooperation' of large automakers was formed in 2004 jointly working on the development of a new hybrid engine technology, which combines a battery-powered electric motor with a conventional gasoline combustion engine (so-called dual-mode hybrid technology).

4.1.2 Bibliographic coupling of patents filed between 2009-2010

Using a similar approach I have built and analyzed a patent map that consists of patents filed between 2009 and 2010. Figure 10 displays the results of the bibliographically coupled patents, where the dots represent the patents and the connections between them represent their BC. In total there are 471 granted utility patents filed in that period of time. Therefore, similar filtering methods, i.e. 'Giant component' and 'Degree range > 10' filters, were applied for the simplification of this patent map. The resulting patent network displays 77 bibliographically coupled patents linked together by 539 edges. Even though the original amount of patents in this two-year period is comparable with the data sample for the years 2007-2008, the patent map for 2009-2010 turned out to be much smaller in terms of the amount of bibliographically coupled patents after filtering. This is explained by a smaller average degree of the network, i.e. 14.0 for the network of 2009-2010 vs. 21.3 for the network of 2007-2008, as discussed further in Table 3. The average degree in social network analysis is defined as the number of links that a given node has to other nodes, or the sum of all node degrees divided by the total number of nodes in the network [43]. This tendency of decreasing average degree of the bibliographically coupled patent network, which is confirmed in the next PCN for 2011-2012, might indicate the diversification of the field with unique innovations appearing that differ from the existing majority of the prior art, therefore not that densely coupled with other patented innovations. The patent citation map consists of three main patent clusters that share similar technological background according to Gephi's community detection mechanism presented in the Methodology section.

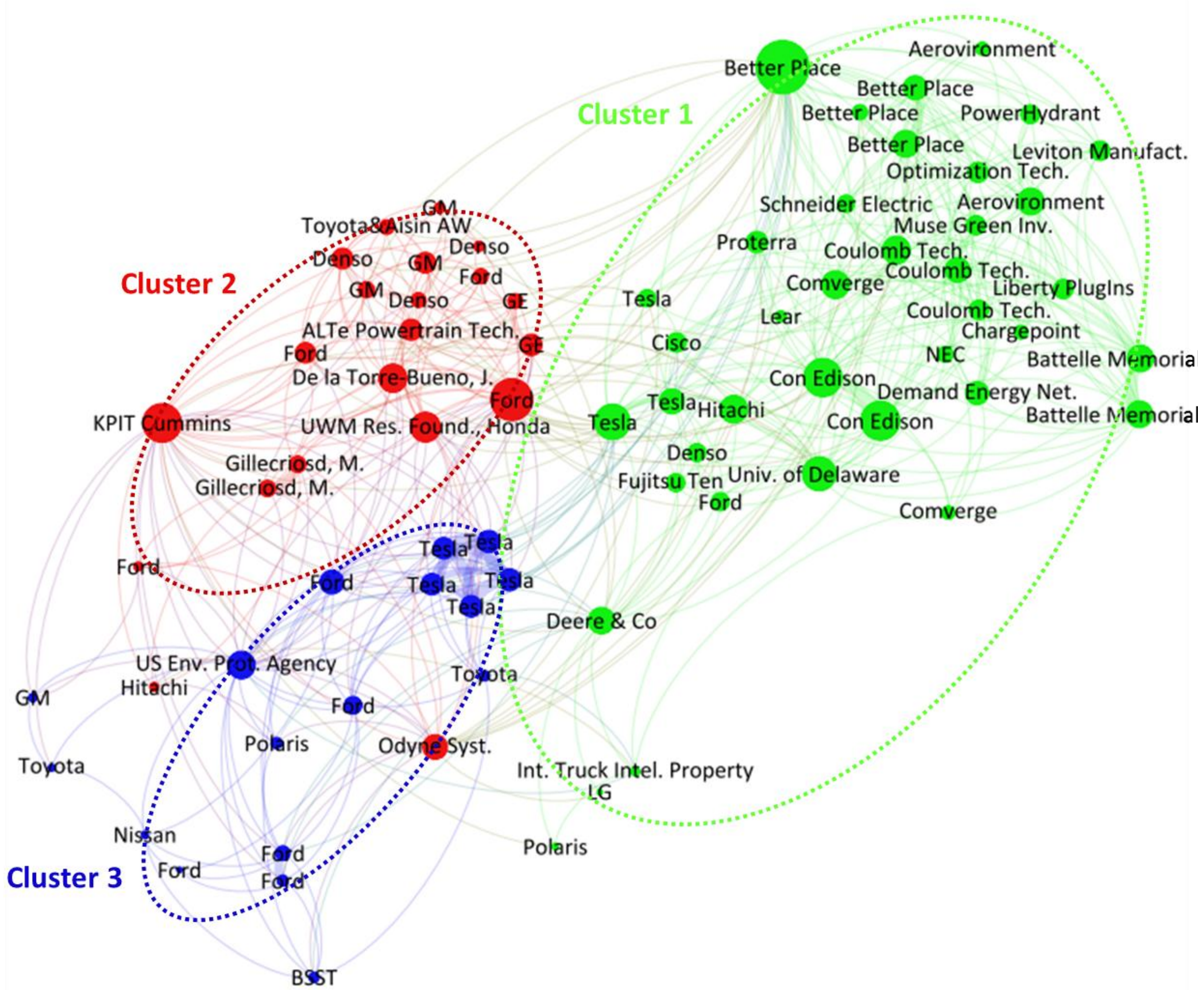
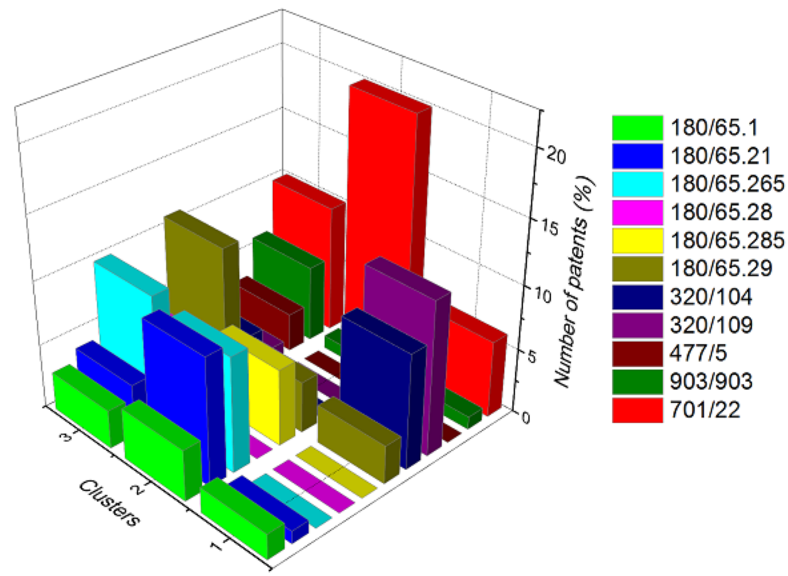


Figure 10: Patent citation map based on BC of patents filed between 2009 and 2010. Patents are clustered using community detection mechanism employed by Gephi. Each cluster contains patents that share similar technological features.

Further in a similar manner as for the earlier patent network, the clusters were analyzed with respect to the distribution of US classes of patents per cluster and more in-depth patent titles/abstracts investigation aided by word frequency analysis. The results are shown in Figures 11 and 12, respectively.

(a)



(b)

US class	Class definition
701/22	Vehicle control, guidance, operation, or indication/Electric vehicle
180/65.1	Motor vehicles/Power: Electric
180/65.21	Motor vehicles/Power: Electric: Hybrid vehicle
180/65.265	Motor vehicles/Power: Electric: Hybrid vehicle: Control of multiple systems specific to hybrid operation
180/65.28	Motor vehicles/Power: Electric: Hybrid vehicle: Control of multiple systems specific to hybrid operation: Control of engine specific to hybrid operation
180/65.285	Motor vehicles/Power: Electric: Hybrid vehicle: Control of individual subunit specific to hybrid operation: Control of motor or generator specific to hybrid operation
180/65.29	Motor vehicles/Power: Electric: Hybrid vehicle: Control of individual subunit specific to hybrid operation: Control of battery specific to hybrid operation
320/104	Electricity: battery or capacitor charging or discharging/One cell or battery charges another: Vehicle battery charging
320/109	Electricity: battery or capacitor charging or discharging/Cell or battery charger structure: Charging station for electrically powered vehicle
477/5	Interrelated power delivery controls, incl. engine control/Plural engines: Electric engine: With clutch control
903/903	Hybrid electric vehicles/ Prime movers comprising electrical and internal combustion motors: Having energy storing means (e.g. battery, capacitor)

Figure 11: (a) Distribution of technology fields per cluster. United States Patent Classification (USPC) categories represent patent technology fields. For comparison the amount of patents that are assigned to a certain technology field are shown in % of the total amount of patents per cluster. (b) The table clarifies the definition of each category as described by USPTO.

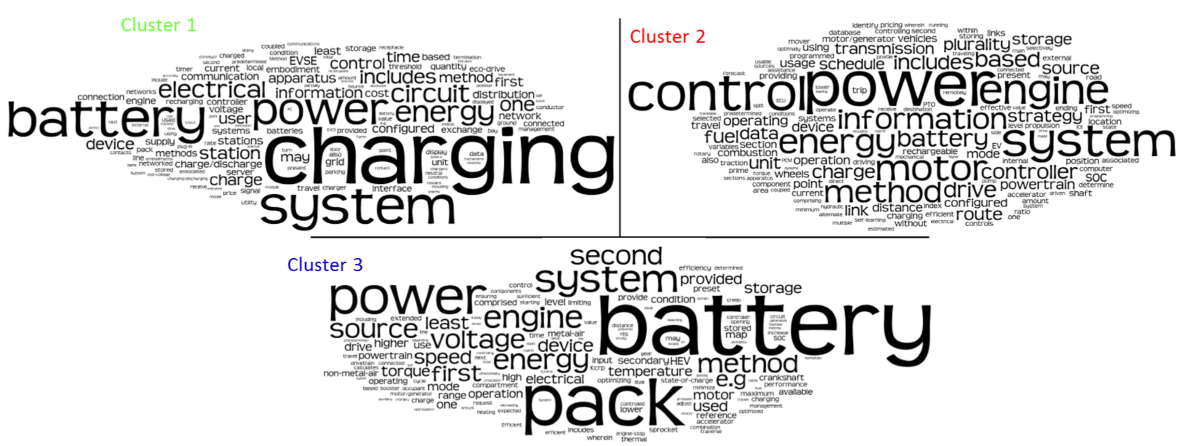


Figure 12: The word-cloud analysis of the most frequently used words in patent titles and abstracts per cluster. The patent search words “electric”, “vehicle/s”, “hybrid” were manually removed from the word cloud representation.

The patent cluster labeling was further verified by the field experts. As a result the largest technology domains in the EV field for the years 2009-2010 can be roughly described as follows. *Cluster 1* combines the technology area related to charging infrastructure (i.e. gas stations for EVs) and in-car charge management systems. This cluster could be further broken down to two separate clusters according to these topics. *Cluster 2* deals mostly with energy and power optimization in hybrid vehicles. *Cluster 3* is a smaller group, which is largely defined by the patent family assigned to Tesla focusing on dual power source/range extenders for EVs. These patents deal with a promising technology that is a new type of hybrid battery pack that combines two different battery technologies, lithium-ion and a metal-air battery pack, to deliver an increase in range of an EV. This cluster demonstrates a significant activity of Tesla in battery technology. Another highly-connected player of this cluster is United States of America as represented by the Administrator of the US Environmental Protection Agency (EPA). EPA’s lab has developed the hydraulic hybrid system allowing for the increases in fuel efficiency as compared with traditional powertrains and reduce the overall greenhouse gas emission. Later on in 2011 Chrysler will partner with EPA to demonstrate the hydraulic series hybrid powertrain for the reduction the size and complexity of hybrid systems.

4.1.3 Bibliographic coupling of patents filed between 2011-2012

In analogy to the previous PCN, the patents filed between 2011 and 2012 were also mapped based on BC (Figure 13). Due to a smaller total amount of patents granted for that time period to start with (i.e. 191 utility patents) as well as a low average degree of the network (i.e. 5.4), I have reduced the degree range filtering criteria for this network to >2 links per node. Therefore, all nodes that have 2 or more links are displayed. The resulting PCN contains 103 nodes and 276 edges. This resolution of the network was found to be optimal for the in-depth technology landscape analysis. The patent citation map consists of 8 different clusters, revealing the network technological richness (see Figures 14 and 15).

