The co-evolution of knowledge and collaboration networks: the role of the technology life-cycle

Article *in* Scientometrics · July 2016

DOI: 10.1007/s11192-017-2579-4

2 authors:

Johannes van der Pol
VIA Inno University of Bordeaux

Rameshkoumar Jean-Paul
University of Bordeaux

The co-evolution of knowledge and collaboration networks: the role of the technology life-cycle

Johannes van der Pol & Jean-Paul Rameshkoumar

June 2, 2016

Abstract

The aim of this paper is twofold. The paper shows that network analysis can be used for the identification of the technology life cycle. In addition we show that there is a correlation between the evolution of the knowledge network and the collaboration network of Structural Composite Materials (SCM) in Aeronautics. First, we analyze the structure of the knowledge network and how it evolves over time. We identify two stages; a first characterized by a continuous increase in clustering showing the formation of the core of the technology. A second stage shows a decrease in clustering and increase in average distance signaling the formation of a periphery around the core of the technology. The overall structure tends towards a core-periphery structure.

Second, we analyze the structure of the collaboration network. We identify a rupture in the structure around the same year as in the knowledge network. The network structure converges to a small world structure from that point on.

Taken together the results show that there is a correlation between the evolution of the structure of the knowledge network and the collaboration network. The correlation coincides with the life-cycle of the technology.

Keywords: Patents; Publications; Network analysis; Small world; Scale-free; Technology Life Cycle; knowledge network; IPC analysis

1 Introduction

Innovations are the driving force behind growing economies and prosperous firms. Achieving innovation is hence at the center of any business strategy. With the growing complexity of new technologies, a single firm can no longer master all the technologies needed for the production of a single product. Accessing knowledge held by other firms becomes of vital importance for firms in order to innovate. Accessing different knowledge sources has proven to be beneficial for the firm (McEvily & Marcus, 2005), for innovation (Kogut & Zander, 1992),(Tsai, 2001) as well as survival and growth (Watson, 2007).

Empiric analyses show that the number of cooperations has increased steadily over the last two decades (paolo Saviotti, 2007). During this same period we can observe that cooperation has evolved from (predominantly) dyadic cooperation to multilateral cooperations. Firms evolve in networks, an interconnection of cooperating firms with a common goal (Pippel, 2013).

Knowledge flows between firms, allowing them to learn from one another. The manner in which firms benefit from the network depends in part on the structure of the network since the structure defines how fast and efficient knowledge flows through the network (Verspagen & Duysters, 2004a). In turn, the diffusion of innovations impact the performance of the firms inside the network. The structure of these networks is hence a variable of paramount importance and has attracted much attention (Ahuja, 2000; Verspagen & Duysters, 2004a; Buchmann & Pyka, 2013; van der Valk *et al.* , 2011). These efforts have resulted in the identification of several network typologies, mainly small worlds(Watts, 1999; Gulati *et al.* , 2012; Tomasello *et al.* , 2013; Baum *et al.* , 2003), scale free networks,(Barabási & Albert, 1999; Van Der Pol, 2015) and nested split graphs (König *et al.* , 2009). The structure taken by a network is highly impacted by its context. Factors such as industry (Salavisa *et al.* , 2012), types of actors included (Nieto & Santamaría, 2007) as well as geography (McKelvey *et al.* , 2003) have shown to have an important impact. However, the role played by the technology life-cycle is still mostly unexplored (Stolwijk *et al.* , 2013). We could expect that different stages of the technology life-cycle require collaborations with different firms or research institutes. A research stage call for basic research and knowledge about fundamental technologies while the development phase requires collaborations with firms that have a more applied approach to the technology. We extend the existing literature on network formation by showing how the life-cycle of the technology impacts the formation of the collaboration network. In addition we extend the existing literature on technology life-cycles by showing that networks can be used to identify the different stages of the life-cycle.

2 Literature and hypothesis

2.1 The technology life-cycle

We aim at identifying a correlation between the evolution of the collaboration network and the technology life-cycle. We use IPC codes present on patents as a proxy for the identification of the evolution of the life-cycle of the technology. Patent and publication data are widely used for the analysis of the technology life-cycle (Alencar *et al.* , 2007; Gao *et al.* , 2013; Trappey *et al.* , 2013), however not in network form. We depart from the Schumpeterian perspective that innovations are the result of the recombination of existing technologies. International Patent Classification codes (IPC codes) present on patents can be used as a proxy for "technology blocks". The presence of different codes on the same patent bares witness to a recombination of technologies. By taking all the IPC codes on patents a network is created.

During the first stage of the life-cycle, fundamental knowledge about the technology is researched. The deposited patents will be characerised by a relatively small number of IPC codes. The patents deposited during this phase form the core of the technology. The low number of codes and high number of deposits result in an IPC network that will be very densely interconnected. From a dynamic perspective we should observe a continuous increase in the clustering coefficient of the IPC network over time.

Hypothesis 1: *The clustering coefficient of the IPC network will increase continuously during the research stage.*

As the technology moves from the research phase to the development phase, patents deposited for incremental innovations combine the core technology with new fields of application. This results in the inclusion of new technology blocks in the network.

The developments on the core technology are relative to a specific application of the technology. We can take the example of a photo camera. The core technology is the camera, an application would be the integration of the camera into laptops, phones and watches. Each application has a specific research direction. As a result, the new blocks connect to the core but only scarcely connect between them. From a structural point of view new nodes are added to the network, creating a periphery around the core created by the research stage. This bring us to our second hypothesis:

Hypothesis 2: *During the development phase, the clustering coefficient decreases continuously due to the addition of nodes in the periphery. This has the associated result of increasing the average distance in the network.*

As the development stage evolves the periphery develops while at the same time reinforcing the core of the technology. This leads us to our third hypothesis:

Hypothesis 3: *The knowledge network has the structure of a core-periphery network.*

2.2 The collaboration network

One of the reasons networks are important for innovation is the hypothesis of knowledge spillovers. During collaborations firms exchange knowledge. The structure of the network affects how efficiently this knowledge flows through the network. From a policy point of view, understanding the structure of a network can help the optimization of R&D investments by reducing redundancy and optimizing knowledge diffusion. Investing in firms or research institutions with a particular (more central) position can result in a faster, more efficient level of knowledge diffusion.

Understanding the structure of a network is highly dependent upon the environment in which it evolves. The collaboration network of the aerospace sector does not have the same structure as the collaboration network of the biotech sector. The first is based on a highly optimized production chain, while the second is a highly competitive horizontal sector. The focus of this paper is on a collaboration network at the level of one specific technology (SCM) inside an existing sector. At this level the life-cycle of the technology intervenes as a defining factor in the structure while it does not at the level of the sector. Since many life-cycles evolve continuously at different stages this would be difficult to track.

We follow the definition of a technology life-cycle based on two stages, a research stage and a development stage (Davide Chiaroni, 2008; Rowley *et al.* , 2000; Virapin & Flamand, 2013) . The first stage is characterized by a research phase in which fundamental knowledge is required to create a new technology. This stage requires collaborations with agents that possess fundament knowledge and are able to conduct basic research. Once the technology has been stabilized, the development stage begins. During this stage, firms start to apply and develop their technology for different applications. During this phase new collaborations are required with other agents with new abilities in other fields. Different applications for the technologies will be developed by different clusters of firms each with their own specialization.

