Exploring networks of proximity for partner selection, firms' collaboration and knowledge exchange. The case of clean-tech industry

Alessandro Marra (corresponding author)

alessandro.marra@unich.it

Associate Professor, University D'Annunzio of Chieti and Pescara, Italy and Explo academic spinoff, Pescara and Rome, Italy

Vittorio Carlei

vittorio.carlei@unich.it

Assistant Professor, University D'Annunzio of Chieti and Pescara, Italy and Explo academic spinoff, Pescara and Rome, Italy

Cristiano Baldassari

cristiano.baldassari@unich.it

Ph.D. candidate, University D'Annunzio of Chieti and Pescara, Italy

Keywords

Partner selection; firms' collaboration; knowledge exchange; M&A; clean-tech

Abstract

Nowadays scholars widely recognize that know-how, capabilities and knowledge needed to generate innovations often reside outside the firm, start-ups are a valuable source, and collaborative networks are a fundamental strategy for innovation. This is true especially for the clean-tech sector, which is characterized by the continuous search for innovative solutions and technological advancements. The purpose of the paper is to provide a methodological support for the screening of potential partners based on network analysis and, then, help firms to select them for collaboration and knowledge exchange. The methodology can be easily adopted by managers and executives to identify firms to monitor with greater attention for future investments. The analysis is on a dataset of 4,782 clean-tech companies operating worldwide. Results highlight that energy companies looking for external sources could investigate their network of business proximity if they intend to specialize in a defined field and/or collaborate with similar partners, while they could explore their network of strategic proximity if they intend to diversify their businesses, that is cooperating and exchanging knowledge with firms with distant but complementary capabilities and resources.

1. Introduction

Suppose there are two entrepreneurs in a Parisienne caffetterie. They are talking about their business. Although the real world is far more varied, we assume that we can summarize the activities carried out by each firm with a couple of keywords: both firms were active for several years in the field of ‘analytics’ applied to ‘energy’, where they offer consulting services and innovative projects. Exchange is very intense, the two men tell each other about their experience and agree on many technical issues and details, related for example to data management, visual interface reporting and predictive models. To make a short story, the capabilities owned by these two firms are very similar, even overlapping in several respects, and it is difficult to imagine that their exchange of ideas could result in new ‘disruptive’ ideas. Rather, the two entrepreneurs might get their returns discussing about how to approach new clients together or about the advantages from merging some corporate functions.

Suppose, instead, that in a coffee shop in Amsterdam, there are two other entrepreneurs with different professional background (below, in inverted commas the keywords describing their capabilities). The former, with a long experience in the energy sector having worked as a manager in major multinational companies, has now launched an ‘e-commerce’ site specialized in the marketing of technologies for ‘home energy efficiency’; the second is a physicist of extraction who has specialized at the university with a PhD in ‘semantics’ and has launched his start-up that employs ‘big data’ technologies and the semantic web to build high-quality ‘knowledge-graph’. Both entrepreneurs talk about their experience and the difficulties they encounter to develop business, and the more the conversation proceeds, the more they become interested about each other’s activities, trying to propose solutions to the difficulties and wondering how they can do something for each other’s success. To sum up, the two men exchange their know-how and generate new knowledge, which may refer to a semantic tool for identifying the next best product or a new index to measure the risk of losing the customers based on their preferences. The same situation might result also into nothing: with the pieces of knowledge brought by each, the meeting does not bring to any new idea or project.

Whether a firm chooses to collaborate with another one is largely determined by the degree to which it possesses the needed capabilities and the degree to which external partners have them. Nonetheless, the firm has to identify who has the ‘right’ capabilities and how much ‘complementary’ such capabilities are with its own. Is it possible to understand the chance that one entrepreneur can actually be useful to another one and that, then, they can identify profitable business synergies, initiate new partnerships or even think about a possible merger? We could think about a graph that links firms based on their know-how, capabilities and knowledge, using firms’ proximity in terms of products, services and technologies in order to screen and, eventually, select the most suitable partners.

The purpose of the paper is to provide a methodological support for the screening of potential partners based on network analysis and, then, help firms to select them for collaboration and knowledge exchange. The methodology can be easily adopted by managers and executives working to identify the firms to monitor with greater attention for future investments. We collect data from CrunchBase, one of the largest databases on new innovative companies in the high technology sector and employ the network analysis. The study is carried out on a dataset of 4,782 clean-tech companies operating worldwide in several industries.

Clean-tech is defined as any product, service or process that delivers value using fewer resources and producing less pollution than current standards require (Cooke, 2008). Any innovation that results in improved environmental performance falls under the clean-tech umbrella: recycling, renewable energy (wind-power, solar-power, biomass, hydropower, biofuels), information technology, green transportation, green buildings, electric motors, green chemistry, lighting, gray-water, and many other energy efficient appliances (Pernick and Wilder, 2007; Dangelico, 2016). As well known, clean-tech companies contribute significantly to a number of environmental problems and are characterized by the continuous search for innovative solutions and technological advancements (Wicki and Hansen, 2019). They need to deliver technological breakthroughs quickly enough and often decide to partner with new innovative start-ups working on disruptive products and services, instead of committing its internal departments in long and uncertain processes of Research, development and deployment (RD&D). As well known, committing to innovation is not an easy task. One way for large corporations to venture new technologies is corporate venture capital (CVC), which is an equity investment by an established corporation in an entrepreneurial initiative. This has received attention in the recent years, especially with regards to information technology and biotechnology. Instead, in clean-tech CVC has been substantially neglected despite rising interest and investments. CB Insights (2018) reports that corporate venture investments by energy companies hit records in 2016 and 2017. These investments have been concentrated in clean-tech companies and technologies. Companies such as Chevron and BP have driven CVC investment by the energy industry over the past decade in alternative energy (such as advanced batteries), alternative materials, and technologies for carbon capture, home energy efficiency, and vehicle efficiency. Companies focus also on mergers and acquisitions (M&A) to increase their core business or shift their scope into related markets. M&A are viewed as a vehicle to accelerate the response to the innovation race in clean-tech. According to Bureau van Dijk (2018) there were 2,813 M&A targeting companies in the clean-tech sector in 2017 with an aggregate value of USD 158,404 million, compared to 2,957 deals totalling USD 144,806 million in 2016. According to KPMG (2018), the capacity of corporates to fund M&A growth is expected to rise by 11 percent for the energy sector and 2 percent for utilities. In fact, many of the companies most active in acquiring clean-tech firms are big energy, utility, and industrial corporations, including General Electric, Cisco, Exelon, and NRG Energy. They make acquisitions to specialize and/or diversify their business, targeting companies that respectively add or complement their own know-how, capabilities and knowledge. Hence, for example, in 2016 Royal Dutch Shell (Shell, an oil and gas company headquartered in the Netherlands) acquired BG Group, a corporation involved in the exploration, development, and production of hydrocarbons, LNG shipping and sale, and operation of LNG import facilities, while in 2018 it acquired a stake in Asia-focused firm Cleantech Solar, which represented the energy giant's second investment in solar power.