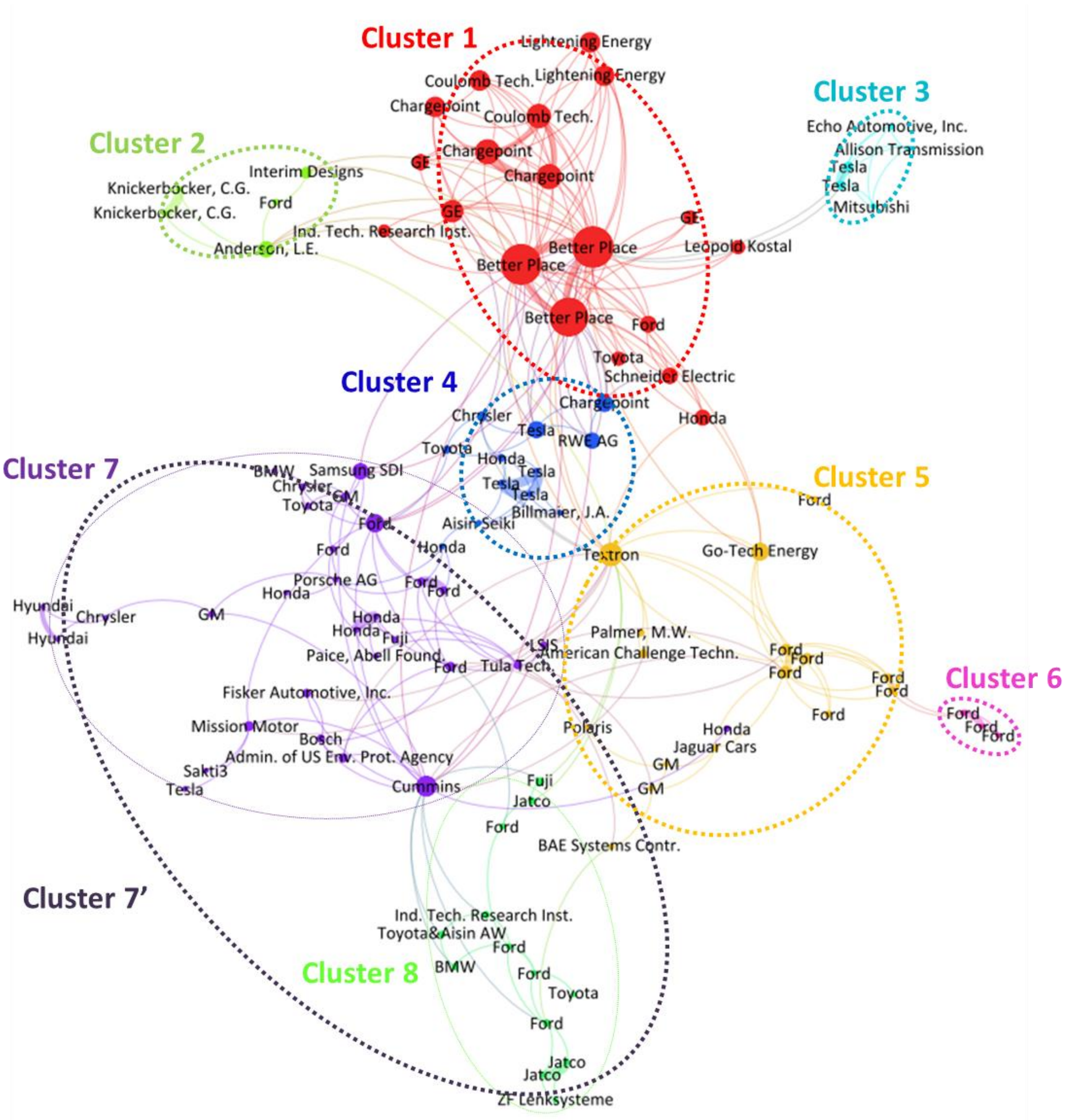
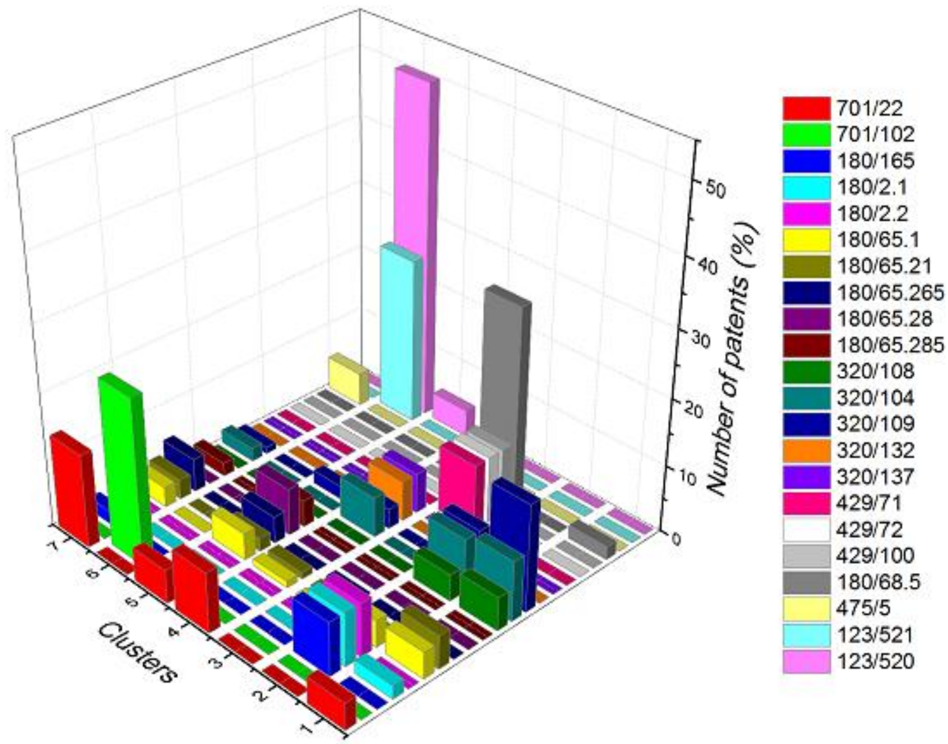


Figure 13: Patent citation map based on BC of patents filed between 2011 and 2012. Patents are clustered using community detection mechanism employed by Gephi. Each cluster contains patents that share similar technological features.

(a)



(b)

US class	Class definition
701/22	Vehicle control, guidance, operation, or indication/Electric vehicle
701/102	Vehicle control, guidance, operation, or indication/With indication or control of power plant/Internal-combustion engine/Digital or programmed data processor
180/165	Motor vehicles/With fluid or mechanical means to accumulate energy (I)Derived from motion of vehicle or (II)Obtained from operation of vehicle motor
180/2.1	Motor vehicles/Motor supplied with power from external source
180/2.2	Motor vehicles/Motor supplied with power from external source/Source comprises or includes energy derived from a force of nature (e.g. sun, wind)
180/65.1	Motor vehicles/Power/Electric
180/65.21	Motor vehicles/Power/Electric/Hybrid vehicle
180/65.265	Motor vehicles/Power/Electric/Hybrid vehicle/Control of multiple systems specific to hybrid operation
180/65.28	Motor vehicles/Power/Electric/Hybrid vehicle/Control of multiple systems specific to hybrid operation/Control of engine specific to hybrid operation
180/65.285	Motor vehicles/Power/Electric/Hybrid vehicle/Control of individual subunit specific to hybrid operation/Control of motor or generator specific to hybrid operation
180/68.5	Motor vehicles/Power/Battery mountings and holders
320/104	Electricity: battery or capacitor charging or discharging/One cell or battery charges another/Vehicle battery charging
320/108	Electricity: battery or capacitor charging or discharging/Cell or battery charger structure/Charger inductively coupled to cell or battery
320/109	Electricity: battery or capacitor charging or discharging/Cell or battery charger structure/Charging station for electrically powered vehicle
320/132	Electricity: battery or capacitor charging or discharging/Battery or cell discharging/With charging/Cycling/With state-of-charge detection/
320/137	Electricity: battery or capacitor charging or discharging/Battery or cell charging
429/71	Chemistry: electrical current producing apparatus, product and process/With system having means to move ventilating fluid
429/72	Chemistry: electrical current producing apparatus, product and process/Having specified venting, feeding or circulation structure
429/100	Chemistry: electrical current producing apparatus, product and process/Cell support for removable cell/Support or holder per se
475/5	Planetary gear transmission systems or components/Including electric motor input
123/521	Internal-combustion engines/Charge-forming device/Having fuel vapor recovery and storage system/Having an adsorbent canister/Responsive to secondary air pressure
123/520	Internal-combustion engines/Charge-forming device/Having fuel vapor recovery and storage system/Having an adsorbent canister/Purge valve controlled by engine parameter

Figure 14: (a) Distribution of technology fields per cluster. United States Patent Classification (USPC) categories represent patent technology fields. For comparison the amount of patents that are assigned to a certain technology field are shown in % of the total amount of patents per cluster. (b) The table clarifies the definition of each category as described by USPTO.

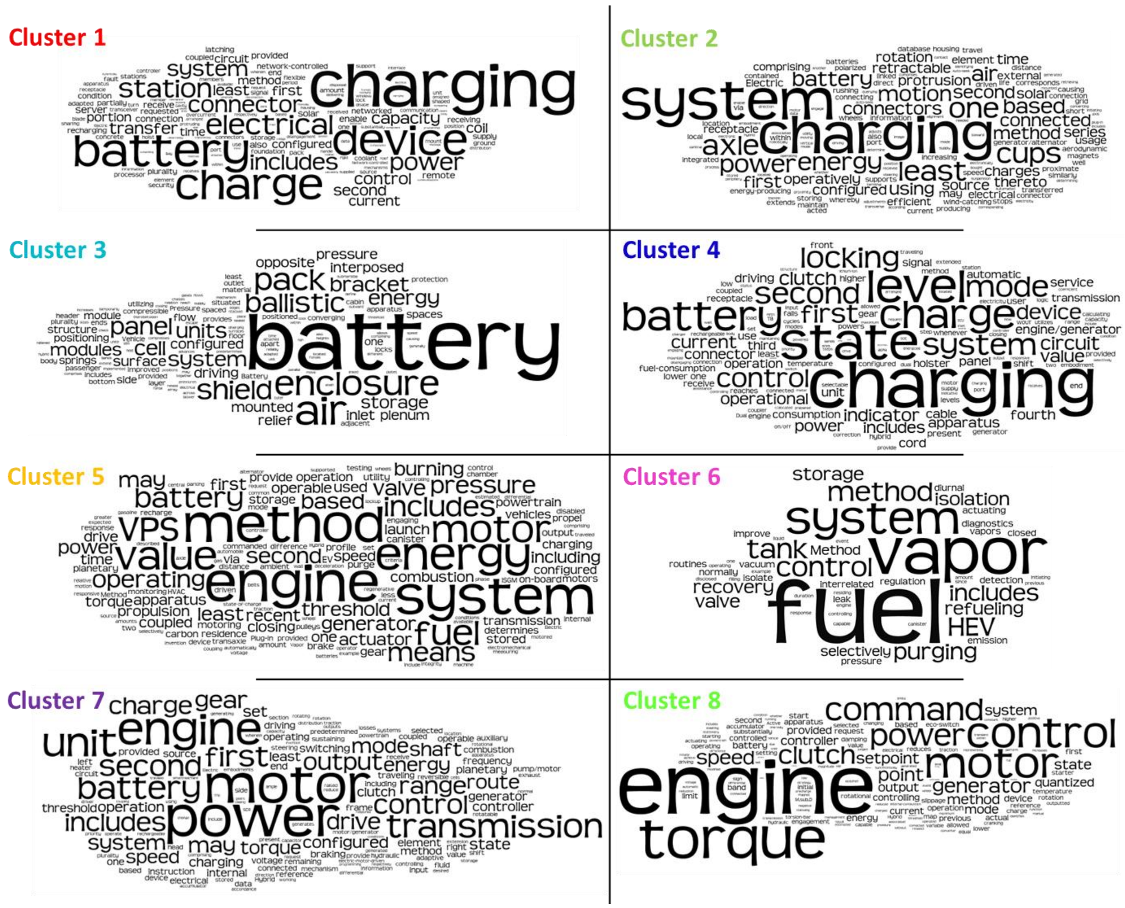


Figure 15: The word-cloud analysis of the most frequently used words in patent titles and abstracts per cluster. The patent search words "electric", "vehicle/s", "hybrid" were manually removed from the word cloud representation.