A dense interconnection of firms for the basic research of the technology is then expected to appear in the early stages of the network. As times goes by, clusters of firms developing applications for the technology will connect to the initial cluster of collaborations during the second phase. This gives us our fourth hypothesis.

Hypothesis 4: *The structure of the collaboration network converges towards a small world structure.*

Since the types of collaborations change around the same time as the stages of the life-cycle we formulate the following, final hypothesis:

Hypothesis 5: *The structure of the collaboration network is correlated with the life-cycle of the technology (regardless of the structure of the collaboration network).*

3 Data and methodology

3.1 Structural Composite Materials

Structural Composite Materials (SCM) were first developed by chemists in the early 20th century and have since been used in sport equipment and the automotive industry (Virapin & Flamand, 2013). It caught the attention of civil aircraft manufacturers during the late 70's. During this period, research programs focusing on the optimization of energy consumption were launched by the European Union and the american government. The aim of these programs was to exploit composite materials in order to increase energy efficiency for aircrafts by the means of weight reduction. This makes SCM the perfect candidate for a study to analyze how a network is structured in order to absorb an existing technology from other sectors and develop it for its own needs.

The aerospace sector has a particular structure, it is organized as a production chain. An aircraft being a multi-technological product, each part of the airplane is developed in a different part of the network.

In the value chain that makes up the sector, a small number of firms occupies a strategic (central) position, these firms assemble intermediary products before sending them to either the final assembler (Airbus, Boeing etc.) or to other firms that use intermediary goods for larger parts. Firms with these specific positions in the value chain are called "pivot firms" (Frigant *et al.* , 2006). These firms have to master all the technologies of the downstream firms in order to complete their part of the aircraft Van Der Pol *et al.* (2014).

The introduction of a new technology such as SCM, can only succeed if the value chain adapts to the technology. Indeed, the introduction of SCM alter the structure of an aircraft in many dimensions. Pivot firms have to adjust their production methods and hence so do the downstream firms. Integrating SCM in the aerospace industry hence implies a thorough understanding of the core and linkage technologies (Prencipe, 1997) by all the actors implicated in aeronautical programs.

We used patent and publication data to generate our networks. Patents were extracted from Orbit while publications were extracted from Web Of Science (there were no geographical restrictions).

In order to extract all relevant patents and publication we started by framing the technologies involved in the production of SCM. The framing process is an iterative process based on discussions with engineers and executives from the aerospace sector.

We conducted an initial search for relevant IPC codes by identifying parts of the aircraft that can be made out of SCM. A detailed search was then conducted (combing IPC codes and key-words) in order to identify which specific products and technologies are involved in the creation of composite materials (resins, matrices). We then discussed these codes and key-words with engineers who would confirm or infirm the relevance. New codes and key-words were identified based upon these discussions and then discussed again. This iterative process allows us to frame the technology and build up a query that extracts patents beyond the scope of keywords which would

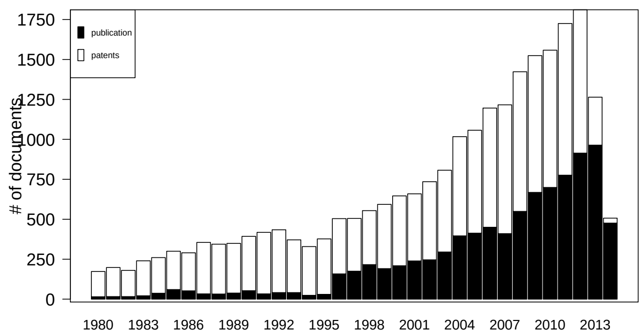


Figure 1: Distribution of the number of patents and publications between 1980 and 2014

result in false positives and unidentified patents and publications. The final query includes both relevant IPC codes and keywords.

A total number of 15313 patents and 9030 publications were identified worldwide between 1980 and 2014. The analysis was initiated in 1980 since it is the point at which SCM caught the attention of aircraft manufacturers. We checked patents and publication before 1980 and confirmed the latter. The results in figure 1 show the evolution of the number of patents and publications identified.

The Orbit database uses algorithms to extract and translate data from patents, this results in terms that get lost in translation and textual mistakes. In addition to this, names on patents often do not match names on publication. For example, we observe the name "Airbus SA" on a patent, "Airbus S.A". on a publication, even mistakes like "Aerhjbus" appear in the data. We used Intellixir to clean any mismatches or textual mistakes. Intellixir is a browser based tool for patent and publication analysis. It has a feature that automatically cleans the data. A final treatment was conducted by hand to clean any remaining problems.

3.2 Methodology

3.2.1 Core-periphery identification

A CP network has a small number of highly connected nodes (the core) and a large number of (relatively) less connected nodes (the periphery). In other words, it is an interconnection of hubs. By representing the network by a Cumulative Frequency Distribution of the number of links we can visualise a network and check for a Core-Periphery structure. A CFD is simply a plot with the frequency of nodes with degree k on the y-axis and the degree on the x-axis. This distribution is then transforms into a cumulative degree distribution as can be seen in figure 2. Figure 2 represents the CFD of the IPC network. One can see that roughly 95% of all nodes have at least two links, 80% has at least 3 links, and so on. If the frequency decreases by a small factor between densities, only a few nodes are lost due to the increase in density. If this CFD has the form of a line then when the degree increases by one, the frequency decreases by a fixed factor. The network is then called a scale-free network (since the diminishing factor is constant). This most sought-after (and found) structure is represented by a power law which has the form: $p(k) = c \cdot k^{-\alpha}$. We also check for another form which is the log-normal function ($ln(k) = \frac{1}{k}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{ln(k)-\mu}{\sigma})}$).

The main difference between the two functions is that they do not represent the same type of network. The scale-free network is a particular form of a core-periphery structure in which the frequency decrease is constant. In other words, when the density increases by one, the frequency drops by a factor k for all values of the density (it has hence the form of a straight line). Other functions such as the log-linear function can have more of a curvature to them. If the function is concave (figure 2) the drop in frequency increases with each increase in the density. The periphery of such a network contains less nodes with a low density. The periphery is less more interconnected than the scale-free network. The inverse would be true if we were to have a convexe function. The shape of the adjusted function informs us about the type of core-periphery structure, ranging from sparse to dense.

In order to conclude to a core-periphery structure we fit a particular function to the data. The functions are fitted using a maximum likelihood estimation. We then use a bootstrapping method in order to assess the goodness of fit which provides us with a p.value. The null hypothesis (data comes from a power-law) is rejected when the p.value is below a fixed value.