The article is organized as follows. Section 2 introduces the literature on the methodologies for partner screening and selection. Section 3 describes CrunchBase and illustrates the dataset. Section 4 explains the methodology. Section 4.1 introduces the network of 'business proximity' between firms, which approximates the former situation at the Parisienne caffetterie. Section 4.2 illustrates a network of firms' capabilities. Section 4.3 presents the network of 'strategic proximity', which reckons the coffee shop in Amsterdam. Section 5 discusses the networks, focuses on the values of proximity, and validates the thesis by comparing them with the values of proximity of clean-tech firms that have actually carried out an M&A operation. Section 6 concludes.

2. Literature

Nowadays scholars widely recognize that know-how, capabilities and knowledge needed to generate innovations often reside outside the firm (Cohen and Levinthal, 1990; Teece et al., 1997), start-ups are a valuable source (Aghion and Tirole, 1994; Kortum and Lerner, 2000; Shane, 2001), and collaborative networks are a fundamental strategy for innovation (Ahuja and Lampert, 2001; Freeman, 1991; Hargadon and Sutton, 1997).

Knowledge exchange and firms' collaboration has attracted much research interest as firms are increasingly relying on partnerships to access to diverse and complementary assets that are not available internally (Mowery et al., 1996; Powell et al., 1996; Argote and Ingram, 2000; Inkpen and Pien, 2006).

We focus on the activities and features of the firms engaged in collaboration and knowledge exchange (Malmberg and Maskell, 2005; Allmendinger and Berger, 2019; Guertler and Lindemann, 2016), where technological proximity plays a crucial role (Knoben and Oerlemans, 2006). Proximity is a concept used in many different ways in the literature. We refer to the definition of technological proximity based on firms' products, services and technologies, where a high level of proximity between firms means a high degree of similarity in firms' know-how, capabilities and knowledge (Guan and Yan, 2016; Knoben and Oerlemans, 2006).

It is argued that companies have to be similar enough to facilitate collaboration and knowledge exchange (Baum et al., 2010; Zeller, 2004). Emden et al. (2006) state that firms' proximity (that is, overlapping assets) is an important common ground to discover and exploit complementarities. Whether a firm chooses to collaborate with another one is largely determined by the degree to which it possesses the needed capabilities and the degree to which external potential partners have the demanded assets. Collaborating on RD&D projects can offer a firm several advantages. First, collaborating can enable a firm to get necessary capabilities and resources more quickly. It is not unusual for a company to lack some of the complementary assets required to transform a body of technological knowledge into a commercial product (Schilling, 2017). A company may be able to gain rapid access to some assets by entering into collaborations (Hamel et al., 1989; Shan, 1990; Pisano, 1990; Venkatesan, 1992). Second, firms' collaboration help reducing firm's commitment, which is important in industries underlying rapid technological change, where product life cycles shorten and innovation generates huge risks (Schilling, 2017). Then, in such a context, firms might choose to become more specialized and activate partnerships with other firms to access know-how, capabilities and knowledge they do not possess internally. Third, collaboration allows to share knowledge: close contact with other firms facilitates the creation and transfer of information that firms could not create working alone (Mowery et al., 1998; Baum et al., 2000; Rosenkopf and Almeida, 2003). If the know-how, capabilities and knowledge of the companies are too similar, there is little to learn. Han et al. (2012) argue that collaborations between similar partners have fewer advantages because of the lack of diversities in terms of capabilities and resources. A low diversity between firms can have a negative impact on the ability to innovate (Nooteboom et al., 2007; Rosenkopf and Almeida, 2003; Sampson, 2007).

In all evidence, the phase of the partner screening and selection is a critical step (Cummings and Holmberg, 2012; Mindruta et al., 2016; Holzmann et al., 2014; Huizingh, 2011; Maurer and Valkenburg, 2014; Yoon and Song, 2014). A growing body of studies supports the use of quantitative techniques. Among the most popular quantitative techniques are fuzzy set logic, mixed integer linear programming, optimization modeling, analytic hierarchical and analytical network processes. Mikhailov (2002) focuses on the partnership selection by means of a fuzzy logic approach and an analytical hierarchy process. Wang and Chen (2007) discuss improvements to the fuzzy preference programming method for partnership screening using decision matrices instead of the exact values of the decision-maker's opinion. Accordingly, this method reduces the number of pairwise comparisons. Also, Herrera-Viedma et al. (2004) provide improvements relating to the fuzzy preference programming method.

The analytical hierarchy process is associated with drawbacks. Notably, it is difficult to assess the relationships between selected partners. Instead, the analytical network process addresses the disadvantage above because it considers interrelationships between partners in the decision-making model. Sarkis et al. (2007) discuss a partner screening tool by adopting the approach of the analytical network process. They elaborate a network of partners and, using inputs from managers, compare them. They calculate the relative ranking and calculate the weights relating to alternative partners. Chen et al. (2008) study an iterative review approach relating to partner selection by employing an analytical network process based on four criteria related to technology capability, resources for R&D, corporate compatibility, and financial conditions. Wu and Su (2005) investigate how the analytical network process can be used to screen partners, considering both tangible and intangible factors that can shape the partner selection decision. Also, Wu et al. (2009) evaluate tangible and intangible factors regarding potential partners using the analytical network process. The main issue in applying this approach stays in the effort to estimate interdependencies between factors, since often information regarding potential partners is limited.

Optimization models can help. Fuqing et al. (2006) apply a multi-objective optimization model, where the important decision is to set the values for each parameter. With the help of a generic algorithm the probabilities to succeed are calculated for each possible partner. The partner with the highest probability to succeed is selected. Cao and Wang (2007) improve the one-stage partner selection model in a two-stage model, where the first step relates to a pilot study screening multiple potential partners and the second step consists in the selection of the best suitable partner. Solesvik and Encheva (2010) propose a mathematical method of formal concept analysis to visualize data and its inherent structures, implications and dependencies, based on mathematical lattice theory. This method allows firms to screen and select the best partners for interfirm cooperation from several possible candidate firms. The formal concept analysis is a simple and versatile method, even though it has been marginally applied to aid managers.

Our methodology employs the network analysis to allow the screening of potential partners. Such a technique is helpful in fulfilling the step of the partner selection whenever firms' proximity has been calculated (Boschma and ter Wal, 2007; Cantner and Graf, 2006; Giuliani, 2007; McEvily and Marcus, 2005; Owen-Smith and Powell, 2004).

We propose two different networks: the former is based on 'business proximity' between firms based on the similarity in terms of products, services and technologies; the latter is a network of 'strategic proximity' where firms are linked based on the existing complementarities developed internally by firms in the observed population. The methodology leverages on qualitative databases, which include metadata and descriptive texts on firms' activities and markets, increasingly used also by managers and executives, and can be easily adopted to identify the most interesting companies to monitor for future investments.