Following the methodology described earlier, I segmented the patent network into 8 distinct clusters that share common technological background. *Cluster 1* is a very well-focused cluster with patents that target optimization of EV charging network. Cluster 2 and Cluster 3 are smaller groups of patents closely positioned to Cluster 1. *Cluster 2* represents a collection of patented innovations mainly by individual patent holders and deals with energy producing systems and automated charging. *Cluster 3* contains patents that are related to battery mounting and cooling. According to the interviewed field experts, the appearance of this cluster, in other words, research and patenting activities in this area, could be due to the press on battery pack vulnerability, e.g. Tesla's Model S fire due to road debris that cause mechanical deformations of lithium-ion battery packs that sit underneath the vehicle's body. Battery mounting and cooling appears to be a subject area of continuous research and development in the EV field, with Mitsubishi being the most active player according to the map of 2007-2008 (Cluster 3) and Tesla Motors as a new-comer in this field in the map of 2011-2012 (Cluster 3). *Cluster 4* presents a mixed group of patents with the core patents belonging to Tesla Motors. These patents deal with dual mode battery and extended range EVs. This cluster and Cluster 3 of the previous patent map (2009-2010) demonstrate continuous patenting activities of Tesla Motors in the field of EV battery technology. *Cluster 5* is focusing on optimizing hybrid technology, i.e. when to operate combustion engine or range extender. *Cluster 6* is a very small cluster of several Ford patents that deal with fuel vapor control. *Cluster 7* combined patents that deal with EV inverter, transmission and power delivery. This cluster is very similar to *Cluster 8*, which focuses on engine torque and transmission control patents on hybrid vehicles. Therefore, cluster 7 and 8 could be combined to form a cluster on EV vehicle control systems.

4.1.4 Topological analysis of patent networks

After visualizing the technology landscape in the recent 6 years using patent networks, I further analyze the topological characteristics of the networks by means of a number of statistical measures. For this analysis the whole network was used without the filtering which was earlier applied for the clarity of patent maps representation. Table 3 displays the results of the topological analysis of the networks.

Table 3: Results of the topological analysis of the three patent citation networks based on BC.

Network by patent filing year	Number of nodes	Number of edges	Average degree	Density	Diameter	Average path length	Average clustering coefficient
2007-2008	486	3298	21.3	0.028	9	3.141	0.449
2009-2010	381	1303	14.0	0.018	9	3.583	0.491
2011-2012	142	311	5.4	0.031	9	3.796	0.529

It is observed that the average degree of the network decreases with time. As mentioned earlier, the average degree represents the average number of links that one node has to other nodes. This might to some extend reveal the structure of the innovation landscape within the field, suggesting a decrease in connectivity between patents. With the same network diameter, or the length of the longest geodesic path in the network, the average path length between any pair of nodes differs between the networks, slightly increasing in recent years. The average path length gives an overall indication of the number of links needed for knowledge dissemination within the network or the efficiency of information transport [43]. In case of BC of patents it can be interpreted as the average number of steps along the shortest path between all possible pairs of bibliographically coupled patents that needs to be taken in order to link patents according to the similarity of their research topic and technological background. An average clustering coefficient indicates the degree to which the nodes in the network tend to cluster together and is calculated as the average ratio of the number of actual links between the node's neighbors to the maximum possible number of links between those neighbors [43]. The slight increase in the average clustering coefficient of our PCNs with the concurrent increase in the average path length in recent years demonstrates the distribution of patents between smaller patent clusters which are connected by longer links, with the majority of patents clustering in those groups and the average distance between them is dominated by the longer distance between the groups. At the same time, as shown earlier in Table 2, the average coupling strength between the patents is also increased in recent years, indicating a strong grouping of patents that share related technological features or applications. According to the methodology of PCN construction in this work, patents are densely linked with other patents if they share references with a large number of other patents in the network, and patents in the network are positioned closer to each other if the amount of shared references is high. Lower values of the average network degree and the average path length between patents indicate the similarity of patents within the network, which share the same and narrow knowledge base. Our observations on the decrease of the average degree of the network and increase of the average path length between the nodes in the recent years might suggest a diversification of the EV field in general with new patented innovations being less widely based on the prior existing innovations or the overall broadening of the research field with slightly more variant topics on a larger scale. It is important to note, however, that graph measures may be influenced by the number of nodes of the network [114] and additional corrections for the size effects are needed to avoid possibly spurious results. Therefore, further study is encouraged in this direction.

4.2 Patent network analysis as a strategic information source

Effective management and planning of technologies is essential for building competitive advantage of a firm. However, technology management is a challenging topic due to the technological complexity and the rate of technological change [16]. Patent-based technology development map was shown to be a useful tool to represent the dynamic history of technology development and support technology management [16]. A prominent study on network dynamics and the evolution of the technology landscape in the field of biotechnology from the perspective of interorganizational collaboration was presented by Powell *et al.* [104]. Looking at the dynamic relationship between patents at the assignee or patented technology levels allows to map the recent developments in the technology focus over the years, changes in the patenting activities of key players within the field and the industry structural changes in general. This information can be useful for decision makers to evaluate the technology landscape including areas of high patenting activities, innovations that pose a threat or present new opportunities for business, and hence support the development of technological strategy, e.g. investment in and development of specific technologies [16]. Using patent maps presented at patent assignee level allows firms to see potential competitors or partners more objectively, analyze the focus of company on a certain technology domain at a particular time and pinpoint strategic changes in their patenting activities. Mapping patents at assignee level can become a useful method to identify the firms that have the legal right to a specific technology and determine the licensing opportunities required for the business operation

[16]. Patent network dynamics reveals the changes in the industry structure, becoming a source of information of the firm's competitive environment.

4.2.1 Industry dynamics from the viewpoint of technologies

Looking at the patent network dynamics constructed using BC approach within the EV field in the last 6 years can reveal a number of interesting observations. It is visible that managing EV charging is a subject of significant patenting activities in the field which in recent years is especially focused on charging infrastructure (i.e. Clusters 1 in all the three maps). Besides a lot of activities in the area of innovative battery technologies and extension of the EV range (Cluster 3 in the map of 2009-2010 and Cluster 4 in the map of 2011-2012) are observed as well as mounting and cooling the battery (Cluster 3 in the map of 2007-2008 and Cluster 3 in the map of 2011-2012). The main issue for EVs to be ready for the mass market and accepted by customers is the availability of a battery technology that could provide the energy for a significant driving range, addressing at the same time car safety issues and competitive costs in comparison with gasoline cars [74]. To meet these requirements, one direction of technological activities target innovations in battery technology to increase the battery capacity. Another direction is focused on developing fast and convenient charging solutions. That is why optimization of EV charging network and developing the required charging infrastructure occupies a prominent position on the patent map in recent years and is expected to grow in the coming years as a prerequisite for a wide market penetration of EVs. Another technology domain of continuous research in the EV field is related to energy and power optimization in hybrid vehicles, i.e. strategies to receive the best outcome when combining ICE and electric engine (Cluster 2 in the map of 2007-2008, Cluster 2 in the map of 2009-2010 and Cluster 5 in the map of 2011-2012). Companies are spending an increasing amounts of development activities in this area due to the limited range of pure electric powertrains associated with the energy density and cost of batteries [74]. In order to have a more detailed assessment of the technological innovations within these main domains a more in-depth patent content analysis per cluster is recommended, which is a valuable study of a further research with the focus on the analysis of certain technologies within the EV field.

Patent mapping presents a valuable tool to ascertain or forecast the likely future technology development trajectories. The case study of this research work provides another example demonstrating the open-endedness surrounding emerging technological fields [7]. Technological breakthroughs from any specific technology, including battery technology, electric motor technology, controlling systems etc., will have a significant impact on the development of the EV industry [72]. We observe that currently the actors within the field of EV technology are betting on different competing technological designs and expectations, instead of focusing on one technological structure. Taking a powertrain configuration for EVs as an example, we observe patents that deal with technologies that concern pure electric vehicles, serial hybrids (range extended), plug-in hybrids that use parallel or power-split hybrid systems as well as micro/mild hybrids. Pure electric vehicles are expected to become mainly city cars, small vans or “fun cars” in the future; range extenders produced as serial hybrids are expected to appear in vehicles of compact segment and small delivery trucks of city use, while plug-in hybrids using parallel or power-split hybrid systems will be applied in upper medium and premium classes, and in SUVs [74]. Further research can advance towards more in-depth analysis of patent clusters within battery-driven vehicles, and widening the data scope including also the competing technologies of hydrogen or fuel cell vehicles. This insight was obtained from a private discussion with the EV field expert. The emerging technology field of EVs is characterized by a high level of government engagement, as opposed to more basic science based technological fields, e.g. nanotechnology, or more engineering- and application-oriented fields, e.g. software [7]. That in part explains the observation that the EV field is mainly represented by firms and large corporations and not so much by research centers or individual inventors.

The unique position of the study field distinguishes the research findings from other previously reported works that consider technological uncertainties within an emerging field, and at the same time demonstrates the generalizability of the earlier demonstrated findings by Gustafsson *et al.* [7], making the results especially interesting for the further analysis with respect to competing guiding images and underlying generalizations in the EV field. Application of the heuristic framework demonstrated by the researchers could help to characterize the future technological developments in the EV field and assess their risks and uncertainty levels [7].

4.2.2 Industry dynamics from the viewpoint of key players

In order to analyze the recent developments in the EV structure over the last 6 years, I have analyzed the PCNs from the perspective of patent holders. All the patent holders were categorized into four groups: automotive OEMs, suppliers, infrastructure providers (e.g. energy suppliers, EV service and service equipment providers, etc.) and other players including government agencies, research centers, universities and individual inventors. The grouping of assignees was verified by the field experts in EV domain. Figures 17-19 present the above demonstrated patent maps based on BC between individual patents, where groups of assignees are color-coded.

2007 – 2008: total 226 patents

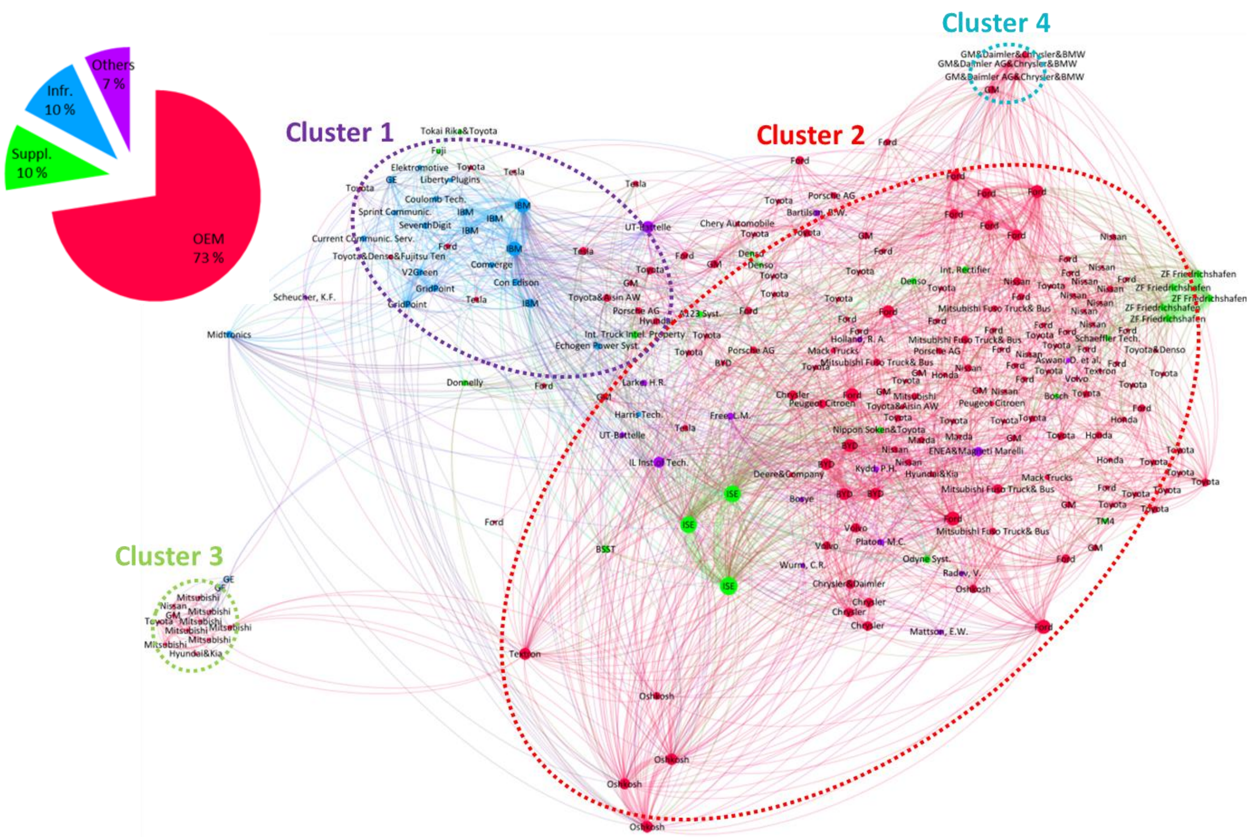


Figure 17: Patent citation map based on BC of patents filed between 2007 and 2008. Patents are clustered using community detection mechanism employed by Gephi and color-coded according to the affiliation of patent assignees to a certain category, i.e. OEMs, suppliers, infrastructure providers or other players. Total amount of patents in this sample assigned to OEMs is 164, to suppliers is 23, to infrastructure-related firms is 23 and other players is 16.