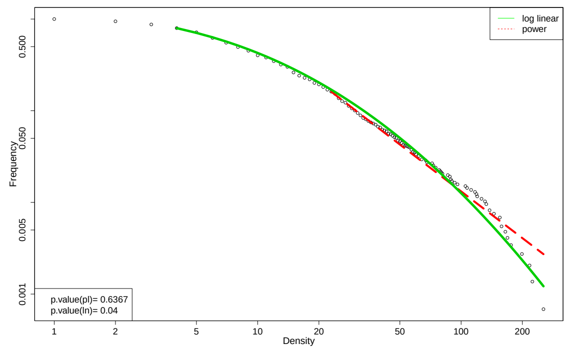


Figure 2: Cumulative Frequency Distribution of the IPC network at the 7-digit level

3.2.2 Small World identification

One of the reasons networks are important for innovation is the hypothesis that knowledge flows through them. The structure of the network plays an important role. In order for knowledge to flow quickly through a network firms need to be a low distance from each other in the network. The higher the average distance in the network, the longer is takes for knowledge to reach all nodes. The latter is a necessary condition, however it is not sufficient. The presence of small communities is also a condition. Within these communities knowledge flows even faster since firms are closely connected to a larger number of other firms. New knowledge developped in these communities spreads fast throughout the community and then to the whole network.

A network structure that has both the characteristics of low average distance and high clustering is a structure called the small world structure. It has been found to be an efficient structure for the diffusion of knowledge through a network (Cowan & Jonard, 2007; Verspagen & Duysters, 2004b). We hence want to know if our network has the particular structure of a small world. In order to check for small world features we need information on the average distance in the network as well as the clustering coefficient.

The clustering coefficient is a measure of cohesiveness in a network, in other words, how well connected the network is. The measure is quite simple; it represents the number of triangles in the network divided by the number of possible triangles.

Consider figure 3, to find the clustering coefficient we need the number of

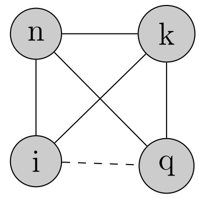


Figure 3: Clustering illustration

triangles in the network. There are two triangles in the network: i-n-k and n-q-k. The number of possible triangles is equal to the number of triangles if the network were a complete network. The dotted link between nodes i and q makes the network a complete network. If this link existed we would have two additional triangles: i-n-q and i-k-q. The number of possible triangles is hence equal to four. The clustering coefficient is then equal to:

$$Clustering = \frac{\sum_{i,j\neq i,k\neq j,k\neq i} g_{ij} \cdot g_{ik} \cdot g_{jk}}{\sum_{i,j\neq i,k\neq j,k\neq i} g_{ij} \cdot g_{ik}} = \frac{2}{4} = 0.5 \tag{1}$$

The same value can be computed at the node level. This would give a measure of the extend to which firms' neighbors are connected. It gives the fraction of the neighbors that are connected.

Whether measured at the level of the node or the network level, the clustering coefficient gives a measure of embeddedness. When clustering equals one all possibles triangles exist, the more it tends towards zero the less triangles are observed.

When a network is studied in a dynamic setting a problem arises with this measure. Instead of measuring simply a clustering coefficient we use a benchmark to which to compare the observed clustering coefficient (Watts & Strogatz, 1998; Baum *et al.* , 2003; Gulati *et al.* , 2012). The observed network is compared to a random network with the same number of links and nodes as the empirical network. Random networks typically have very low clustering since there is no reason why triangles would form a random. Clustering is then defined by the ratio :$\frac{C}{C_r}$. We use the same method for the average distance $\frac{L}{L_r}$.

Given that random graphs have low to no clustering and a low average distance we want a small world to show:

$$\frac{C}{Cr} >> 1 \qquad and \qquad \frac{L}{Lr} \approx 1 \tag{2}$$

3.3 Network dynamics

In order to track the evolution of the networks we use two different methods. For the collaboration network we start with all patents and publications that appeared in 1980 and extract collaborations. The latter is done by creating a link between firms that have co-deposited a patent or co-published a scientific article. Then, for the next year we add the collaborations that appeared that year. The network in 1985 hence contains all collaborations between 1980 and 1985. Co-patenting and co-publication collaborations are treated equally in the network.

In order to identify the life-cycle of the technology we use IPC codes present on the patents. Whenever two or more IPC codes are present on the same patent a link is created between them. The International Patent Classification (IPC) classifies patented technologies according to different technological fields. The system itself is crescendo is nature. The more digits, the more precise the codes become from a technological point of view. A 4-digit level defines a broad definition of technologies, for instance B64C defines "Aeroplanes; helicopters". The 7-digit level goes one step further into detail by specifying for B64C1, " Fuselages; Constructional features common to fuselages, wings, stabilizing surfaces, or the like". Going a step further, at the 9-digit level we will find "Floors" for code B65C1/18.

We use 3 different digit levels for our analysis, the 4, 7 and 9 digit levels. In order to obtain a 4-digit network, all IPC codes are reduced to their 4-digit format. For example B64C001/23 is a 9-digit code, in order to obtain the corresponding 4-digit code, one simply reduced the code to 4-digits: B64C. The same goes for the 7-digit code which would be B64C001 in this example.

Because of the hierarchy in the classifications, the 4-digit network will represent the interconnection of broad technological domains, while the 9-digit network pertains to more precise technological applications. The best fit for our analysis is hence the 9-digit network. However, the 4-digit network should have some interesting characteristics, it should show how different technological domains are interconnected.

4 Results

4.1 Structure of the knowledge network

4.1.1 The technology life-cycle

The 4-digit knowledge network represents the interconnection of broad technological fields. As shown in figure 4a, the number of codes increases each year. Even though the number of codes added decreases each year, the number of links increases steadily over the same period of time. This shows that the major areas for the application of the technology are identified early on. New recombinations are however found between the technological fields. The high level of clustering shown in figure 5a shows that there was a dense interconnection of the different fields from the start. The fields that are added over time do not reinforce the core but rather expand it. The average distance between the nodes in the network hence increases as shown in figure 5b. We notice here that the 4-digit network takes the particular structure of a small world. Figure 5a shows an adjusted clustering coefficient exceeding 1, even though declining over the period while figure 5b shows an adjusted average distance around the value of 1 around 1990. From that point on the network takes the structure of a small world and remains a small world during the rest of the period. This shows that the technological fields are locally clustered while all being at a close average distance from one another. Figure 7 shows the core of the network (the nodes with the highest number of links) The colors represent different communities in the complete 4-digit network. The network shows that even though these fields are densely interconnected they all can be divided into different communities.