3. Dataset

We collected data from CrunchBase, the world's most comprehensive database on high-tech industries operated by TechCrunch, one of the most highly regarded web blogs providing information on technological innovation. CrunchBase contains information on 500 thousand companies located in two hundred different countries, is partially crowd-sourced and is updated on a daily basis by investors and community contributors. As reported in Dalle et al (2017), the crowd-sourcing process, the partnership with investors, and the validation with sophisticated algorithms represent an added value compared to other commercial databases and public data sources.

CrunchBase is increasingly used in research (Waldner et al. 2012; Werth and Boert 2013; Homburg et al. 2014; Dalle et al. 2017; Marra et al. 2015; Cassetta et al. 2017; Ratzinger et al., 2018; Spiekermann, 2019; Moschner et al., 2019; Sussan et al., 2018). Dalle et al. (2017) compare the coverage of CrunchBase comparing it with the OECD Entrepreneurship Financing Database and find that the pattern across years and countries are substantially similar across the two data sources.

CrunchBase provides a powerful tagging system and lots of metadata that can be used to describe and classify businesses. CrunchBase does not require a huge amount of data handling before it can be used for econometric analysis (Dalle et al. 2017). Nonetheless, a check has been carried out on metadata, undertaking where requested a term normalisation process and revising typos.

Metadata are proxies of firms' products, services, technologies and, more generally, refer to the know-how, capabilities and knowledge on which firms build their own specializations. Such information, bottom-up (generated by companies' owners and employees, and other contributors) and up-to-date, is at a very detailed level and much more informative than SIC codes, as showed by the relevant and recent literature (Nathan and Rosso, 2015; Nathan et al., 2014; Papagiannidis et al., 2017; Tech City UK, 2017, 2015; Marra et al., 2015, 2017).

We built a dataset for all clean-tech start-ups in CrunchBase, including 4,782 companies operating worldwide. The 66% of the companies observed have been founded after the year 2000 (see Table 1).

Table 1

For each company CrunchBase provides information on the location of the company (country, state, region, city, latitude and longitude). The available data also includes information on the founding of the company, the industry category code, funding money, rounds and timing, number of employees, and metadata on firms' features and capabilities. Most companies in the dataset are from USA (see Table 2).

Table 2

Most of the companies observed are high-tech, according to the focus of CrunchBase on young companies operating in new emerging industries, such as software, clean-tech, nano-tech, consumer electronics, natural resources, marketing, financial services and mobile app sectors. All the companies in our dataset are mainly active in the energy industry, more generally in the field of sustainability, but also in other sectors and technological areas such as natural resources, science and engineering, hardware, manufacturing, and software (see Table 3).

Table 3

Despite its advantages, the dataset presents two potential limitations. The first issue is that the dataset focuses on start-ups active in high technology sectors and, therefore, there is a tendency to underestimate the presence of large and consolidated companies that have been operating on the market for a longer time, which are equally active in the clean-tech arena. Further research should be carried on in order to isolate and investigate the RD&D strategies of incumbent firms and large corporations. Another concern might rise from the fact that the observed clean-tech companies and, then, their keywords, are not weighted by means of parameters concerning their economic and financial size. As a consequence, the proximity of products, services and technologies can suggest either a promising market or technology or an already consolidated one. In any case, on these areas, a considerable number of new innovative companies are investing in RD&D.

4. Methodology

This Section explains the methodology. Section 4.1 introduces the network of 'business proximity' between clean-tech firms. Section 4.2 illustrates a network linking firms' tags on products, services and technologies (instead of companies). Section 4.3 presents the network of 'strategic proximity' between clean-tech firms. All networks are homogeneous and links between nodes are undirected and weighted.

4.1. A network of business proximity

We consider a network $N(V, E)$ with *V* and *E* being respectively nodes and edges. For every node $f_j \in V$ we have a set of unique *m* capabilities and resources, even called tags:

$$f_j = \{\tau_0 \dots \tau_m\}_j, \tag{1}$$

and we suppose that cardinality m could be different for each firm *j.*
We use a dyadic comparison between firms, and the set (1) is used to determine the relationship between the two firms f_i and f_j

$$w_{ij} = |f_i \cap f_j|, \tag{2}$$

where the business proximity w_{ij} is the cardinality of intersection set between firms f_i and f_j and we define w_{ij} as the weight of the edge $e_{ij} \in E$ for the network $N(V, E)$.
In other words, the nodes represent firms and the links between firms are generated whenever they share a tag, which are proxies of firms' capabilities in the same products, services and technologies. The higher the number of tags that the firms have in common, the closer they will be interconnected. The level of interconnection between firms (2) shows clusters of firms specialized in a set of products, services and technologies and, then, able to share know-how, capabilities and knowledge, collect external information, elaborate on it and exploit it for commercial purposes.

4.2. A network of firms' capabilities

We consider a network $\bar{N}(\bar{V}, \bar{E})$, where the tag τ_j are the nodes of the network. For every node $\tau_j \in \bar{V}$ we have a set *j* of firms *f* associated with it:

$$\tau_j = \{f_0 \dots f_m\}_j. \tag{3}$$

We use a dyadic comparison between capabilities, and the set is used to determine the relationship between the two tags τ_i and τ_j

$$\bar{w}_{ij} = |\tau_i \cap \tau_j|, \quad (4)$$

where the proximity $\bar{w}_{ij}$ is the cardinality of intersection between tags' sets and we define $\bar{w}_{ij}$ as the weight of the edge $\bar{e} \in \bar{E}$ for the network $\bar{N}(\bar{V}, \bar{E})$.

For example, capabilities 'energy efficiency' and 'manufacturing' are linked in the network if they coexist in the same firm, and the weight $\bar{w}_{ij}$ is heavier if the number of companies in which the two features coexist is larger. Therefore, for 'energy efficiency' and 'manufacturing', the weight of the edge 'energy efficiency'–'manufacturing' is ten as these coexist in ten different firms, and the weight of the edge 'energy efficiency'–'solar' is four as these coexist in four different firms, and so on. Heavier edges embody the complementarities between firms' assets and show the prevailing common channels by which firms can develop their business, exchange information and collaborate. We deepen the understanding of this aspect in the next Section. For the moment, it is sufficient to say that if some assets coexist internally in the same firm it is highly likelihood that such capabilities might represent the drivers by which two firms can grow externally, network and collaborate.

4.3. A network of strategic proximity

The third network is a network of strategic proximity, where firms are linked based on the existing complementarities developed internally by firms in the observed population. Intuitively, this last network is calculated by looking at whether and to what extent two firms can be considered close in the network of firms' capabilities. We provide a short exemplification. Even though the link between 'energy efficiency' and 'manufacturing' is at the firm level, the assumption is that the heavier the weight of the edge between the two capabilities (in terms of number of firms having both features) the higher the chance that a firm with capability 'energy efficiency' might find suitable to pursue growth strategies towards 'manufacturing', as well as to network or cluster with firms with capability 'manufacturing'. Contra, if the number of firms with capabilities in both 'solar' and 'data analytics' is low, then the chance that a firm with 'solar' might develop their business towards 'data analytics', or exchange information or collaborate with a firm with 'data analytics' is low.