2009 – 2010: total 77 patents

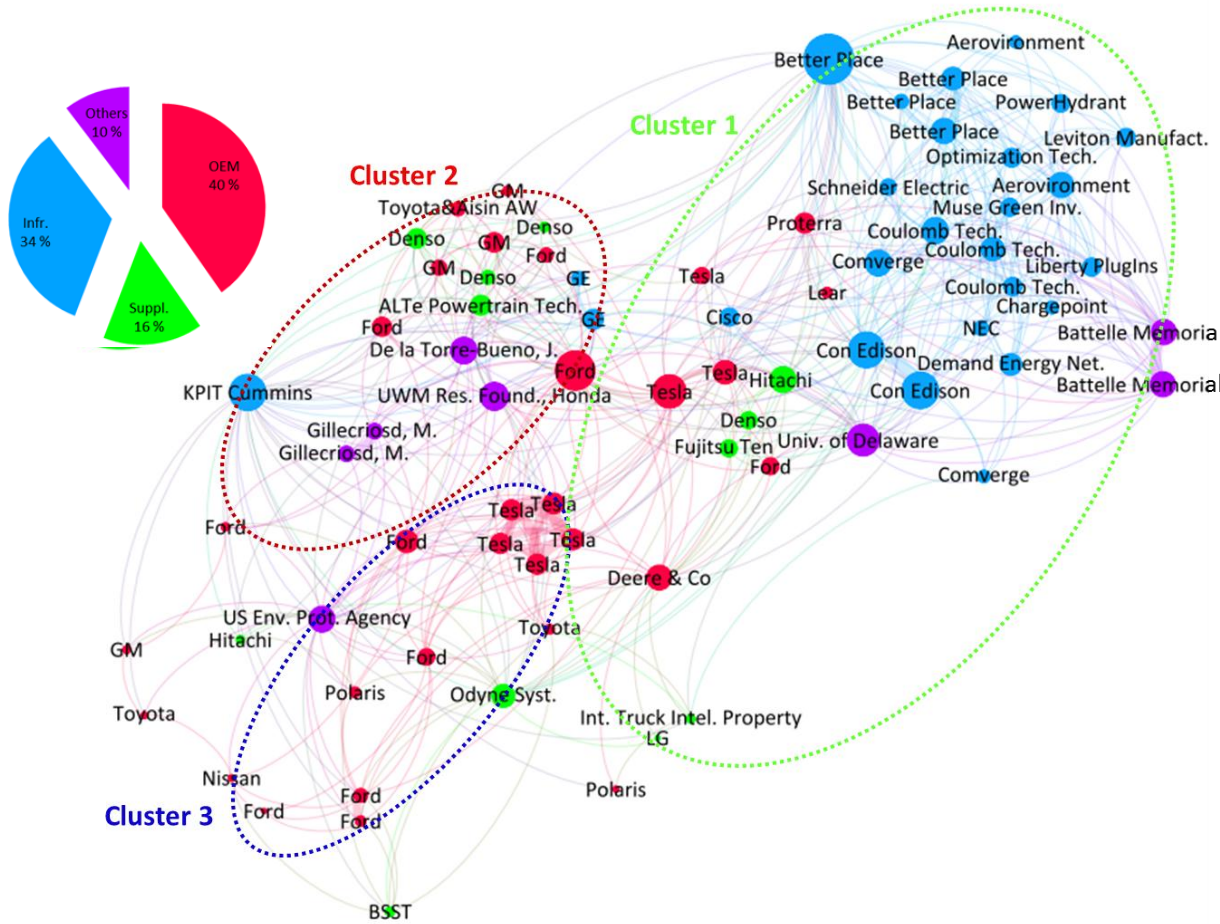


Figure 18: Patent citation map based on BC of patents filed between 2009 and 2010. Patents are clustered using community detection mechanism employed by Gephi and color-coded according to the affiliation of patent assignees to a certain category, i.e. OEMs, suppliers, infrastructure providers or other players. Total amount of patents in this sample assigned to OEMs is 31, to suppliers is 12, to infrastructure-related firms is 26 and other players is 8.

2011 – 2012: total 134 patents

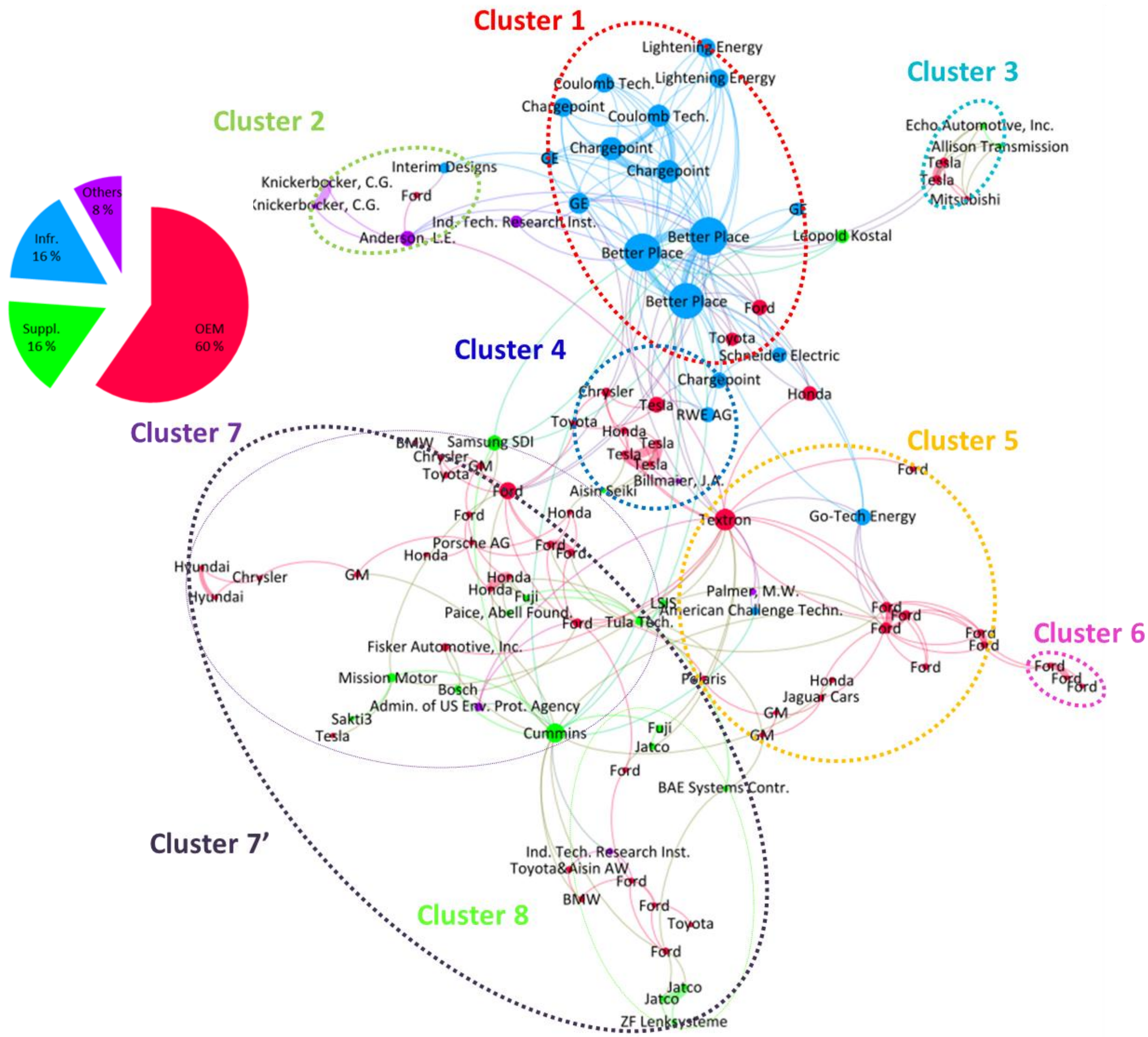


Figure 19: Patent citation map based on BC of patents filed between 2011 and 2012. Patents are clustered using community detection mechanism employed by Gephi and color-coded according to the affiliation of patent assignees to a certain category, i.e. OEMs, suppliers, infrastructure providers or other players. Total amount of patents in this sample assigned to OEMs is 80, to suppliers is 22, to infrastructure-related firms is 21 and other players is 11.

Analysis of the patent network dynamics from the viewpoint of key players reveals a significant role of OEMs as the major players in the field of EVs. Even though established OEMs have a greater knowledge background in vehicle design in general, they start losing their advantage to the newcomers within the EV field that specialize in battery technology or the design of electric motors [74]. Tesla Motors, Proterra, Fisker Automotive are the examples of the firms being the recent newcomers to the field. Traditional car makers experiencing the shift from a comparatively stable technology environment to an emerging EV technology are inclined to collaborate with many external firms outside the boundaries of the traditional automotive sector and supply environments [9]. A prominent tendency of the increase in the role of automotive suppliers, infrastructure providers and other industry players in recent years is clearly observed comparing PCNs of this thesis. There is a time lag between the active participation of OEMs in patenting activities within the EV field (see patent map of 2007-2008) and the active involvement of infrastructure providers (see patent map of 2009-2010). This signifies crucial changes in the automotive industry in recent year from largely vertical, when key automakers manage their products across the value chain, to predominantly horizontal with lower entry barriers for independent industry players as discussed in Section 3.1.

Analyzing patent clusters from the viewpoint of assignees' role within the field reveals that certain clusters are mainly represented by one of these assignees' categories. In the patent map of 2007-2008 Cluster 1 that deals with charging and power control in EV vehicles is partly represented by a group of infrastructure related firms, e.g. IBM, Consolidated Edison, GridPoint, V2Green, Current Communications Services, General Electric, Coulomb Technologies, Electromotive, SeventhDigit Corporation, Echogen Power Systems, Midtronics. The other part of the cluster includes OEMs, namely, Tesla Motors, Toyota Jidosha Kabushiki Kaisha, GM Global Technology Operations, Ford Global Technologies, Porsche AG, Hyundai Motor Company. As suggested by EV field experts, grouping the major players in the field provides additional insight into cluster verification and refinement of cluster labelling. For example, Cluster 1 in the patent map of 2007-2008 could be broken down into two clusters according to the assignee type, i.e. infrastructure providers dealing with EV charging and OEMs dealing with power control in EV vehicles. Cluster 2 of this network is mainly represented by OEMs, with the main players being Ford, BYD, Nissan, General Motors, Chrysler etc. BYD is a good example of a manufacturing company from emerging market (China) who moved from being a newcomer in the EV field to a global leader in rechargeable batteries, that successfully produces plug-in hybrids and pure electric vehicles [74]. Looking at the patent map of 2009-2010, Cluster 1 exhibits a separate sub-cluster formed by infrastructure providers that focus on managing EV charging networks. The key influential players include Better Place, Consolidated Edison Company and Coulomb Technologies. Other industry players including OEMs, like Tesla Motors, Ford, Proterra, Deere and Company, and several automotive suppliers, like Fujitsu Ten, Hitachi, Denso, form a second sub-cluster that mainly deals with in-car charge management systems. Among other significant players of this cluster are Battelle Memorial Institute and University of Delaware. The same cluster refinement guided by the role of patent assignees as suggested earlier could be applied here. Cluster 1 of the patent map of 2011-2012 is already more focused uniting mainly infrastructure providers as patent assignees, including Better Place, Chargepoint, Coulomb Technologies, General Electric, Lightening Energy, Schneider Electric, RWE AG, Co-Tech Energy. By working closely with OEMs, infrastructure providers are determined to provide end customers with attractive solutions for recharging their batteries at competitive prices [74].

A tendency of certain firms to form agglomerations of strongly BC patent pairs was observed from PCNs, for example, in the map of 2007-2008 such separate small patent agglomerations are formed by patents assigned to IBM, ISE, Oshkosh, BYD, ZF Friedrichshafen; in the map of 2008-2009 by Tesla's patents and in the map of 2011-2012 by the patents assigned to Better Place, for example. These patents that belong to the same organization possess strong intra-group interactions (visualized by thick edges that link the patents) as well as dense inter-group interactions within the cluster, as demonstrated by the short average path length for the networks. The observed grouping of patents assigned to the same organization originates from self-citation or citing other patents of the same organization. This can in part be explained that patent applicants from the same organization draw on the same existing knowledge base and these patents share closely related technological features of the same innovative technology that the company is developing. It was discussed in literature that firms and inventors patent strategically, choosing citation with respect to potential infringement and holdup threats, and perform self-citation in patents in part by the threat of litigation by owners of related patents [37]. Understanding patenting strategies is important for the correct interpretation of PCNs. Owing to the growing importance of patenting in the technology strategy, firms may choose a certain citation strategy in an attempt to defend themselves against the intellectual property of rivals [88]. Identifying such strong clusters in the field allows a firm that turns to a PCN for a support in strategic decision-making to pinpoint a technological domain in which other firms are actively protecting their innovative ideas through patents.