More precisely, during the first decade (1980-1990) the IPC codes that had the largest number of deposits were:

C08L: COMPOSITIONS OF MACROMOLECULAR COMPOUNDS

C22C: ALLOYS

C08F: MACROMOLECULAR COMPOUNDS OBTAINED BY REACTIONS ONLY INVOLVING CARBON-TO-CARBON UNSATURATED BONDS

C07D: HETEROCYCLIC COMPOUNDS

C22F: CHANGING THE PHYSICAL STRUCTURE OF NON-FERROUS METALS OR NON-FERROUS ALLOYS

These codes relate to the fundamental development of the technology. In the succeeding decades, developments of the technology are added, baring witness to the start and evolution of the development phase. During these years, the IPC codes have changed to more diverse applications of the technology:

B82Y: SPECIFIC USES OR APPLICATIONS OF NANO-STRUCTURES

B82B: NANO-STRUCTURES FORMED BY MANIPULATION OF INDIVIDUAL ATOMS

D07B: ROPES OR CABLES IN GENERAL

B29C SHAPING OR JOINING OF PLASTICS

B64C AEROPLANES; HELICOPTERS

B64D EQUIPMENT FOR FITTING IN OR TO AIRCRAFT

F02C AIR INTAKES FOR JET-PROPULSION PLANTS; CONTROLLING FUEL SUPPLY IN AIR-BREATHING JET-PROPULSION PLANTS

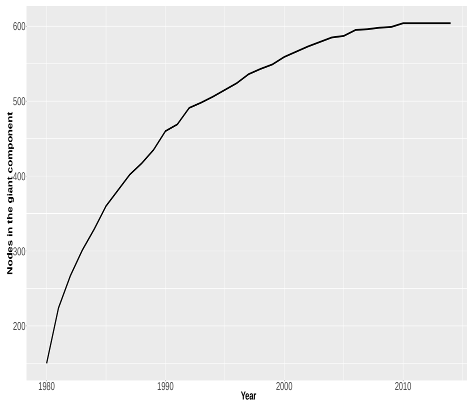
The latter clearly show a switch from the research phase to the development phase of the technology.

When we increase the number of digits, the precision of the technological fields increases. As a result, the core of the technology takes longer to stabilize. The 7-digit network stabilizes around the year 1990 while the 9-digit network stabilizes around the year 2000. For the identification of the stages of the life-cycle of the technology the 9-digit network is the most adequate. It represents the most detailed information about the domain of the technology and can hence be used for the identification of applications .

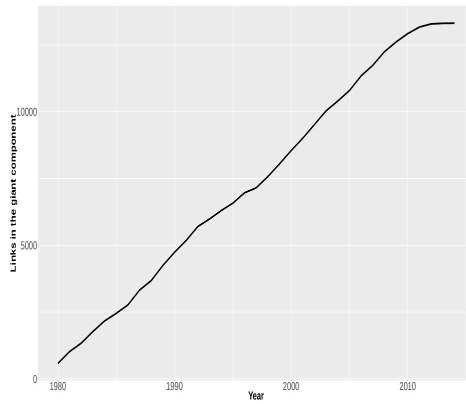
Figures 5c and 5e show an increase in the level of clusterin from the beginning of the period. This represents the first phase of the technology life-cycle: the research phase. The fundamental technologies are interconnected until creating the core of the technology. Since an additional link increases the overall clustering of the graph we can deduce that the core is being reinforced as long as clustering increases. This observation is reinforced by figures 5d and 5f that show a decrease in the average distance of the network during that first phase. Technologies are interconnecting densifying the core, reducing the distance separating them. From these observation we can confirm hypotheses 1 and 2.

The second stage of the technology life-cycle starts once the clustering coefficient of the network starts decreasing. New technologies are added to the network but they are not reinforcing the core, they stay in the periphery. The latter results in a decrease of the clustering coefficient and increases the average distance of the network. The codes that are added in the periphery of the network are applications of the technology. They do no connect to all the different IPC codes but rather to a specific fraction.

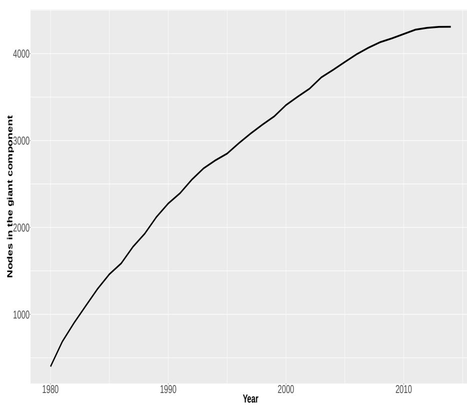
The increase of the average distance is difficult to observe in the 9-digit network which shows a stabilization of the average distance rather than an increase. In order to check the hypothesis that the structure is indeed a core-periphery structure we will use a statistical test on the degree distribution of the network.



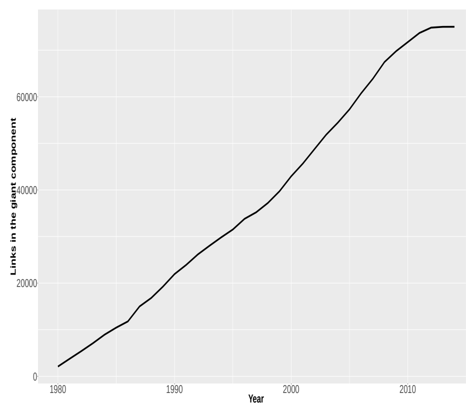
(a) Evolution of the number of IPC codes (4-digits)



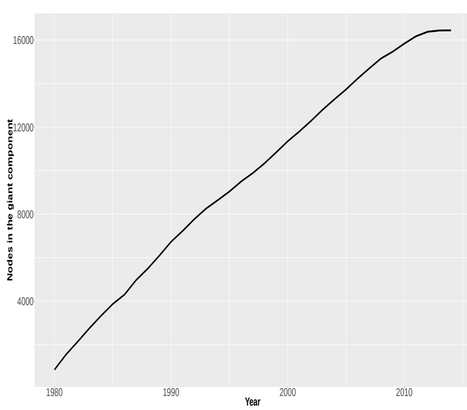
(b) Evolution of the number of links between IPC codes (4-digits)



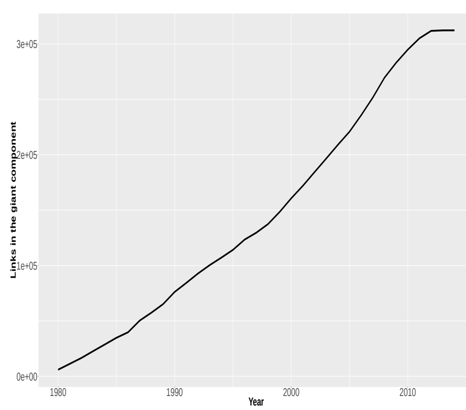
(c) Evolution of the number of nodes in the IPC network (7-digits)



(d) Evolution of the number of links in the IPC network (7-digits)

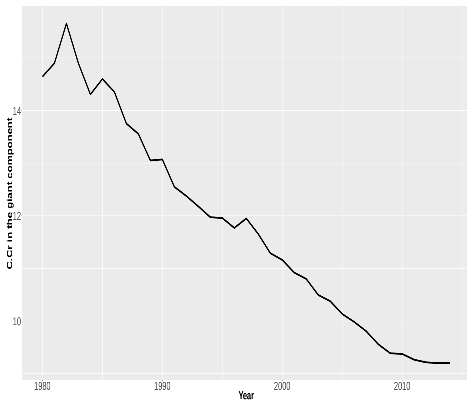


(e) Evolution of the number of nodes in the IPC network (9-digit)