The network of 'strategic proximity' is $\tilde{N}(\tilde{V},\tilde{E})$. Consider all the possible subsets made by the couples of capabilities, or tags, for a single firm f_j, already defined in (1) as:

$$f_j = \{\tau_0 \dots \tau_m\}_j, \tag{1}$$

Next, we define the tags symmetrical co-occurrence square matrix Φ as follows:

$$\Phi = \begin{pmatrix} 0 & \cdots & \varphi_{ks} & \cdots & \varphi_{kz} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \varphi_{st} & \cdots & 0 & \cdots & \varphi_{sz} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \varphi_{zt} & \cdots & \varphi_{zs} & \cdots & 0 \end{pmatrix} \tag{5}$$

with each element on the diagonal of the matrix Φ imposed to be null (meaning that same tags are intentionally excluded), and with cell element φ_{jk} representing the frequency indicating the number of co-occurrences of a couple of tags in the firms.

Then, we define the strategic proximity as a product between two firms

$$def\colon f_a \times f_b = \frac{\sum_{i,j}\varphi_{ij}}{n_a} + \frac{\sum_{i,j}\varphi_{ij}}{n_b} = \mu_a + \mu_b = \widetilde{w}_{ab} \tag{6}$$

with n_a and n_b being the cardinality of the sets associated respectively with f_a and f_b and μ_a, μ_b the average values of the φ_{jk}. This last formula represents the dyadic comparison between firms in the network $\tilde{N}(\tilde{V},\tilde{E})$ and so the value of the weight $\widetilde{w}_{ab}$ for the network of strategic proximity $\tilde{N}(\tilde{V},\tilde{E})$.

5. Results

The network of business proximity has 4,782 nodes (the firms in the dataset) and 10,876,222 edges (the links between firms). The average degree is 2,267 and the average weighted degree is 4,526. More specifically, the number of edges between nodes is high, on average each firm is linked with more than 2 thousand other firms, and these edges are strong since firms share many capabilities. As a consequence, the graph density is high and equal to 0.4. Modularity is also high at 0.1. Figure 1 shows a network of business proximity for a sample of 30 companies randomly extracted from the observed dataset. Within the network it is possible to appreciate two pairs of companies that have, de facto, proceeded to perform an operation of M&A. These are Evoqua Water Technologies and ADI Systems on the one side, and Entelios AG and EnerNOC on the other.

Figure 1

In 2017 Evoqua acquired ADI Systems. Evoqua, the global leader in helping municipalities and industrial customers protect and improve the water, has a long-lasting innovative experience and market-leading expertise in transforming water and wastewater. Its cost-effective and reliable treatment systems and services ensure uninterrupted quantity and quality of water, enable regulatory and environmental compliance and increase efficiency through water reuse. ADI Systems, founded in 1989, offers wastewater treatment and waste-to-energy solutions, using a wide range of technologies specifically designed for industrial wastewater treatment. The combination of ADI with Evoqua gives the latter the widest-ranging industrial wastewater offerings.

In 2014 EnerNOC acquired Entelios AG. The former is a leading provider of energy intelligence software and related solutions, the latter is Europe’s leading solutions provider for demand response and virtual power plants. This acquisition accelerated EnerNOC’s entry into continental Europe with Entelios’ existing relationships with leading grid operators, utilities, retailers, and commercial, institutional, and industrial customers. More specifically Germany, one of Europe’s largest potential markets for demand response and energy intelligence software, has a growing need for resources which can balance power supply and demand given the increasing level of intermittent renewable energy resources in its electricity system.

To notice that the value of the business proximity between Evoqua Water Technologies and ADI Systems on the one side, and Entelios AG and EnerNOC on the other, is equal to respectively 3 and 4, much higher than the average value of the pairs in the network (1.80).

The network of firms’ capabilities has 627 nodes (the total number of firms’ capabilities in the observed dataset) and 12,934 edges. The average degree for the entire network is 20 and the average weighted degree is 32. The graph density is 0.06 and modularity is 0.05.

The network of strategic proximity has 4,782 nodes and 11,431,366 edges (counting the number of all possible links between nodes). The average degree is high and equal to 2,389, and the average weighted degree is 52,292. The graph density is very high at 0.9 and modularity reaches 0.2 level. Again, the level of interconnection between firms is a proxy of their ability to collaborate and strategically share their know-how, capabilities and knowledge. Even though the interpretation of the third network is more speculative, the insights that might derive from its investigation provide considerable value.

Figure 2 shows network of a strategic proximity for a sample of 30 companies. Within the network it is possible to appreciate two pairs of companies that have, de facto, proceeded to perform an operation of M&A. These are GridPoint and Twenty First Century Utilities, on the one side, and Sunible and MyDomino on the other.

Figure 2

In 2015 Twenty First Century Utilities (TFCU), a company founded by power industry veterans, acquired GridPoint, a data-driven energy management solutions provider based in Arlington, VA. TFCU drives utilities towards mass adoption of clean, low cost energy producing and saving technologies, while optimizing the grid to enhance reliability, resiliency and security. A few years ago, GridPoint was the hottest start-up company in the cleantech stable. The GridPoint acquisition is the first action in TFCU’s campaign to transform regulated utilities and facilitate wide-scale adoption of emerging energy technologies, including distributed energy and other behind the meter applications.

In 2015, Sunible Inc. acquired MyDomino. Sunible is an online platform for solar, which is radically simplifying the home solar buying experience. More specifically, Sunible makes it easy for homeowners to see activity in their area, compare firms, get quotes, and go solar. Installers benefit from high-quality leads, lower customer acquisition costs, and market intelligence. MyDomino is a one-stop site that makes it easy for consumers to achieve clean, low-carbon living, and save money in the process. MyDomino guides individuals through actions they can take to enhance their lives, as well as help clean our air and water, and achieve energy independence. MyDomino and Sunible are joining forces to bring solar and other clean energy solutions to millions more people around the world while drastically reducing global carbon emissions.

The value of the strategic proximity between GridPoint and Twenty First Century Utilities on the one side, and Sunible and MyDomino on the other, is respectively 31,27 and 35,07, higher than the average value of the pairs in the network (21,80).

Firms usually look for external sources depending on their business proximity (i.e., collaborating with similar firms) and on their strategic proximity (i.e., firms exchange knowledge and cooperate when their capabilities are complementary but distant). The choice between the two different measures of proximity depends on the single firm's strategy: if the firm intends to focus on their core business and specialize in a specific field, then the firm will find useful the first measure of proximity, what we called business proximity. Instead, if the firm aims at growing externally through diversification, that means approaching new markets not necessarily correlated to the primary industry, then the firm will find more suitable the latter measure of proximity, strategic proximity.

It is interesting to look at the values of proximities for all the firms in the dataset and, more specifically, focus on those firms that effectively carried out an operation of M&A, which could have been promoted by specialization or diversification purposes. If we look at the values of the business proximity for all pairs of clean-tech firms in the dataset (10,876,222 observations) we get a frequency distribution similar to that in Figure 3 (attained from a random extraction of 50 thousand pairs of firms), with minimum value equal to 1, maximum value equal to 14, average that equals 1.8 and standard deviation of 0.9.