Finally, I present an example of patent network analysis from the viewpoint of geographical affiliation of patent assignees. Such networks based on BC of patents reveal the knowledge flow and commonality of technological background (prior art) between countries. Figure 20 displays a PCN based on BC where patents are labeled and color-coded by the country to which the patent assignees are affiliated. It can be observed that the majority of US patents are coupled with other US patents meaning that the knowledge from within the US plays a more important role in EV research and development as compared to knowledge from other countries [115]. At the same time US patents affect almost all other patents assigned to international organizations (such as Japan, Germany, Taiwan, Switzerland), who at the same time exchange knowledge between each other [115]. It important to note that since the data sample was extracted from USPTO, a 'home advantage' effect undermines the validity of conclusions drawn from such PCNs from the geographical distribution perspective. In this case worldwide patent data is required. Moreover, direct citation or CC analysis is beneficial in this study in order to observe direct knowledge flows and main knowledge diffusion centers.

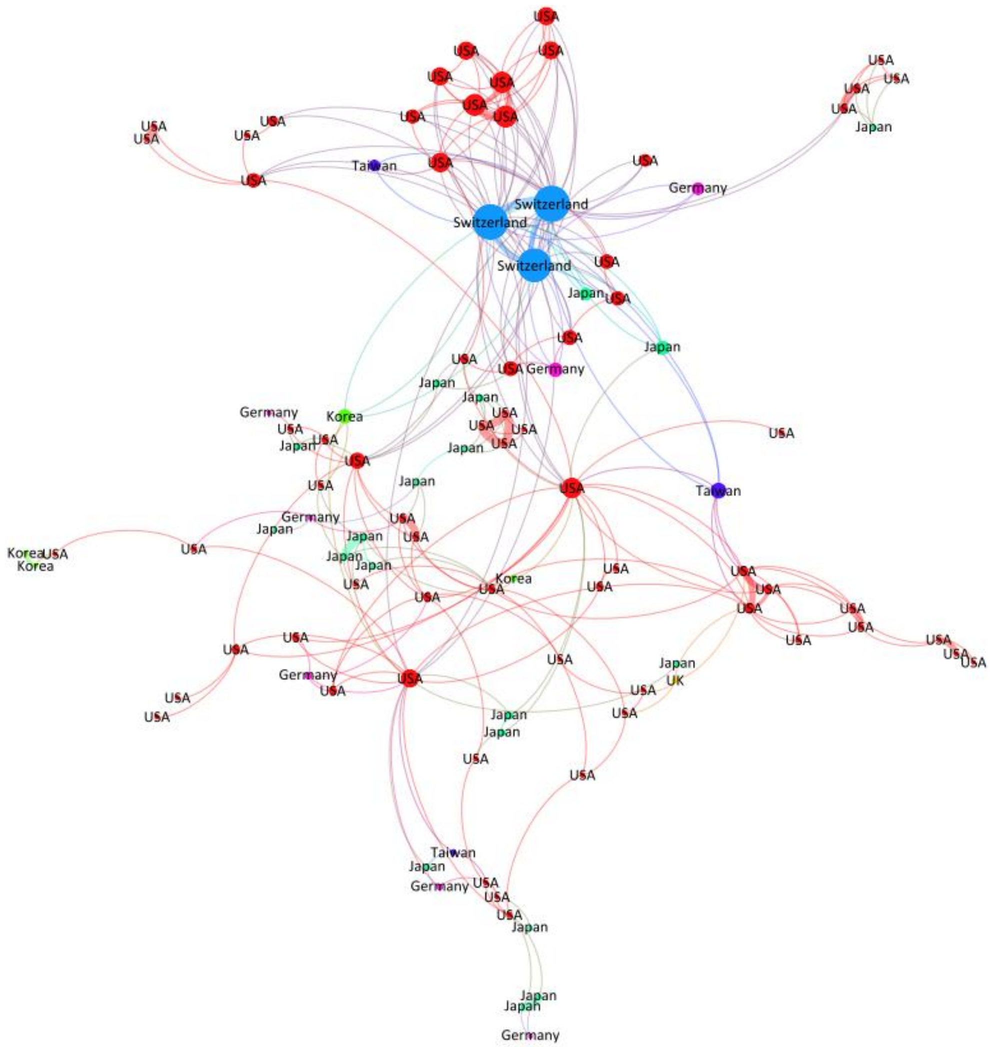


Figure 20: Patent citation map based on BC of patents filed between 2011 and 2012 where patents are labeled by the country to which the patent assignee is affiliated.

4.3 Limitations of patent mapping analysis

The research work on patent mapping based on BC shows a wide range of possibilities provided by this method and encourages future research in this area. However, it is important to raise awareness about the existing limitations of this and similar research works. This section is devoted to the discussion of existing methodological limitations of patents and patent citation analysis that also constitute the limitation of this particular research work and are generalizable to similar kinds of studies. Figure 21 summarizes the main sources of limitations within this analysis approach as elaborated further.

Figure 21: Sources of patent citation network analysis limitations.

One of the underlying limitations of any research that deals with the analysis of patent citations is that it reflects a social process, largely dependent on changing patent laws, habits of patent examiners, the pace of economic growth, changing government regulations (e.g. in case of EV field) and many other factors that influence the results of technology developments shown by PCNs [27]. Therefore, the analysis methodology enables to capture mainly a larger picture of technology environment based on a number of assumptions and provides a glimpse into a relatively short-term future. Besides, patents as a data source to indicate technology development have two important limitations. First of all, not all innovations are actually patented because not all of them satisfy the criteria of patentability [20]. Therefore PCNs do not include pure scientific discoveries that do not show immediate commercial applicability. Second, not all patentable innovations are actually patented since patenting is a strategic decision that differs across companies and industries [16]. For example, some firms might prefer to rely on secrecy or other appropriate means to protect their property right, which also depends on the industry under consideration. Therefore, one of the basic limitations that need to be considered when carrying out a patent-based research is that patents constitute only a part of all research outcomes and technology development in general. However, it is generally accepted that these limitations are not severe when looking at technology trends and developments at a larger scale [20].

A number of methodological limitations have been partially shown in the Methodology section of this work. Here for the purpose of summary, I discuss these and other additional limitations, stemming from the chosen data source and methods of analysis. Basing the research on patent citations implies a certain bias to the reason those citations are given associated with the subjective judgment by the patent examiner and inventor [68]. The assumption is that a patent citation represents a piece of previously existing knowledge upon which the citing patent is built on and over which it cannot have a claim [20]. Due to the role of the government patent examiner, who is an expert in the area, and the legal significance of the patent citations, patent citations are believed to be less likely contaminated by other motives of citations as compared to the citations in the scientific literature, for example, and therefore provide a more objective source of information [20, 22]. Here I refer the reader to a comprehensive review paper by M. Meyer who investigated in detail the similarities and differences between patent and research paper citations [22].

One of the main methodological limitations is related to the choice of the patent data source. When working with patents that originated in different countries, it is important to consider the difference in the approaches to the examination of patent applications. As mentioned in the Methodology section, due to the US law on the duty of disclosure, US-originated patents have a higher citation frequency as compared to European-originated ones. This requirement to ensure that all the aspects of the patent claim are backed up by the existing prior art, including the reference to the competing patent or a patent that deals with a similar end application, introduces some noise from patents that have no strong relevance to the core of the patented invention and certain bias toward the citations given by the applicants themselves. On the other hand, European patents might contain much less citations but assumable possess a higher cognitive relevance to the citing patent [22]. Therefore, depending on the purpose of the research work and data availability, one needs to consider these differences in patenting between different countries, especially when comparing international patents from a variety of patent data sources. Earlier studies have shown the existence of a ‘home advantage’ effect in both USPTO and EPO databases that has a considerable effect on measuring the patenting activities of foreign organizations, underestimating the patenting activities of firms operating outside the US and European market, respectively, and overestimating the patenting activities of the local companies [116]. This limitation should also be considered when choosing a source of data. It was shown that using patent families, i.e. patents that have been filed in different countries to protect the same invention, in order to construct PCNs might compensate for the 'home advantage' effect [116].

Another issue is associated with the time difference between the actual invention and the time when a patent is granted. Most of the patents mapping studies, including this work, base their analysis only on patents that have been granted. However, there always exists a time delay between the application for the patent and granting of the patent which is around 2-3 years as shown in literature [20] and also demonstrated in this work (see Table 1 of the Methodology section). One alternative data source is patent applications, which allows accessing the most recent data on technological inventions. However, as described in the Methodology section, dealing with patent applications contains a number of own limitations related to the aim of the patent application and a possibility that it is eventually not going to be granted. In case of patent applications, the value of PCNs as a strategy input could be improved by considering the different reasons that citation to prior art are made in patent applications. Patent citations that define the state of the art in a technology field and provide technological contribution should be distinguished from the so-called 'blocking citation' that challenge the novelty of the patent under examination [117]. European Patent Office (EPO) database provides details on how the cited document is applied to the patent application being examined [118]. Unfortunately, USPTO database does not provide the categorization of references cited by the examiner, and therefore, additional commercial services (e.g. Patent River) should be consulted to obtain information on blocking citations [119]. Recent studies have shown that patents highly cited as state of the art represent valuable contribution in the field [120]. On the other hand blocking references in patent filings of other firms are especially valuable for their owners since they allow to prevent the competitors from entering the market by legally blocking them from obtaining patents on related inventions or narrowing the scope of their patent claims [120]. Obtaining information on different types of references would allow to focus the PCN analysis in the right direction with respect to the research question of interest.

Additionally to that, there also exists a certain delay between the point in time when the invention was made and the time when patent application becomes public. In most countries patent application is typically published 18 months after it is filed. However, introduction of a product based on a patent in the market usually happens after the granting of a patent, and therefore patent application is often still considered as the earliest publicly known information about the invention [121].

After the choice of patent data source is made, the next methodological challenge is the extraction of a relevant data sample. Various patent search strategies have been proposed, e.g. search by the keywords, classification codes, or both at the same time [3, 84]. The choice of the search strategy largely depends on the field of study and the research goals. A search by the key word was chosen in case of this work, which allowed obtaining a broad dataset that covers all the patents dealing with electric or hybrid vehicles in the title or abstract, and avoid a bias toward specific area of the field. This type of bias is always present when a manual choice of patent classification codes is done to restrict the search criteria within a certain technology field. Keyword search refined by the classification code of the relevant patents could provide a method to obtain a well-focused data sample in terms of a specific technology. However, this patent search strategy is appropriate only when knowing most of all the possible classification codes that the relevant patent might be assigned to. Based on the data sample of this study extracted from USPTO, each patent might contain as much as 20 IPC or 32 US classification codes. One possibility is to perform data search refinement based on the classification codes after the examination of a broader data sample that is extracted based on the key words.

A number of approaches for patent mapping have been proposed. One of them is the method used in this work, i.e. BC or co-occurrence of references in patents. This way patents are linked based on the shared references, in other words, based on the similarity of certain technological or application features of the patent claim. Other patent mapping approach is based on linking patents through the similarity of their classification codes that are specified on every granted patent document, so-called co-classification patent mapping approach. According to this approach, classes that occur frequently together are shown very close to each other on the map, whereas less related classes are displayed further apart [7]. Using co-classification based mapping, groups of related patent classes (from the viewpoint of technological classes or patent holders) aggregated to a guiding image of technology development can be identified [7]. Co-classification patent mapping allows for a global overview of the technological landscape and illustration of the cognitive structure of an area, therefore is more effective for studying the developments at a larger level since it may miss more specific local developments of the field important when analyzing smaller datasets [7]. Appendix 5 demonstrates one possible approach to couple patents based on their classification. Each patent is labeled by the patent number, and the link between the patents indicates the weighted amount of US classes they have in common, signifying the strength of the knowledge relationship between the patents (since as discussed earlier each patent can be classified into more than one patent class). A number of studies have also shown patent mapping from the viewpoint of patent holders (assignees) or countries to which those assignees belong [43]. Often these mapping approaches are based on CC analysis presented in the Literature review of this work. Besides the time delay associated with citations accumulation unavoidable with the CC analysis methodology as often mentioned in literature, another challenge is associated with the presence of double or even triple assignees of a patent. A second assignee might be, for example, a subsidiary of the first assignee, e.g. Hyundai Motor Company and Kia Motors often appear together as patent holders (Hyundai owns 32.8 % of Kia); or an organization that helps an innovating technology firm to achieve commercial success, e.g. Abell Foundation is a second patent holder of Paice Corporation; or just collaborating partners, e.g. GM, Daimler, Chrysler and BMW corporations hold several patents together. Therefor when the patent is referred to by the name of the first author in analogy to the scientific publication, it does not reveal all the other co-inventors who have equal rights for the invention. In our research work for the best transparency of the presented results I have displayed all the patent assignees when labeling patent on PCNs. This aspect of multiple assignees is not that often discussed in the literature dealing with the subject, however, it still remains a minor issue on the scale of the large data sample analyzed (e.g. in our dataset there are about 187 utility patents out of 3683 with more than one assignee per patent). The choice of patent mapping approach depends on the analysis purpose when patent mapping is applied as a support tool for strategic decision-making, e.g. analyzing the technology landscape, hot topics and trends of technological development, assessing the competitive environment and looking for potential partners based on their patenting activities, compare patenting within a certain technological area internationally and gain a deeper insight into the global technological changes or specific national legal changes etc.