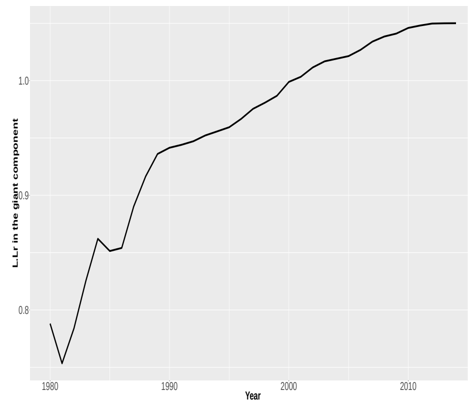


(f) Evolution of the number of links in the IPC network (9-digit)

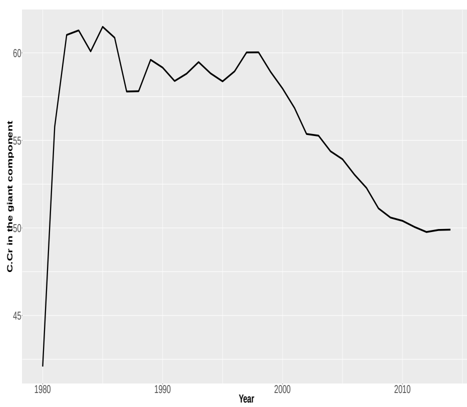
Figure 4: Dynamics of the IPC network for SCM



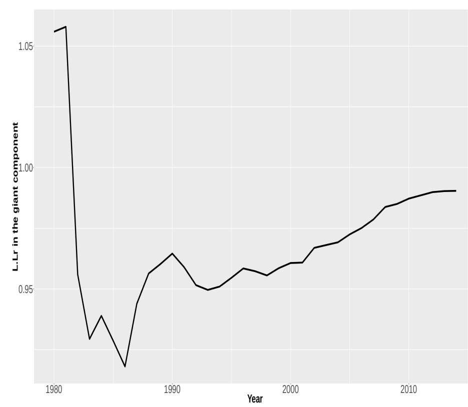
(a) Evolution of C/Cr in the 4-digit network



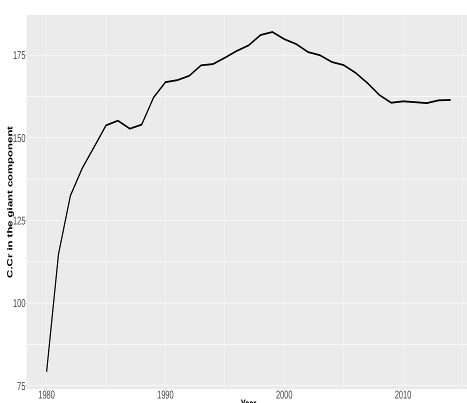
(b) Evolution of L/Lr in the 4-digit network



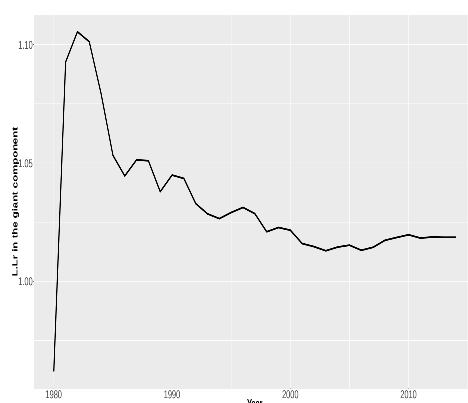
(c) Evolution of C/Cr in the 7-digit network



(d) Evolution of L/Lr in the 7-digit network



(e) Evolution of C/Cr in the 9-digit network



(f) Evolution of L/Lr in the 9-digit network

Figure 5: Dynamics of the IPC network for SCM

4.1.2 Core-periphery identification

Figure 6 contains the degree distribution for teh 7-digit and 9-digit networks. The green (plain) line is the fitted log-normal distribution, the red (dotted) line is the power law fit. The lower left corner of the graph contains the p.values. When the p.value is lower than 5% we reject the null hypothesis and conclude that the degree distribution has a core-periphery structure.

A first observation is that we reject the power-law fit for both the 7-digit and the 9-digit networks. The networks are not scale-free.

The log-normal fit is not significant in the first years of the 7-digit network, the fit becomes significant in 1995 (p.value = 0.27), and remains significant until the end of the period. Towards 2009, the parameters of the log-normal fit stabilize. Recall that the parameters of the log-normal distribution are the average and the standard-deviation. The variance tends towards a value of 1.96.

The 9-digit network appears to have a core-periphery structure quite early on, but the structure is not stable. The log-normal fit implies that the difference between the core and the periphery is less clean-cut as would be the case in a scale-free structure. The periphery is quite densely connected. The parameters of the network stabilize around the year 2000, the variance stabilizes around 3.24.

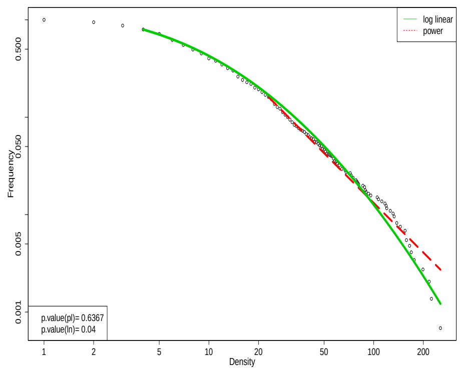
These observations allow us to conclude that the knowledge network for SCM technologies takes the structure of a core-periphery network and hence validate hypothesis 3.

We have now identified the different stages of the technology life-cycle as well as the stabilization of the network structures.

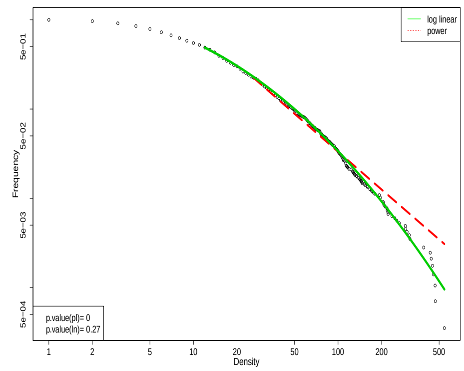
4.2 Structure of the collaboration network

Since we use publication data and patent data we start by identifying the structure of each type of network separately before turning to the complete network. Publications have a higher average number of co-assignees than patents do. Performing the analysis on the separate networks allows a better understanding of the structure of the network as a whole.