Figure 3

Instead, if we look at the values of the business proximity for the only pairs of clean-tech firms that carried out an operation of M&A (with 191 observed pairs of firms) we get the frequency distribution in Figure 4, with minimum value equal to 3, maximum value equal to 6, average that equals 3.7 and standard deviation of 0.7.

To notice that the average value of proximity for the M&A firms is +110% higher than the average value of proximity for the entire clean-tech network worldwide. This confirms that the methodology incorporates some elements that are effectively taken into account by firms in their process of screening and selection of partners for collaboration. Hence, it represents a useful methodology that helps firms looking for partners among a large and diversified set of external sources when they intend to specialize rather than diversify.

Figure 4

If we look at the values of the strategic proximity for all pairs of clean-tech firms in the dataset (11,431,366 observations) we get a frequency distribution similar to that in Figure 5 (attained from a random extraction of 50 thousand pairs of firms), with minimum value equal to 0.0, maximum value equal to 273.5, average that equals 21.8 and standard deviation of 10.0.

Figure 5

Instead, if we look at the values of the strategic proximity for all pairs of clean-tech firms carrying out an operation of M&A (140 observed pairs of firms) we get the frequency distribution in Figure 6, with minimum value equal to 24.6, maximum value equal to 61.4, average that equals 33.1 and standard deviation of 6.1.

To notice that the average value of proximity for the M&A firms is +51% higher than the average value of proximity for the entire clean-tech dataset. Given a lower gap between the average values, someone might conclude that this last methodology based on the strategic proximity is less effective than that based on the business proximity. However, we well understand that predicting firms' diversification strategies is much more difficult than forecasting possible trajectories of external growth based on specialization on the core businesses. Accordingly, we confirm that the methodology of strategic proximity is especially useful to advice firms looking for external partners among a large and diversified set of external sources when the goal of the firm is to vary and open up its technological scenarios and RD&D projects.

Figure 6

6. Conclusions

The purpose of the paper was to provide a methodological support for the screening of potential partners based on network analysis and, then, help firms to select them for collaboration and knowledge exchange.

An important aspect that needs to be considered in order to fully grasp the potential of the proposed methodology concerns the use of metadata on firms' products, services and technologies. We use metadata to calculate the proximity through which investigate the position of each firm with each other within networks based on qualitative data. Since metadata refer to the know-how, capabilities and knowledge on which firms build their own specializations, they contribute to provide information on companies' ability to exploit external sources, which is critical for example for RD&D. It follows that the completeness of the information collected on companies' activities, as well as the ability to systematize them into homogeneous categories, can result in an extremely informative database that can be used alternatively to more traditional data, such as those on firms' patents and research projects, to screen and select partners for collaboration and knowledge exchange.

We build networks linking firms based on their assets, using firms' proximity in terms of products, services and technologies in order to screen and select the most suitable partners. More specifically, we derived three different graphs. The first one, business proximity, represents a useful methodology that helps firms when they intend to specialize rather than diversify. The second graph, which was instrumental to calculate the third graph, is a network of firms' capabilities and shows complementarities between them. The last graph, which was more complex to calculate, is called strategic proximity. In our opinion it allows to realize how close are two companies with apparently different know-how, capabilities and knowledge. This methodology is especially useful to assist firms in looking for external partners when the goal of the firm is to diversify and open up its innovative scenarios. The methodology can be easily adopted by managers and executives.

Results highlight that companies looking for external sources could investigate their business proximity if they intend to specialize in a defined field and/or collaborate with similar partners, while they could explore their strategic proximity if they intend to diversify their businesses, that is cooperating and exchanging knowledge with companies with distant but complementary assets.

The present work shows interesting elements compared to the existing literature, mainly due to the novel use of network analysis for the screening and selection of ‘distant’ partners for inter-firm cooperation and knowledge exchange. More specifically, the originality of the third methodology lies in the fact that it allows the selection of partners, even on large networks, on the basis of existing complementarities developed internally by firms in the observed population, rather than on the similarity between firms (which, instead, characterizes the first network). In our view, this is a very important aspect. Through the network of strategic proximity, it is possible to identify pairs of companies otherwise not identifiable through the most common measures of proximity adopted in the literature, which are based on the similarity between companies. As a result, two companies with different capabilities, distant on the network, can still be studied in pairs by the researcher and/or analyst on the basis of the ease with which it is possible to find a considerable number of companies that individually present these two capabilities together.

To notice that the chance that two firms can identify areas of synergy, within which to start new collaborations and exchange knowledge, does not imply that any cooperation started should also be considered successful in terms of, for example, the level of innovation developed, or financial returns achieved. This last methodology aims to propose a tool to support the selection of partners with which it is more likely to find areas of synergy to be grasped, rather than a predictive tool aimed at estimating the degree of novelty or economic success of a possible partnership. In fact, on closer inspection, the higher the success of a partnership (as well as the exchange of knowledge), the lower the number of companies already working in those areas.

The proposed methodology is particularly useful in clean-tech, which are characterized by firms’ continuous search for innovative solutions and technological advancements. As well known, energy companies need to deliver technological breakthroughs quickly enough and might decide to partner with new innovative start-ups, instead of committing itself in long and uncertain processes of RD&D. Many organizations around the world are moving in the direction of supporting firms in the phase of the partner selection. The Clean Energy Group and the Meridian Institute, for example, are focusing on the creation of a collaborative platform to address the challenges of climate change (Clean Energy Group and the Meridian Institute, 2009).

References

Aghion, P, Tirole, J., 1994. The management of innovations. Quarterly Journal of Economics 109: 1185 – 1210.

Ahuja, G., Lampert, C.M., 2001. Entrepreneurship in the Large Corporation: A Longitudinal Study of How Established Firms Create Breakthrough Inventions. Strategic Management Journal 22 (2001), 521–43.

Allmendinger, M. P. and Berger, E. S. C. (2019). Selecting corporate firms for collaborative innovation: entrepreneurial decision making in asymmetric partnerships. International Journal of Innovation Management, 2050003.

Argote, L. and Ingram, P. (2000). 'Knowledge transfer: a basis for competitive advantage in firms'. Organizational Behavior and Human Decision Processes, 82, 150–69

Baum, J.A.C., Calabrese, T., Silverman, B.S., 2000. Don't Go It Alone: Alliance Network Composition and Startups' Performance in Canadian Biotechnology. Strategic Management Journal 21, p. 267.

Baum, J.A.C., Cowan, R., Jonard, N., 2010. Network-independent partner selection and the evolution of innovation networks. Manag. Sci. 56 (11), 2094–2110.

Boschma, R. A. and ter Wal, A. L. J. (2007). 'Knowledge networks and innovative performance in an industrial district: the case of a footwear district in the South of Italy'. Industry and Innovation, 14, 177– 99.