Finally, a possible limitation of the patent mapping based research may originate from the data visualization and clustering techniques. According to Kessler, BC approach allows to group patents mechanically with no knowledge of science or judgment of content, automatically processing a large number of papers according to a rigorously defined criterion of coupling [49]. In order to aggregate patents into related groups to facilitate data visualization and interpretation, various clustering techniques can be applied. However, cluster analysis itself is an exploratory tool which may result in artifacts or instability of the cluster structure [122]. Mathematical functions on which clustering techniques are based can deal only with overall trends at the macro level [16]. The clustering method provides only candidates for the identification of technology areas, since it is not possible to determine a priori the amount of clusters within a given PCN [27]. In our case this refers to the resolution of the community detection algorithm used in Gephi network analysis software, the choice of which also depends on the sample size as described in the methodology section. Therefore validation of the results at different stages of the analysis for the qualitative assessment of cluster compositions is required for the interpretation of the clusters. A systematic approach that combines BC with a cluster method supported by experts' evaluations at several iterative stages was shown for science mapping and research front identification, for example by Jarneving [54]. The PCNs constructed for different years are also not homogeneous with respect to the number of patents, average number of citations per patent, etc. This causes additional implications for the formation and analysis of technology cluster developments over time.

Keeping the limitations of this and similar study explicitly in mind, there are great potential benefits of patent mapping analysis for its practical applicability in support of strategic decision making. Therefore, further research in this area is encouraged.

5 Conclusions and avenues for further research

Analysis of patents and patent citations hold intriguing possibilities as a data source for the support of strategic decision-making. In this study I touched upon only one aspect of this research direction. Application of complex network analysis allowed to reflect the linkages between the patented innovations based on the common prior art that they refer to. Thus technological relationships between patents that share a similar knowledge base or belong to a similar research front have been visualized. Labeling each node on the patent citation map by patent assignee demonstrated the association among firms that hold the patents and their active patenting position within certain technology domains on the global technology landscape in real time. This study indicates that BC can be a useful tool in exploring the relationships between patents, providing a glimpse of the direction of EV technology development and supporting business strategy formation of a firm including competition and collaboration possibilities. The usefulness of this analysis is in the ability to raise awareness of the patenting activities in the field and company's R&D focus at a particular time, and identify emerging technology areas before they are otherwise recognized. However, visualization-based patent analysis serves only as a starting point for gaining insights into the development of technologies. Further in-depth analysis based on the patents' content and expert interviews are important to examine the underlying technologies and to validate the results. This study also describes in detail the possible limitations of such patent analysis stemming from the chosen analysis methodology, approach of data acquisition and patent network construction and interpretation. Some of the limitations are the perplexities for the future study to overcome.

In this study patents are linked based on their shared references to other patents. Such analysis method is especially useful when analyzing smaller data samples and identifying specific details in the development of the field [7]. In our case of a larger sample size containing a broader technological area, an alternative approach could involve co-classification analysis as a starting point for pinpointing the technology development at a lower resolution, further followed by in-depth analysis on the level of individual patents within certain classification clusters based on BC. Future work could also explore text-mining methods for strengthening the results of cluster labeling. As an example, keyword vector mapping could be integrated into the current research framework as an additional layer to construct a comprehensive PCN based on BC [42].

The cognitive resemblance of bibliographically coupled technology clusters can be further examined by analyzing the citation frequency of shared references, adding an additional analysis layer. Earlier studies showed a significantly higher similarity of the word profile within the groups that share a citation to a highly cited publication than between documents without such a relationship [123].

The construction of PCN could be enhanced by analyzing the citation relationship within bibliographically coupled patent pairs as recently demonstrated by Yeh *et al.* [68]. The researchers developed a technique for the analysis of PCN, dividing the bibliographically coupled patent pairs into two groups: “BC pairs with citation” and “BC pairs without citation” This approach allowed to improve PCN, excluding less relevant bibliographically coupled pairs and adding the relevant uncited patent citations, overall making the network citation links of all technology clusters more concentrated and disclosing prior invisible technology clusters in the field [68]. The future study is proposed to examine the knowledge in-flow or knowledge out-flow between the BC patents using information on the citation relationship between patents as an additional analysis layer. BC method on its own considers only the undirected relation between the firms as a proxy for the amount of shared-ness of technological features [2].

The results of patent cluster analysis revealed that EV technology is multiplex and interpretation of some technology clusters is not obvious. This research was conducted at the macro level, considering the whole EV industry sector. Alternatively, a certain technology area within the emerging EV sector could be considered for the micro level analysis.

Using the current dataset the future study could attempt to perform a comparative analysis based on several clustering and graph partitioning algorithms. Hierarchical methods are preferred since they do not require the number of clusters within the network to be specified [27]. Structural relationships between patent clusters and their evolution over time could be analyzed by comparing the dendrogram structure resulting from hierarchical clustering methods. Using this approach Erdi *et. al.* have recently demonstrated a methodology based on the analysis of PCN, allowing to detect the new emerging cluster recombination within a certain technology field (technological branching) and to predict emerging new technology clusters [27]. Dynamic technology clusters could be analyzed from the perspective of patent assignees, i.e. observing firms that are new-comers to the industry, firms that exit or change the technology area of patenting activities, etc. The maturity of clusters could be evaluated by comparing companies or patent assignees that hold the patents within the cluster, i.e. the balance of patents between large/old and small/new companies.

The results of this study show that PCNs provide an objective understanding and visualization of where EV industry currently lies and what are the expectations for its development, which is useful for the support of strategic decision-making. We observe that currently the actors within the field of EV technology are betting on different competing technological designs and expectations, instead of focusing on one technological structure. This presents a typical case of existing uncertainties within an emerging technological field. Further examination of such open-endedness of expectations within the field is encouraged in collaboration with the field experts. In this respect, the scope of the data acquisition part of the analysis could be broadened, including e.g. fuel cell and hydrogen vehicles as well as battery-driven cars to allow for the comparison with alternative technologies as potentially competing platforms.

There are many other possible avenues for future research depending on the desired input in support of strategic decision-making. It may, for instance, be interesting to compare the behavior of the PCN internationally to gain an insight into the emerging technological areas and competitors worldwide. However, the effects of global technological change and those of nation-specific legal changes should be taken into account. Analyzing the change in technology within the EV sector starting from earlier time periods could be another interesting approach. In case of large data samples, the network size could be limited by the core documents, which have strong links to many other documents. The concept of core documents in the context of BC was introduced by Glänzel and Czerwon in 1996 [63, 70].

This research work raises awareness of a wide range of possibilities offered by PCNs based on BC, provides an input to the community's understanding of patent analysis to assist in strategic decision making, and encourages future research in this area.

Appendices

Appendix 1. The percentage share of patents in EV technology field with further details about section H based on the IPC classification system.

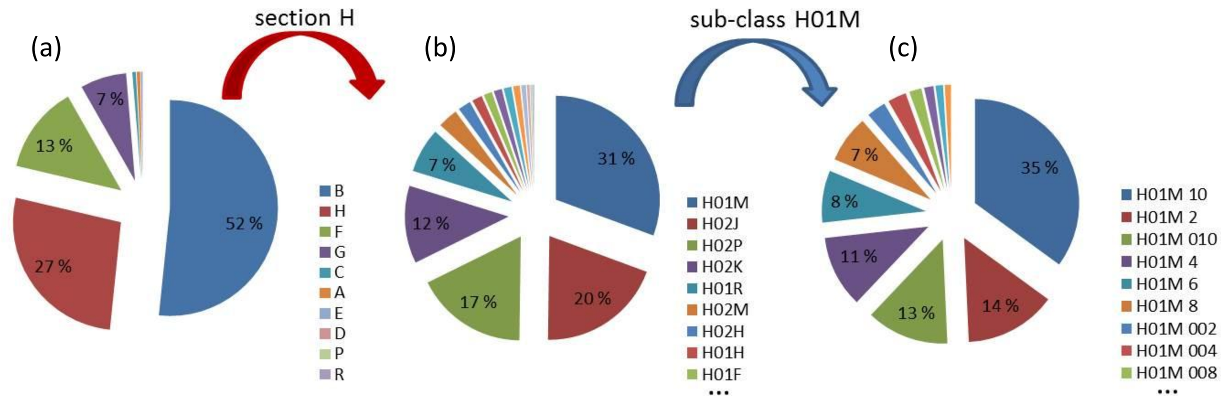


Figure A1: The percentage share of (a) patents in EV technology; (b) patent subclasses within section H; (c) patent groups within sub-class H01M. The main subclasses of section H ("Electricity") include battery (sub-class H01M, 31%) and battery management technology (sub-class H02J, 20%), motor (H02K, 12%) and motor control technology (H02P, 17%) as well as electrically-conductive connections technology (H01R, 7%). Within the largest sub-class H01M, the key domains belong to H01M 10, 35% (Secondary cells or accumulators receiving and supplying electrical energy by means of reversible electrochemical reactions, and their manufacturing).

Appendix 2. Patent application-grant lag distribution.

Table A2: Patent application-grant lag distribution in the field of EV for the period between 2007 and 2012 of patent filing years.

	Application years
	2007-2012
Lag	**Distribution of lags (in %)**
Up to 1 year	5,6
From 1 to 2 years	16,8
From 2 to 3 years	33,2
From 3 to 4 years	26,5
From 4 to 5 years	14,9
From 5 to 6 years	2,5
From 6 to 7 years	0,4
Total	100,
Number of patents	**1244**
Mean and standard deviation	
Mean	**2,93**
S.d.	13,74

Appendix 3. Distribution of normalized bibliographic coupling strength in EV technology for patents in three patent citation networks of this study.

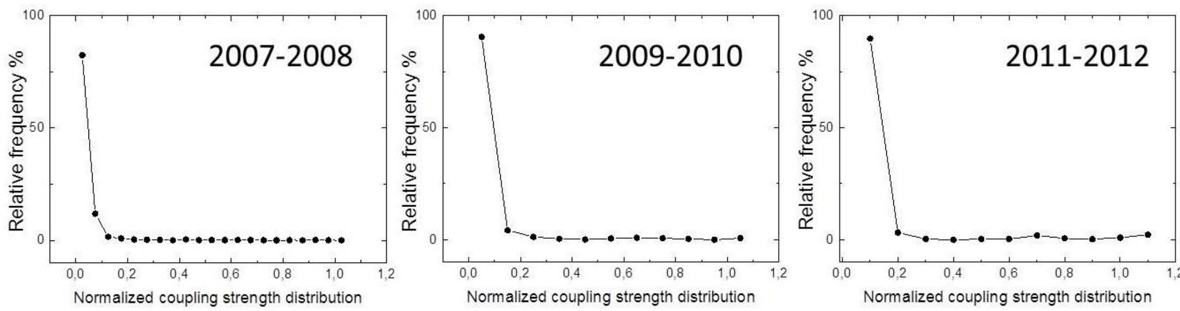


Figure A3: Normalized bibliographic coupling strength distribution corresponding to the patent filing years of 2007-2008, 2009-2010 and 2011-2012. It exhibits a Pareto distribution with high-frequency population followed by low-frequency population that gradually tails off asymptotically [67].

Appendix 4. Patent citation mapping based on bibliographic coupling analysis of patents filed between 2007 and 2012.

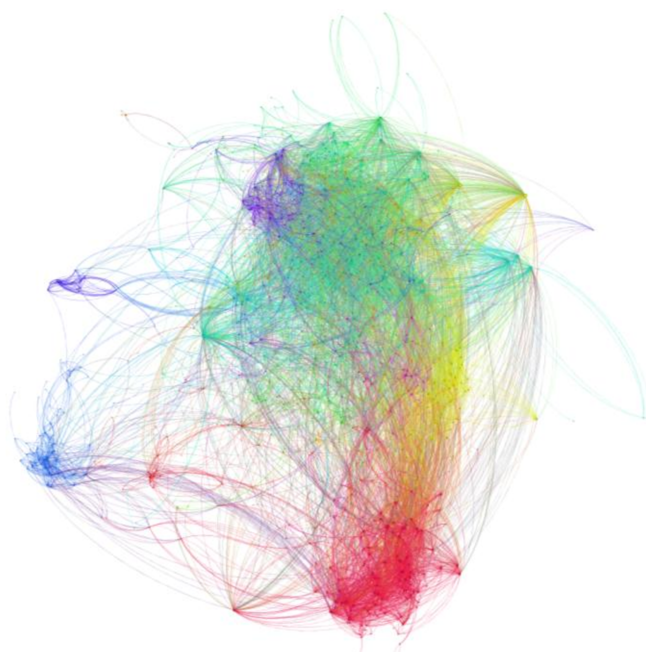


Figure A4: Patent citation mapping based on bibliographic coupling analysis of patents filed between 2007 and 2012. Gephi's 'Giant component' filter was applied. Nodes are ranked by their eigenvector centrality; patents are partitioned by modularity class (resolution of 5).