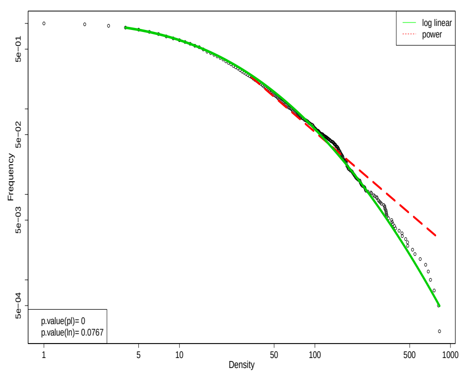
The plots in figure 8 show that the patent network is structured early on in the analysis. The number of nodes and links is computed on the the largest component of the network. The publication network is build up from a large



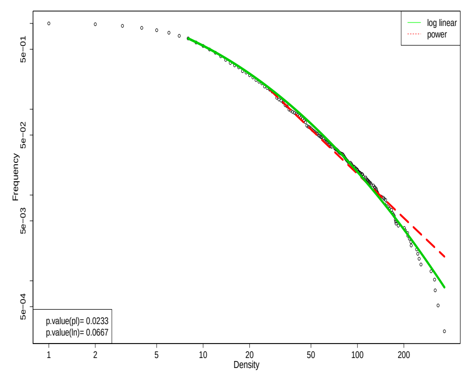
(a) Powerlaw fit for the 7-digit network for 1985



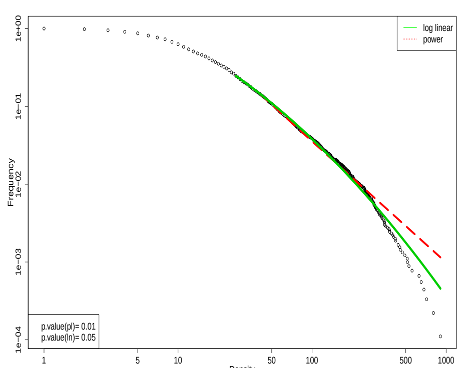
(b) Powerlaw fit for the 7-digit network for 1995



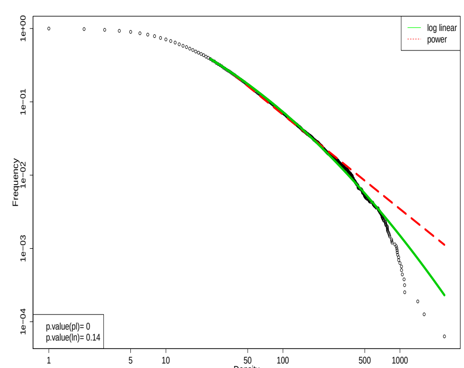
(c) Powerlaw fit for the 7-digit network for 2010



(d) Powerlaw fit for the 9-digit network for 1985



(e) Powerlaw fit for the 9-digit network for 1999



(f) Powerlaw fit for the 9-digit network for 2010

Figure 6: Dynamics of the IPC network for SCM

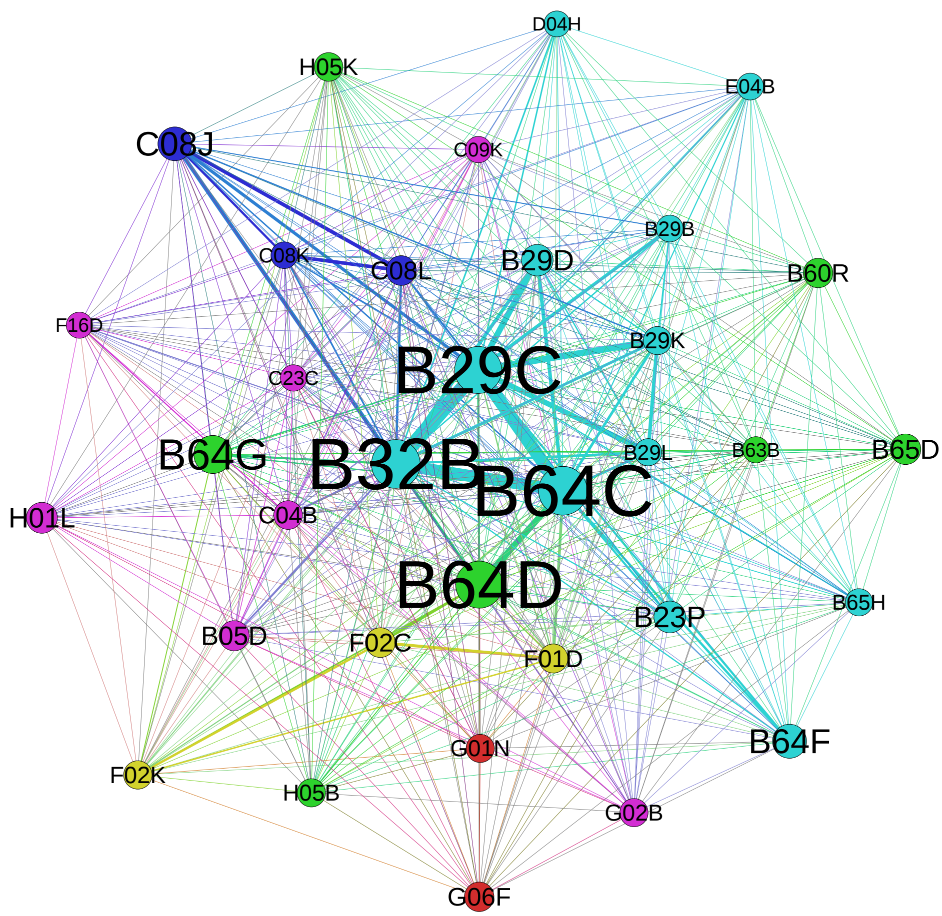


Figure 7: The core interconnections of SCM technologies

number of very small components that start to interconnect just before the year 2000.

Around the year 2008, the number of nodes in the publication network exceeds that of the patent network. The number of links is exceeded in 2005. In terms of links and nodes there is no clear cutoff point that indicates a switch from the research phase to the development phase. When we go into more detail we notice that there are many publications during the years 1980-1997 (see figure 1). The authors of these publications are actors of the space and defense sector that published mostly alone. The patents were deposited by the same type of actors (Boeing, US air force, Lockheed, Aerospace, NASA) but with collaboration. When we check the IPC codes on these patents we notice that they relate to the fundamentals of composite materials. In this case then, the research phase of the technology was accomplished by the private sector instead of the Universities.

The small number of collaborations in the publication network results in a low average distance as shown in figure 8c. The average distance in the publication network is higher showing a larger diversity in firms. The distance increases over time because of new actors entering the network. In addition smaller clusters start to interconnect. Around the year 2000 the average distance in the patent network stabilizes. Less firms deposit patents while the number of publications increases. Actors from academia are developing the technology, the number of universities entering the network increases at a steady pace over that period of time.

The stabilization and decline of the patent network starts in the year 2000 and is the result of existing actors in the network collaborating more extensively. New links are added between firms already in the network resulting in a decline in average distance and increase in clustering.

By connecting the patent network to the publication network we obtain the complete collaboration network. Firms that both deposit patents and publish in scientific journals will create connections between the two types of networks. The firms interconnecting the type network have a type of "gatekeeper" position. The gatekeepers are listed in table **??** accompanied by the year of first appearance as a gatekeeper. We notice here that it are mostly the large companies that interconnect both networks rather than large research institutions.