Bureau van Dijk, 2018. M&A Review - M&A Portal, available at http://www.mandaportal.com/getattachment/55b16d87-dc98-426f-bc41-32a76fece333/Cleantech

Cantner, U. and Graf, H. (2006). ‘The network of innovators in Jena: an application of social network analysis’. Research Policy, 35, 463–80

Cao, Q. and Wang, Q. (2007), “Optimizing vendor selection in a two-stage outsourcing process”, Computers & Operations Research, Vol. 34, pp. 3757-68.

Cassetta, E., Marra, A., Pozzi, C., Antonelli, P., 2017. Emerging technological trajectories and new mobility solutions. A large-scale investigation on transport-related innovative start-ups and implications for policy. Transportation Research Part A: Policy and Practice, 106, 1–11.

CB Insights, 2018. Oil And Gas Corporates Are Investing In Clean Tech, Analytics, And The Internet Of Things, available online at https://www.cbinsights.com/research/oil-gas-corporate-venture-capital-investment/

Chen, S.-H., Lee, H.-T. and Wu, Y.-F. (2008), “Applying ANP approach to partner selection for strategic alliance”, Management Decision, Vol. 46 No. 3, pp. 449-65.

Clean Energy Group and the Meridian Institute, 2009. Accelerated Climate Technology Innovation Initiative (ACT II): A New Distributed Strategy to Reform the U.S. Energy Innovation System, available at https://www.cleanegroup.org/ceg-resources/resource/accelerated-climate-technology-innovation-initiative-act-ii-a-new-distributed-strategy-to-reform-the-u-s-energy-innovation-system/

Cohen, W.M., Levinthal, D.A., 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly, March, 128–52.

Cooke, P., 2008. Cleantech and an analysis of the platform nature of life sciences: further reflections upon platform policies. Eur. Plan. Stud. 16, 375–393.

CrunchBase, 2018. Database behind API permission at https://www.crunchbase.com, extraction March 2018.

Cummings, J.L., Holmberg, S.R., 2012. Best-fit alliance partners: the use of critical success factors in a comprehensive partner selection process. Long Range Plan. 45 (2–3), 136–159.

Dalle, J., den Besten, M., Menon, C., 2017. Using Crunchbase for economic and managerial research, OECD Science, Technology and Industry Working Paper 2017/08, OECD Publishing, Paris. Retrieved from https://www.oecd-ilibrary.

org/industry-and-services/using-crunchbase-for-economic-and-managerial-research_6c418d60-en

Dangelico, R. M. (2016). Green product innovation: Where we are and where we are going. Business Strategy and the Environment, 25 (8), 560– 576.

Emden, Z., Calantone, R., Droge, C., 2006. Collaborating for new product development: selecting the partner with maximum potential to create value. J. Prod. Innov. Manag. 23 (4), 330–341.

Freeman C., 1991. Networks of Innovators: A Synthesis of Research Issues. Research Policy. 20, 499–514.

Fuqing, Z., Yi, H. and Dongmei, Y. (2006), “A multi-objective optimization model of the partner selection problem in a virtual enterprise and its solution with generic algorithms”, International Journal of Advanced Manufacturing Technology, Vol. 28, pp. 1246-53.

Giuliani, E. (2007). ‘The selective nature of knowledge networks in clusters: evidence from the wine industry’. Journal of Economic Geography, 7, 139–68.

Guan, J. C., and Yan, Y. (2016). Technological proximity and recombinative innovation in the alternative energy field. Research Policy, 45(7), 1460–1473. https://doi.org/https://doi.org/10.1016/j.respol.2016.05.002

Guertler, M. R. and Lindemann, U. D. O. (2016). Identifying open innovation partners: a methodology for strategic partner selection. international journal of innovation management, 20(05), 1640011.

Hamel, G., Doz, Y.L., Prahalad, C. K. 1989. Collaborate with Your Competitors—and Win. Harvard Business Review, January–February, 133–39.

Han, K., Oh, W., Im, K.S., Chang, R.M., Oh, H., Pinsonneault, A., 2012. Value cocreation and wealth spillover in open innovation alliances. MIS Q. 36 (1), 291–315.

Hargadon, A.B., Sutton, R.I., 1997, Technology Brokering and Innovation in a Product Development Firm, Administrative Science Quarterly 42, 716–49.

Herrera-Viedma, E., Herrera, F., Chiclana, F. and Luque, M. (2004), “Some issues on consistency of fuzzy preference relations”, European Journal of Operational Research, Vol. 154, pp. 98-109.

Holzmann, T., Sailer, K., Katzy, B.R., 2014. Matchmaking as multi-sided market for open innovation. Tech. Anal. Strat. Manag. 26 (6), 601–615.

Homburg, C., Hahn, A., Bornemann, T., Sandner, P., 2014. The role of chief marketing officers for venture capital funding: Endowing new ventures with marketing legitimacy. Journal of Marketing Research, 51(5), 625–644.

Huizingh, E.K., 2011. Open innovation: state of the art and future perspectives. Technovation 31 (1), 2–9.

Inkpen, A. C. and Pien, W. (2006). ‘An examination of collaboration and knowledge transfer: ChinaSingapore Suzhou Industrial Park’. Journal of Management Studies, 43, 779–811.

Knoben and Oerlemans, 2006. Proximity and inter-organizational collaboration: A literature review J. International Journal of Management Reviews.

Kortum, S, Lerner, J., 2000. Assessing the contribution of venture capital to innovation? Rand Journal of Economics 31(4): 674–692.

KPMG, 2018. M&A Predictor 2018: Global Energy Sector, available online at https://home.kpmg/xx/en/home/insights/2018/04/m-and-a-predictor-2018-gobal-energy-sector.html

Malmberg, A. and Maskell, P. (2005). Localized Learning Revisited. DRUID Working Paper No. 05-19, Copenhagen Business School, Copenhagen.

Marra, A., Antonelli, P., Dell’anna, L., Pozzi, C., 2015. A network analysis using metadata to investigate innovation in clean-tech: Implications for energy policy. Energy Policy, 86, 17–26.

Marra, A., Antonelli, P., Pozzi, C., 2017. Emerging green-tech specializations and clusters: A network analysis on technological innovation at the metropolitan level. Renewable and Sustainable Energy Reviews, 67, 1037–1046.

Maurer, C., Valkenburg, R., 2014. Approaches to networked innovation. Int. J. Innov. Technol. Manag. 11 (01), 1–18.

McEvily, B. and Marcus, A. (2005). ‘Embedded ties and the acquisition of competitive capabilities’. Strategic Management Journal, 26, 1033–55.

Mikhailov, L. (2002), “Fuzzy analytical approach to partnership selection in formation of virtual enterprises”, Omega, Vol. 30 No. 5, pp. 393-401.

Mindruta, D., Moeen, M., Agarwal, R., 2016. A two-sided matching approach for partner selection and assessing complementarities in partners' attributes in inter-firm alliances. Strateg. Manag. J. 37 (1), 206–231.