Appendix 5. Patent mapping based on co-classification approach.

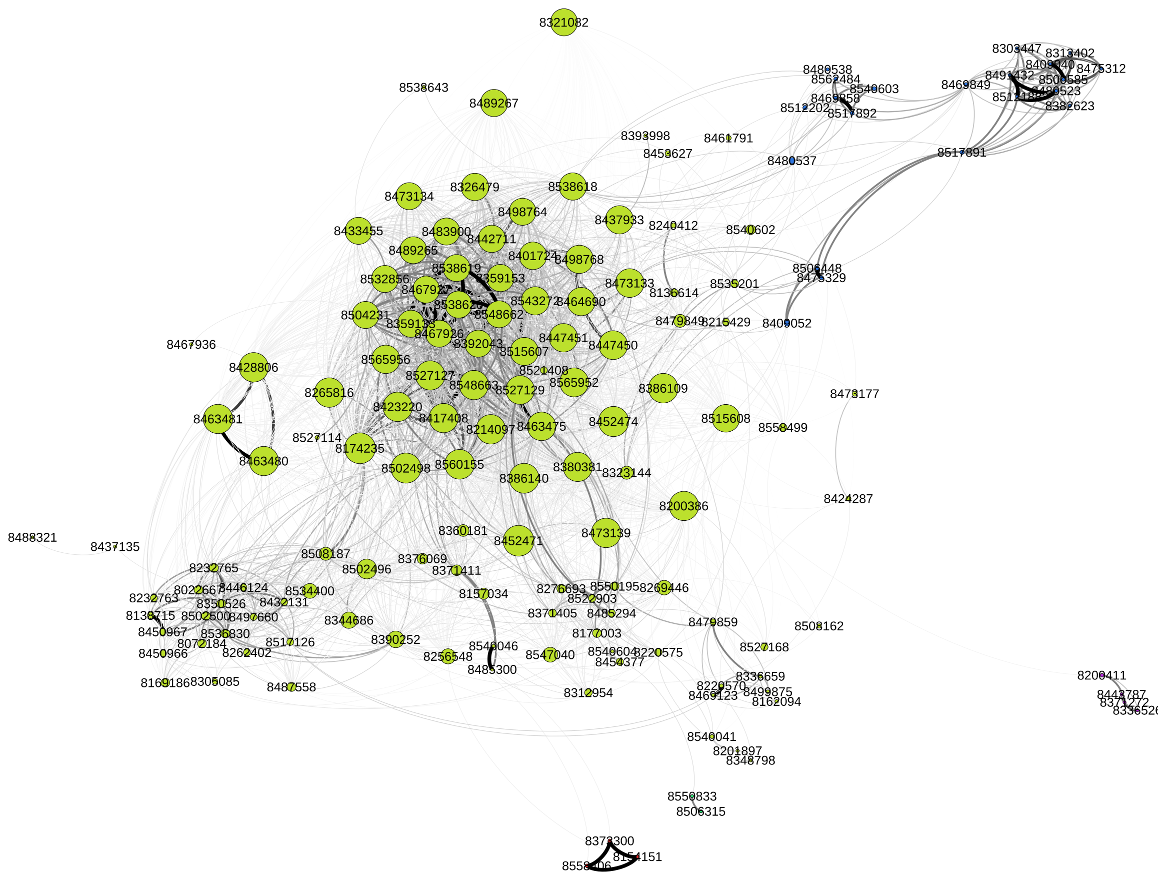


Figure A5: Visualization of the co-classification based coupling of patents filed between 2011 and 2012. Each node is labeled by the patent number, and two nodes are linked if they share a US classification code. Nodes are ranked by their eigenvector centrality, and edges are ranked by their weight (weight is calculated using Jaccard index). Patents are partitioned by modularity class (resolution of 5 in Gephi). The Gephi's 'Giant component' filter was applied for clarity. The patent map displays patents that have more common classification codes (in other words belong to a similar technological field) very close to each other, and less related patents are displayed further apart. After identifying the patents that are closely related in terms of the technological field (e.g. using weight filter etc), more in-depth analysis can be undertaken based on e.g. the bibliographic coupling of these patents.

References

1. Soos, S., *Age-sensitive bibliographic coupling with an application in the history of science.* Scientometrics, 2014. **98**(1): p. 23-51.

2. von Wartburg, I., T. Teichert, and K. Rost, *Inventive progress measured by multi-stage patent citation analysis.* Research Policy, 2005. **34**(10): p. 1591-1607.

3. Kuusi, O. and M. Meyer, *Anticipating technological breakthroughs: Using bibliographic coupling to explore the nanotubes paradigm.* Scientometrics, 2007. **70**(3): p. 759-777.

4. Boyack, K.W. and R. Klavans, *Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately?* Journal of the American Society for Information Science and Technology, 2010. **61**(12): p. 2389-2404.

5. Schiebel, E., *Visualization of research fronts and knowledge bases by three-dimensional areal densities of bibliographically coupled publications and co-citations.* Scientometrics, 2012. **91**(2): p. 557-566.

6. Chang, S.-B., *Using patent analysis to establish technological position: Two different strategic approaches.* Technological Forecasting and Social Change, 2012. **79**(1): p. 3-15.

7. Gustafsson, R., O. Kuusi, and M. Meyer, *Examining foundations of competing guiding images in emerging technological fields: the case of cellulosic ethanol.* Technological Forecasting and Social Change, Available online 6 March 2014.

8. Ma, R., *Author bibliographic coupling analysis: A test based on a Chinese academic database.* Journal of Informetrics, 2012. **6**(4): p. 532-542.

9. Pilkington, A., R. Dyerson, and O. Tissier, *The electric vehicle: patent data as indicators of technological development.* World Patent Information, 2002. **24**(1): p. 5-12.

10. Scherer, F.M., *Firm size, market structure, opportunity, and the output of patented inventions.* The American Economic Review, 1965. **55**(5): p. 1097-1125.

11. Schmookler, J., *Invention and economic growth*. Vol. 26. 1966: Harvard University Press Cambridge, MA.

12. Griliches, Z., *Patent statistics as economic indicators: a survey*, in *R&D and productivity: the econometric evidence*. 1998, University of Chicago Press. p. 287-343.

13. Basberg, B.L., *Patents and the measurement of technological change: a survey of the literature.* Research policy, 1987. **16**(2): p. 131-141.

14. Hall, B.H., A.B. Jaffe, and M. Trajtenberg, *Market value and patent citations: A first look*. 2000, National Bureau of Economic Research Working Paper No. 7741.

15. Tang, J., et al. *Patentminer: Topic-driven patent analysis and mining*. in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2012. ACM.

16. Choi, C. and Y. Park, *Monitoring the organic structure of technology based on the patent development paths.* Technological Forecasting and Social Change, 2009. **76**(6): p. 754-768.

17. Narin, F., E. Noma, and R. Perry, *Patents as indicators of corporate technological strength.* Research policy, 1987. **16**(2): p. 143-155.

18. Mogee, M.E., *Using patent data for technology analysis and planning.* Research-Technology Management, 1991. **34**(4): p. 43-49.

19. Huang, M.-H., L.-Y. Chiang, and D.-Z. Chen, *Constructing a patent citation map using bibliographic coupling: A study of Taiwan's high-tech companies.* Scientometrics, 2003. **58**(3): p. 489-506.

20. Jaffe, A.B. and M. Trajtenberg, *Patents, citations, and innovations: A window on the knowledge economy*. 2002: MIT press.

21. Boucher, P.M., et al., *Recent developments in US patent law.* Physics Today, 2012. **65**(1): p. 27.

22. Meyer, M., *What is special about patent citations? Differences between scientific and patent citations.* Scientometrics, 2000. **49**(1): p. 93-123.

23. Daim, T.U., et al., *Forecasting emerging technologies: Use of bibliometrics and patent analysis.* Technological Forecasting and Social Change, 2006. **73**(8): p. 981-1012.

24. Watanabe, C., Y.S. Tsuji, and C. Griffy-Brown, *Patent statistics: deciphering a 'real'versus a 'pseudo'proxy of innovation.* Technovation, 2001. **21**(12): p. 783-790.

25. Lee, C., et al., *A stochastic patent citation analysis approach to assessing future technological impacts.* Technological Forecasting and Social Change, 2012. **79**(1): p. 16-29.

26. Almeida, P. and A. Phene, *Subsidiaries and knowledge creation: The influence of the MNC and host country on innovation.* Strategic Management Journal, 2004. **25**(8-9): p. 847-864.

27. Érdi, P., et al., *Prediction of emerging technologies based on analysis of the US patent citation network.* Scientometrics, 2013: p. 1-18.

28. Trajtenberg, M., *A penny for your quotes: patent citations and the value of innovations.* The Rand Journal of Economics, 1990: p. 172-187.

29. Breitzman, A. and P. Thomas, *Using patent citation analysis to target/value ma candidates.* Research-Technology Management, 2002. **45**(5): p. 28-36.

30. Harhoff, D., et al., *Citation frequency and the value of patented inventions.* Review of Economics and statistics, 1999. **81**(3): p. 511-515.

31. Gay, C. and C. Le Bas, *Uses without too many abuses of patent citations or the simple economics of patent citations as a measure of value and flows of knowledge*.* Economics of Innovation and New Technology, 2005. **14**(5): p. 333-338.

32. Harhoff, D., F.M. Scherer, and K. Vopel, *Citations, family size, opposition and the value of patent rights.* Research Policy, 2003. **32**(8): p. 1343-1363.

33. Hagedoorn, J. and M. Cloodt, *Measuring innovative performance: is there an advantage in using multiple indicators?* Research Policy, 2003. **32**(8): p. 1365-1379.

34. Jaffe, A.B., M. Trajtenberg, and R. Henderson, *Geographic localization of knowledge spillovers as evidenced by patent citations.* The Quarterly Journal of Economics, 1993. **108**(3): p. 577-598.

35. Moretti, E., *Workers education, spillovers, and productivity: Evidence from plant-level production functions.* The American Economic Review, 2004. **94**(3): p. 656-690.

36. Roach, M. and W.M. Cohen, *Lens or Prism? Patent Citations as a Measure of Knowledge Flows from Public Research.* Management Science, 2013. **59**(2): p. 504-525.

37. Alcacer, J. and M. Gittelman, *Patent citations as a measure of knowledge flows: The influence of examiner citations.* The Review of Economics and Statistics, 2006. **88**(4): p. 774-779.

38. Almeida, P., *Knowledge sourcing by foreign multinationals: patent citation analysis in the US semiconductor industry.* Strategic Management Journal, 1996. **17**: p. 155-165.

39. Podolny, J.M., T.E. Stuart, and M.T. Hannan, *Networks, knowledge, and niches: Competition in the worldwide semiconductor industry, 1984-1991.* American journal of sociology, 1996: p. 659-689.

40. Marco, A.C., *The option value of patent litigation: Theory and evidence.* Review of Financial Economics, 2005. **14**(3): p. 323-351.

41. Chang, S.-B., K.-K. Lai, and S.-M. Chang, *Exploring technology diffusion and classification of business methods: Using the patent citation network.* Technological Forecasting and Social Change, 2009. **76**(1): p. 107-117.

42. Park, I. and B. Yoon, *Identifying Potential Partnership for Open Innovation by using Bibliographic Coupling and Keyword Vector Mapping.* World Academy of Science, Engineering and Technology, 2013. **74**.

43. Choe, H., et al., *Patent citation network analysis for the domain of organic photovoltaic cells: Country, institution, and technology field.* Renewable and Sustainable Energy Reviews, 2013. **26**: p. 492-505.

44. Stuart, T.E., *Network positions and propensities to collaborate: An investigation of strategic alliance formation in a high-technology industry.* Administrative Science Quarterly, 1998: p. 668-698.

45. Sternitzke, C., A. Bartkowski, and R. Schramm, *Visualizing patent statistics by means of social network analysis tools.* World Patent Information, 2008. **30**(2): p. 115-131.

46. Verspagen, B., *Mapping technological trajectories as patent citation networks: A study on the history of fuel cell research.* Advances in Complex Systems, 2007. **10**(01): p. 93-115.