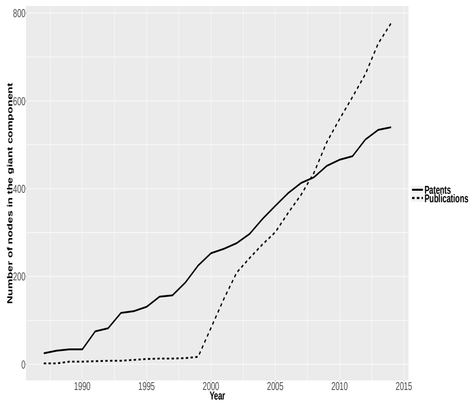
The evolution of the complete network can be found in figure 8e. For both networks we identify an inverted U-shape. We can hence distinguish two phases, a first phase staring in 1980 and ending around the year 2000, a second phase till the end of the period. During the first phase, identified as the research phase in the IPC network, the average distance in the network continuously increases. Mainly large multinational firms (e.g Honda, Taylormade golf, Rio Tinto, Saab, Astrium, Constellium, Daimler) enter the network during this phase. Since these actors have they own communities in the network the average distance increases during this period (the communities are interconnecting). When universities start to enter the network around the year 2000, the distance between the firms decreases. Universities typically take a central position and tend to have a large number of cooperations. In this sens they connect to many firms in different communities in the network reducing the distance. The appearance of research institutions also marks the start of the development phase of the technology. The fundament technology has been developed by the firms in the previous stage. Collaborations to find applications for the technology are launched in the second phase. The IPC network already showed that the patents deposited during this phase relate to developments of the fundamental technologies.

The data show then that the structure of the complete collaboration network converges towards a small worlds structure, this structure is reached around the year 2005. We can hence not validate hypothesis 4. The data does not show a convergence to / divergence from relation as we expected. However, hypothesis 5 appears to be validated. When the IPC network changes from the research stage to the development stage we appear to observe a change in the structure of the collaboration as well. In the case of SCM the entrance of universities in the network reduce the overall distance and cause the network to converge to a small world network. This means that once the technology is ready to be diffused, the structure of the network is optimal for knowledge diffusion.

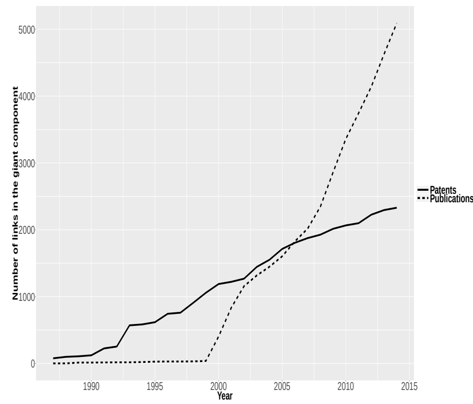
5 Conclusion and discussion

This paper shows that the evolution of the IPC network can identify different stages of the technology life-cycle. This method can be used to identify at which stage of the life-cycle a technology is positioned. Firms can use this information in their business strategy when it comes to the identification of potential partners for example.

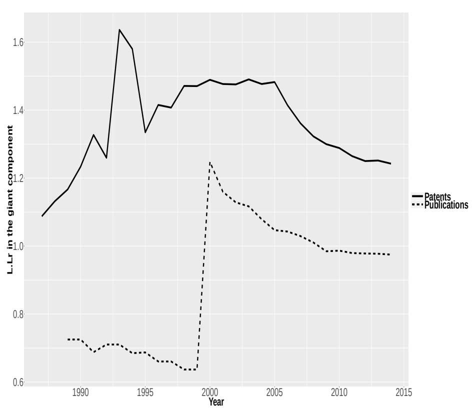
The IPC network has also shown that the initial technology was not developed by research institutions by the firms from the defense tier of the



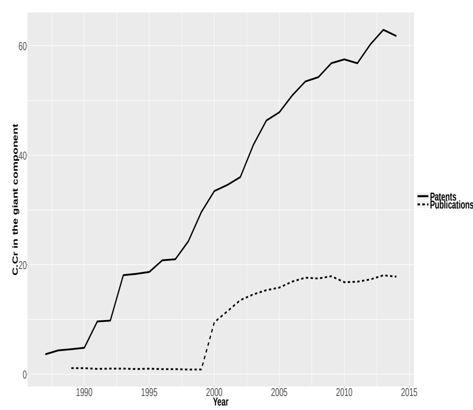
(a) Evolution of the number of nodes in each network



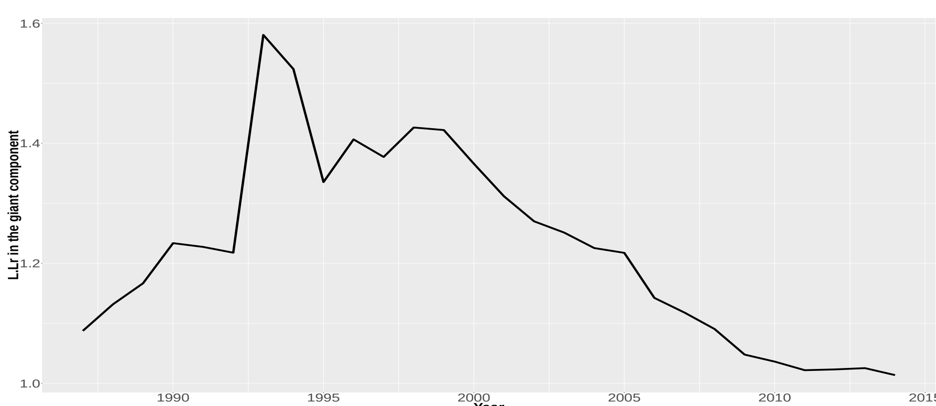
(b) Evolution of the number of links in each network



(c) Evolution of adjusted average distance in each network



(d) Evolution of adjusted clustering in each network



(e) Evolution of adjusted distance in the combined network

Figure 8: Dynamics of the collaboration network for SCM

aerospace sector. Of course, composite materials have existed before for other applications, but the results show that firms brought the technology to the aerospace sector.

This has led to patents in chemistry deposited by large firms in the aerospace sector. The development phase is characterized by the appearance of research institutions who played a vital role in the development of applications for the technology.

The evolution of the knowledge network presented the hypothesized characteristics, the first three hypotheses are hence validated. A core is formed during the research phase, while the development phase shows the appearance of a periphery around the core.

In the collaboration network the evolution of the structure was less clearcut than expected. In the early stages the average distance was too high to be a small world. The high average distance was due to a weak interconnection of the different clusters. These clusters started to regroup during the development phase reducing the average distance and resulting in the appearance of a small world structure, confirming hypothesis four.

When comparing the knowledge and collaboration networks we can identify the year 2000 as point from which structures change. The switch in the knowledge network states a change in the type of technology deposited (as proven by the IPC codes on the deposited patents during that phase)

while the change in the collaboration network show the entrance of new firms and universities (as shown by both the patent depositors and the publication authors). There hence is a clear correlation between the evolution of the collaboration network and the knowledge network in the case of SCM technologies.

References

Ahuja, Gautam. 2000. Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative science quarterly*, **45**(3), 425–455.

Alencar, MSM, Porter, AL, & Antunes, AMS. 2007. Nanopatenting patterns in relation to product life cycle. *Technological Forecasting and Social Change*, **74**(9), 1661–1680.