Moschner, S.-L., Fink, A. A., Kurpjuweit, S., Wagner, S. M., & Herstatt, C. (2019). Toward a better understanding of corporate accelerator models. Business Horizons, 62(5), 637–647. https://doi.org/https://doi.org/10.1016/j.bushor.2019.05.006.

Mowery, D. C., Oxley, J. E. and Silverman, B. S. (1996). ‘Strategic alliances and interfirm knowledge transfer’. Strategic Management Journal, 17, 77–91

Mowery, D.C., Oxley J.E., Silverman B.S. (1998).Technological Overlap and Interfirm Cooperation: Implications for the Resource-Based View of the Firm, Research Policy 27, 507–24.

Nathan M., Rosso A. (2015), Mapping digital businesses with big data: some early findings from the UK. Research Policy 44(9): pp. 1714-1733.

Nathan M., Rosso A., Bouet F. (2014), Mapping Information Economy Business with Big Data: Findings from the UK. Nesta Working Paper 14/10, London: Nesta.

Nooteboom, B., van Haverbeke, W., Duysters, G., Gilsing, V., van den Oord, Ad, 2007. Optimal cognitive distance and absorptive capacity. Res. Policy 36 (7), 1016–1034.

Owen-Smith, J. and Powell, W. W. (2004). ‘Knowledge networks as channels and conduits: the effects of spillovers in the Boston biotechnology community’. Organization Science, 15, 5–21.

Papagiannidis S., See-To E. W.K, Assimakopoulos D. G., Yang Y. (2017), Identifying industrial clusters with a novel big-data methodology: Are SIC codes (not) fit for purpose in the Internet age? Computers and Operations Research, Elsevier.

Pernick, R., Wilder, C., 2007. The Clean Tech Revolution: The Next Big Growth and Investment Opportunity. HarperCollins, New York.

Pisano, G.P., 1990. The R&D Boundaries of the Firm: An Empirical Analysis. Administrative Science Quarterly 35, 153–76.

Powell, W. W., Koput, K. W. and Smith-Doerr, L. (1996). 'Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology'. Administrative Science Quarterly, 41, 116–46.

Ratzinger, D., Amess, K., Greenman, A., & Mosey, S. (2018). The impact of digital start-up founders' higher education on reaching equity investment milestones. The Journal of Technology Transfer, 43(3), 760–778. https://doi.org/10.1007/s10961-017-9627-3.

Rosenkopf, L., Almeida, P., 2003. Overcoming local search through alliances and mobility. Manag. Sci. 49 (6), 751–766.

Sampson, R.C., 2007. R&D alliances and firm performance: the impact of technological diversity and alliance organization on innovation. Acad. Manag. J. 50 (2), 364–386.

Sarkis, J., Talluri, S. and Gunasekaran, A. (2007), "A strategic model for agile virtual enterprise partner selection", International Journal of Operations & Production Management, Vol. 27 No. 11, pp. 1213-34.

Schilling, M., 2017. Strategic management of technological innovation. McGraw-Hill, Boston.

Shan, W., 1990. An Empirical Analysis of Organizational Strategies by Entrepreneurial High-Technology. Strategic Management Journal 11, 129–39.

Shane, S., 2001. Technological opportunities and new firm creation. Management Science 47(2): 205–220.

Solesvik, M. Z. and Encheva, S. (2010) , "Partner selection for interfirm collaboration in ship design", Industrial Management & Data Systems Vol. 110 No. 5, 2010 pp. 701-717.

Spiekermann, M. (2019). Data Marketplaces: Trends and Monetisation of Data Goods. Intereconomics, 54(4), 208–216. https://doi.org/10.1007/s10272-019-0826-z.

Sussan, F., Sloboda, B., & Hall, R. (2018). Is There a Path from Sin City to Tech City? The Case for Las Vegas. In A. O'Connor, E. Stam, F. Sussan, & D. B. Audretsch (Eds.), Entrepreneurial Ecosystems: Place-Based Transformations and Transitions (pp. 153–171). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-63531-6_8.

Tech City UK (2015), Tech Nation. Powering the digital economy.

Tech City UK (2017), Tech Nation. At the forefront of global digital innovation.

Teece, G., Pisano, G., Shuen, R., 1997. Dynamic capabilities and strategic management. Strategic Management Journal 18(7): 509–533.

Venkatesan, R., 1992. Strategic Sourcing: To Make or Not to Make. Harvard Business Review 70, no. 6, 98–107.

Waldner, F., Zsifkovits, M., Heidenberger, K., 2012. Emerging service- based business models in the music industry: An exploratory survey. In M. Snene (Ed.), Exploring Services Science (321–329). Berlin: Springer.

Wang, T.-C. and Chen, Y.-H. (2007), “Applying consistent fuzzy preference relations to partnership selection”, Omega, Vol. 35, pp. 384-8.

Werth, J.C., Boert, P., 2013. Co-investment networks of business angels and the performance of their start-up investments. International Journal of Entrepreneurial Venturing, 5, 240–256.

Wicki, S, Hansen, EG. Green technology innovation: Anatomy of exploration processes from a learning perspective. Bus Strat Env. 2019; 28: 970– 988.

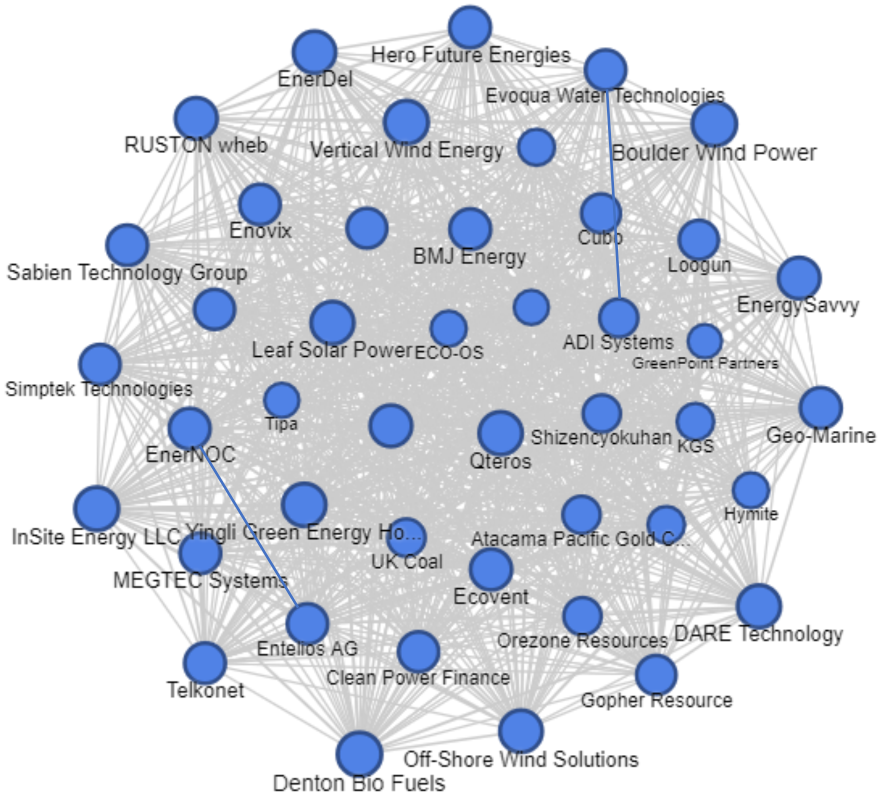
Wu, N. and Su, P. (2005), “Selection partners in virtual enterprise paradigm”, Robotics and Computer-Integrated Manufacturing, Vol. 21, pp. 119-31.