47. No, H.J. and Y. Park, *Trajectory patterns of technology fusion: Trend analysis and taxonomical grouping in nanobiotechnology.* Technological Forecasting and Social Change, 2010. **77**(1): p. 63-75.

48. Kessler, M.M., *An experimental study of bibliographic coupling between technical papers*. 1962, Massachusetts Institute for Technology, Lincoln Laboratory.

49. Kessler, M.M., *Bibliographic coupling between scientific papers.* American documentation, 1963. **14**(1): p. 10-25.

50. Kessler, M., *Concerning some problems of intrascience communication*. 1958: Massachusetts Institute of Technology, Lincoln Laboratory.

51. Kessler, M.M., *Bibliographic coupling extended in time: ten case histories.* Information storage and retrieval, 1963. **1**(4): p. 169-187.

52. Cleverdon, C.W., J. Mills, and M. Keen, *Factors determining the performance of indexing systems.* 1966.

53. Sen, S. and S. Gan, *A mathematical extension of the idea of bibliographic coupling and its applications.* Annals of Library Science and Documentation, 1983. **30**(2): p. 78-82.

54. Jarneving, B., *Bibliographic coupling and its application to research-front and other core documents.* Journal of Informetrics, 2007. **1**(4): p. 287-307.

55. Glänzel, W. and H.-J. Czerwon. *A new methodological approach to bibliographic coupling and its application to research-front and other core documents*. *in International Society for Scientometrics and Informetrics. International conference proceedings.* 1995.

56. Weinberg, B.H., *Bibliographic coupling: A review.* Information Storage and Retrieval, 1974. **10**(5): p. 189-196.

57. Vladutz, G. and J. Cook. *Bibliographic coupling and subject relatedness*. in *Challenges to an information society. Proceedings of the 47th ASIS annual meeting. White Plains: Knowledge Industry Publications*. 1984.

58. Jarneving, B., *A comparison of two bibliometric methods for mapping of the research front.* Scientometrics, 2005. **65**(2): p. 245-263.

59. Morris, S.A., et al., *Time line visualization of research fronts.* Journal of the American society for information science and technology, 2003. **54**(5): p. 413-422.

60. Boyack, K.W., *Using detailed maps of science to identify potential collaborations.* Scientometrics, 2009. **79**(1): p. 27-44.

61. Nicolaisen, J. and T.F. Frandsen, *Consensus formation in science modeled by aggregated bibliographic coupling.* Journal of Informetrics, 2012. **6**(2): p. 276-284.

62. Soós, S., *Age-sensitive bibliographic coupling reflecting the history of science: The case of the Species Problem.* Scientometrics, 2014. **98**(1): p. 23-51.

63. Glänzel, W. and H.-J. Czerwon, *A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level.* Scientometrics, 1996. **37**(2): p. 195-221.

64. Glänzel, W., *Bibliometrics as a research field.* K.U. Leuven, F.E.T.E.W.: Leuven, 2003. **25**: p. 115.

65. Chen, S.-H., et al., *Detecting the temporal gaps of technology fronts: A case study of smart grid field.* Technological Forecasting and Social Change, 2012.

66. Small, H., *Co-citation in the scientific literature: A new measure of the relationship between two documents.* Journal of the American Society for information Science, 1973. **24**(4): p. 265-269.

67. Chen, D.-Z., et al., *Identifying missing relevant patent citation links by using bibliographic coupling in LED illuminating technology.* Journal of Informetrics, 2011. **5**(3): p. 400-412.

68. Yeh, H.-Y., et al., *The bibliographic coupling approach to filter the cited and uncited patent citations: a case of electric vehicle technology.* Scientometrics, 2013. **94**(1): p. 75-93.

69. Huang, M.-H., D.-Z. Chen, and H.-R. Dong. *Identify technology main paths by adding missing citations using bibliographic coupling and co-citation methods in photovoltaics*. in *Technology Management in the Energy Smart World (PICMET), 2011 Proceedings of PICMET'11*. 2011. IEEE.

70. Glänzel, W. and B. Thijs, *Using 'core documents' for the representation of clusters and topics.* Scientometrics, 2011. **88**(1): p. 297-309.

71. Brown, S., D. Pyke, and P. Steenhof, *Electric vehicles: The role and importance of standards in an emerging market.* Energy Policy, 2010. **38**(7): p. 3797-3806.

72. Yang, L.-f., J.-h. Xu, and P. Neuhäusler, *Electric vehicle technology in China: An exploratory patent analysis.* World Patent Information, 2013. **35**(4): p. 305-312.

73. Harrop, P. and R. Das, *Hybrid and pure electric vehicles for land, water and air 2013-2023: forecasts, technologies, players*. 2013: http://www.idtechex.com/.

74. Valentine-Urbschat, M. and W. Bernhart. *Powertrain 2020–Challenges and opportunities for OEMs and suppliers*. in *Proc. The International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium*. 2009.

75. Zhou, L., et al. *Charging ahead: battery electric vehicles and the transformation of an industry*. Deloitte Review, 2010.

76. Porter, A.L. and N.C. Newman, *Patent profiling for competitive advantage*, in *Handbook of Quantitative Science and Technology Research*. 2005, Springer. p. 587-612.

77. U.S. Patent and Trademark Office, T.P.T.M.T. *U.S. Patent Statistics Report*. (accessed in December 2013).

78. Meyer, M., J.T. Utecht, and T. Goloubeva, *Free patent information as a resource for policy analysis.* World Patent Information, 2003. **25**(3): p. 223-231.

79. Narin, F. and E. Noma, *Is technology becoming science?* Scientometrics, 1985. **7**(3): p. 369-381.

80. He, Z.-L. and M. Deng, *The evidence of systematic noise in non-patent references: A study of New Zealand companies' patents.* Scientometrics, 2007. **72**(1): p. 149-166.

81. Patel, P., *Localised production of technology for global markets.* Cambridge Journal of Economics, 1995. **19**(1): p. 141-153.

82. Leydesdorff, L., D. Kushnir, and I. Rafols, *Interactive overlay maps for US patent (USPTO) data based on International Patent Classification (IPC).* Scientometrics, 2012: p. 1-17.

83. Meyer, M., *Between technology and science: exploring an emerging field*. 2003: Universal-Publishers.

84. Kang, I.-S., et al., *Cluster-based patent retrieval.* Information processing & management, 2007. **43**(5): p. 1173-1182.

85. Chiu, T.-F., *A proposed IPC-based clustering method for exploiting expert knowledge and its application to strategic planning.* Journal of Information Science, 2013: p. 0165551513507404.

86. Hong, S., *The magic of patent information.* World Intellectual Property Organization, www. wipo. int/sme/en/documents/patent_information. htm, 2007.

87. Simmons, E.S., *Trends disrupted—patent information in an era of change.* World Patent Information, 2005. **27**(4): p. 292-301.

88. Guellec, D., M. Zuniga, and C. Martinez, *Blocking patents: What they are and what they do*. 2008, OECD No., Paris.

89. Nagaoka, S., *Assessing the R&D management of a firm in terms of speed and science linkage: evidence from the US patents.* Journal of Economics & Management Strategy, 2007. **16**(1): p. 129-156.

90. Cherven, K., *Network Graph Analysis and Visualization with Gephi*. 2013: Packt Publishing Ltd.

91. Bastian, M., S. Heymann, and M. Jacomy. *Gephi: an open source software for exploring and manipulating networks*. in *ICWSM*. 2009.

92. Jacomy, M., et al., *ForceAtlas2, A Continuous Graph Layout Algorithm for Handy Network Visualization.* in preparation (webatlas.fr/tempshare/ForceAtlas2_Paper.pdf), 2012.

93. Kim, C., et al., *Identifying core technologies based on technological cross-impacts: An association rule mining (ARM) and analytic network process (ANP) approach.* Expert Systems with Applications, 2011. **38**(10): p. 12559-12564.

94. Leydesdorff, L., *On the normalization and visualization of author co-citation data: Salton's Cosine versus the Jaccard index.* Journal of the American Society for Information Science and Technology, 2008. **59**(1): p. 77-85.

95. Schneider, J.W. and P. Borlund, *Matrix comparison, Part 1: Motivation and important issues for measuring the resemblance between proximity measures or ordination results.* Journal of the American Society for Information Science and Technology, 2007. **58**(11): p. 1586-1595.

96. Schneider, J.W. and P. Borlund, *Matrix comparison, Part 2: Measuring the resemblance between proximity measures or ordination results by use of the Mantel and Procrustes*

statistics. Journal of the American Society for Information Science and Technology, 2007. **58**(11): p. 1596-1609.

97. Hamers, L., et al., *Similarity measures in scientometric research: the Jaccard index versus Salton's cosine formula.* Information Processing & Management, 1989. **25**(3): p. 315-318.

98. Wasserman, S., *Social network analysis: Methods and applications*. Vol. 8. 1994: Cambridge university press.

99. Borgatti, S.P. and M.G. Everett, *A graph-theoretic perspective on centrality.* Social networks, 2006. **28**(4): p. 466-484.

100. Bonacich, P., *Some unique properties of eigenvector centrality.* Social Networks, 2007. **29**(4): p. 555-564.

101. Valdez, A.C., et al., *Using mixed node publication network graphs for analyzing success in interdisciplinary teams*, in *Active Media Technology*. 2012, Springer. p. 606-617.

102. Newman, M.E., *A measure of betweenness centrality based on random walks.* Social networks, 2005. **27**(1): p. 39-54.

103. Hung, S.-W. and A.-P. Wang, *Examining the small world phenomenon in the patent citation network: a case study of the radio frequency identification (RFID) network.* Scientometrics, 2010. **82**(1): p. 121-134.

104. Powell, W.W., et al., *Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences1.* American journal of sociology, 2005. **110**(4): p. 1132-1205.

105. Jacomy, M., *Force-Atlas Graph Layout Algorithm.* URL: http://gephi.org/2011/forceatlas2-the-new-version-of-our-home-brew-layout, 2009.

106. Paranyushkin, D., *Identifying the Pathways for Meaning Circulation using Text Network Analysis.* Available online: noduslabs.com/research/pathways-meaning-circulation-text-network-analysis, 2011.

107. Fortunato, S., *Community detection in graphs.* Physics Reports, 2010. **486**(3): p. 75-174.

108. Blondel, V.D., et al., *Fast unfolding of communities in large networks.* Journal of Statistical Mechanics: Theory and Experiment, 2008. **2008**(10): p. P10008.

109. *Gephi Tutorial Visualization*. http://gephi.org/tutorials/gephi-tutorial-visualization.pdf, Last updated May 26th, 2010.

110. Lambiotte, R., J.-C. Delvenne, and M. Barahona, *Laplacian dynamics and multiscale modular structure in networks.* arXiv preprint arXiv:0812.1770, 2008.

111. Feinberg, J. and I. Corp. *www.wordle.net*. 2009 (accessed in January 2014).

112. McNaught, C. and P. Lam, *Using Wordle as a supplementary research tool.* The qualitative report, 2010. **15**(3): p. 630-643.

113. Cook, J. *US Patent Classification System, www.intellogist.com/wiki/US_Patent_Classification_System*. 2009 (accessed in Januray 2014).

114. van Wijk, B.C.M., C.J. Stam, and A. Daffertshofer, *Comparing brain networks of different size and connectivity density using graph theory.* PLoS One, 2010. **5**(10): p. e13701.

115. Li, X., et al., *Patent citation network in nanotechnology (1976–2004).* Journal of Nanoparticle Research, 2007. **9**(3): p. 337-352.

116. Criscuolo, P., *The'home advantage'effect and patent families. A comparison of OECD triadic patents, the USPTO and the EPO.* Scientometrics, 2006. **66**(1): p. 23-41.

117. Czarnitzki, D., K. Hussinger, and B. Leten, *The market value of blocking patent citations.* Available at SSRN 1868847, 2011.

118. Webb, C., et al., *Analysing European and International Patent Citations*. 2005: OECD Science, Technology and Industry working papers.

119. Trippe, T., *All citations are not the same - exploring examiner citations from US patent documents*. 2013, Patent strategy and analytics advisors, Patinformatics, LLC: www.patinformatics.com.

120. Czarnitzki, D., K. Hussinger, and B. Leten. *Improving citation-based patent value measures*. in *Management of Innovation and Technology (ICMIT), 2012 IEEE International Conference on*. 2012. IEEE.

121. Hong, S. *The Magic of Patent Information*. Available at www.wipo.int/sme/en/documents/patent_information_fulltext.html (accessed in Januray 2014).

122. Aldenderfer, M.S. and R.K. Blashfield, *Cluster Analysis*. Beverly Hills, CA: Sage Press, 1984

123. Peters, H.P.F., R.R. Braam, and A.F.J. van Raan, *Cognitive resemblance and citation relations in chemical engineering publications.* Journal of the American Society for Information Science, 1995. **46**(1): p. 9-21.