Barabási, Albert-László, & Albert, Réka. 1999. Emergence of scaling in random networks. *science*, **286**(5439), 509–512.

Baum, Joel AC, Shipilov, Andrew V, & Rowley, Tim J. 2003. Where do small worlds come from? *Industrial and Corporate change*, **12**(4), 697–725.

Buchmann, Tobias, & Pyka, Andreas. 2013. *The evolution of innovation networks: The case of a German automotive network*. Tech. rept. FZID discussion papers.

Cowan, Robin, & Jonard, Nicolas. 2007. Structural holes, innovation and the distribution of ideas. *Journal of Economic Interaction and Coordination*, **2**(2), 93–110.

Davide Chiaroni, Vittorio Chiesa & Federico Frattini. 2008. Patterns of collaboration along the biopharmaceutical innovation process. *Journal of Business Chemistry*, **5**(1), 7–22.

Frigant, Vincent, Kechidi, Med, & Talbot, Damien. 2006. Les territoires de l'aéronautique: EADS, entre mondialisation et ancrage.

Gao, Lidan, Porter, Alan L, Wang, Jing, Fang, Shu, Zhang, Xian, Ma, Tingting, Wang, Wenping, & Huang, Lu. 2013. Technology life cycle analysis method based on patent documents. *Technological Forecasting and Social Change*, **80**(3), 398–407.

Gulati, Ranjay, Sytch, Maxim, & Tatarynowicz, Adam. 2012. The rise and fall of small worlds: Exploring the dynamics of social structure. *Organization Science*, **23**(2), 449–471.

Kogut, Bruce, & Zander, Udo. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization science*, **3**(3), 383–397.

König, Michael, Tessone, Claudio J, & Zenou, Yves. 2009. A dynamic model of network formation with strategic interactions.

McEvily, Bill, & Marcus, Alfred. 2005. Embedded ties and the acquisition of competitive capabilities. *Strategic Management Journal*, **26**(11), 1033–1055.

McKelvey, Maureen, Alm, Håkan, & Riccaboni, Massimo. 2003. Does co-location matter for formal knowledge collaboration in the Swedish biotechnology–pharmaceutical sector? *Research Policy*, **32**(3), 483–501.

Nieto, María Jesús, & Santamaría, Lluis. 2007. The importance of diverse collaborative networks for the novelty of product innovation. *Technovation*, **27**(6), 367–377.

paolo Saviotti, Pier. 2007. On the dynamics of generation and utilisation of knowledge: The local character of knowledge. *Structural change and economic dynamics*, **18**(May), 387–408.

Pippel, Gunnar. 2013. The impact of R&D collaboration networks on the performance of firms: a meta-analysis of the evidence. *International Journal of Networking and Virtual Organisations*, **12**(4), 352–373.

Prencipe, Andrea. 1997. Technological competencies and product's evolutionary dynamics a case study from the aero-engine industry. *Research policy*, **25**(8), 1261–1276.

Rowley, Tim, Behrens, Dean, & Krackhardt, David. 2000. Redundant governance structures: An analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic management journal*, **21**(3), 369–386.

Salavisa, Isabel, Sousa, Cristina, & Fontes, Margarida. 2012. Topologies of innovation networks in knowledge-intensive sectors: Sectoral differences in the access to knowledge and complementary assets through formal and informal ties. *Technovation*, **32**(6), 380–399.

Stolwijk, CCM, Ortt, JR, & den Hartigh, E. 2013. The joint evolution of alliance networks and technology: A survey of the empirical literature. *Technological Forecasting and Social Change*, **80**(7), 1287–1305.

Tomasello, Mario Vincenzo, Napoletano, Mauro, Garas, Antonios, & Schweitzer, Frank. 2013. The rise and fall of R&D networks. *arXiv preprint arXiv:1304.3623*.

Trappey, Charles V, Wang, Tong-Mei, Hoang, Sean, & Trappey, Amy JC. 2013. Constructing a dental implant ontology for domain specific clustering and life span analysis. *Advanced Engineering Informatics*, **27**(3), 346–357.

Tsai, Wenpin. 2001. Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of management journal*, **44**(5), 996–1004.

Van Der Pol, Johannes. 2015. Structural dynamics of the French aerospace sector: A network analysis.

Van Der Pol, Johannes, Rameshkoumar, Jean-Paul, Virapin, David, & Zozime, Bernard. 2014. A preminary analysis of knowledge flows: The case of structural composite materials in aeronautics.

van der Valk, T., Chappin, M.H., & Gijsbers, G. W. 2011. Evaluating innovation networks in emerging technologies. *Technological Forecasting & Social Change*, **78**, 25–39.

Verspagen, B., & Duysters, G. 2004a. The small worlds of strategic technology alliances. *Technovation*, **24**, 563–571.

Verspagen, Bart, & Duysters, Geert. 2004b. The small worlds of strategic technology alliances. *Technovation*, **24**(7), 563–571.

Virapin, David, & Flamand, Marina. 2013. *L'innovation dans les matériaux composites, quelle diffusion autour de quels acteurs?* P2i conference 2013.

Watson, John. 2007. Modeling the relationship between networking and firm performance. *Journal of Business Venturing*, **22**(6), 852–874.

Watts, Duncan J. 1999. Networks, dynamics, and the small-world phenomenon 1. *American Journal of sociology*, **105**(2), 493–527.

Watts, Duncan J., & Strogatz, Steven H. 1998. Collective dynamics of 'small-world' networks. *Nature*, **393**(4), 440–442.

Appendix

	1996	2000	2004	2008	2014
AIRBUS	X	X	X	X	X
BOEING	X	X	X	X	X
GE	X	X	X	X	X
NASA	X	X	X	X	X
NORTHWESTERN UNIV	X	X	X	X	X
PECHINEY	X	X	X	X	X
SICMA	X	X	X	X	X
UNIV OF DELAWARE	X	X	X	X	X
UNIV OF VIRGINIA	X	X	X	X	X
ASTRIUM		X	X	X	X
DASSAULT		X	X	X	X
EUROCOPTER		X	X	X	X
HONEYWELL		X	X	X	X
ROLLS ROYCE		X	X	X	X
SAINT GOBAIN		X	X	X	X
BAE SYSTEMS			X	X	X
CNES			X	X	X
EADS			X	X	X
FRAUNHOFER			X	X	X
INDUSTRIELLE DU PONANT			X	X	X
MITSUBISHI ELECTRIC INFORMATION TECHNOLOGY			X	X	X
NIPPON STEEL CORP			X	X	X
ALSTOM				X	X
DAHER SOCATA				X	X
EUROP PROPULSION				X	X
MESSIER BUGATTI DOWTY				X	X
RENAULT				X	X
SNECMA				X	X
TECHSPACE				X	X
UNIV ORLEANS				X	X
CNRS					X
HISPANO HUREL					X
UNIV LORRAINE					X
UNIV REIMS					X

Table 1: Gatekeepers between the publication and patent collaboration networks