Wu, W.Y., Shih, H.-A. and Chan, H.-C. (2009), “The analytic network process for partner selection criteria in strategic alliances”, Expert Systems with Applications, Vol. 36, pp. 4646-53.

Yoon, B., Song, B., 2014. A systematic approach of partner selection for open innovation. Ind. Manag. Data Syst. 114 (7), 1068–1093

Zeller, C. (2004). North Atlantic innovative relations of Swiss pharmaceuticals and the proximities with regional biotech arenas. Economic Geography, 80(1), 83 –111.

Figure 1: The 'business proximity' network (sample of 30 firms)

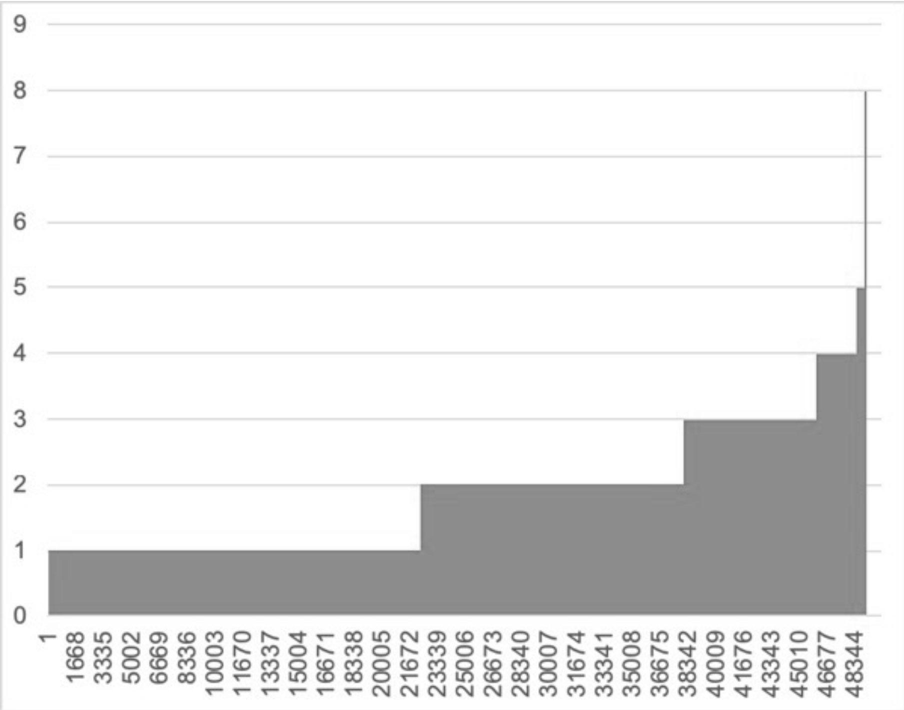


Source: Own elaboration on CrunchBase (2018)

Figure 2: The ‘strategic proximity’ network (sample of 30 firms)

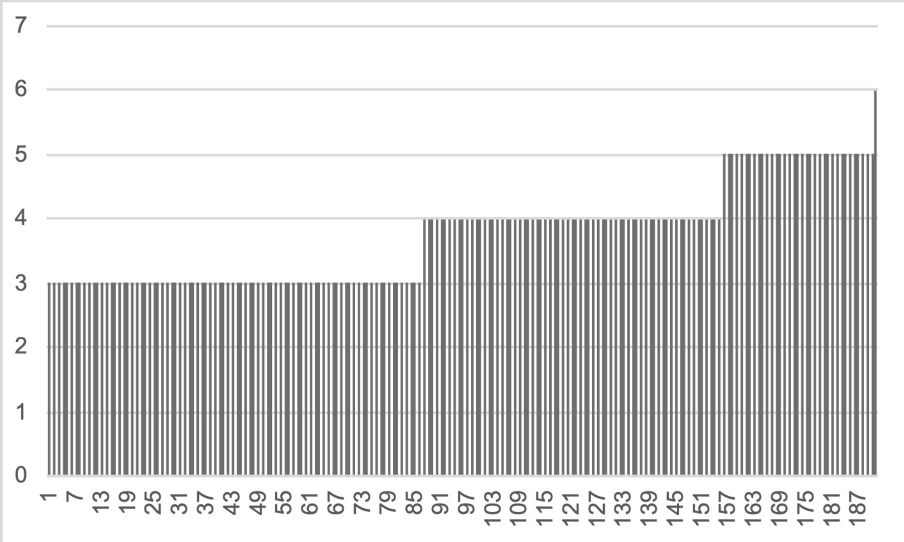
Source: Own elaboration on CrunchBase (2018)

Figure 3: Frequency distribution for all firms in the network of business proximity



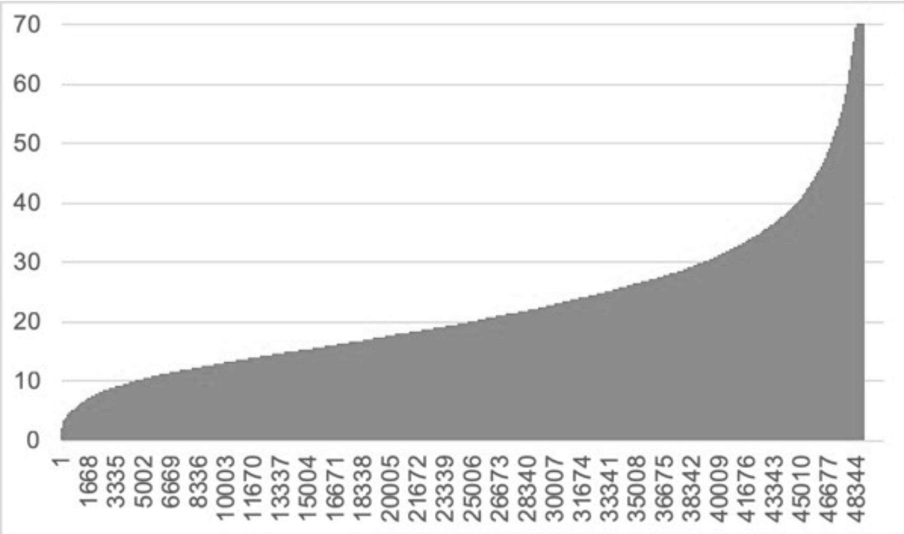
Source: Own elaboration on CrunchBase (2018)

Figure 4: Frequency distribution for M&A firms in the network of business proximity



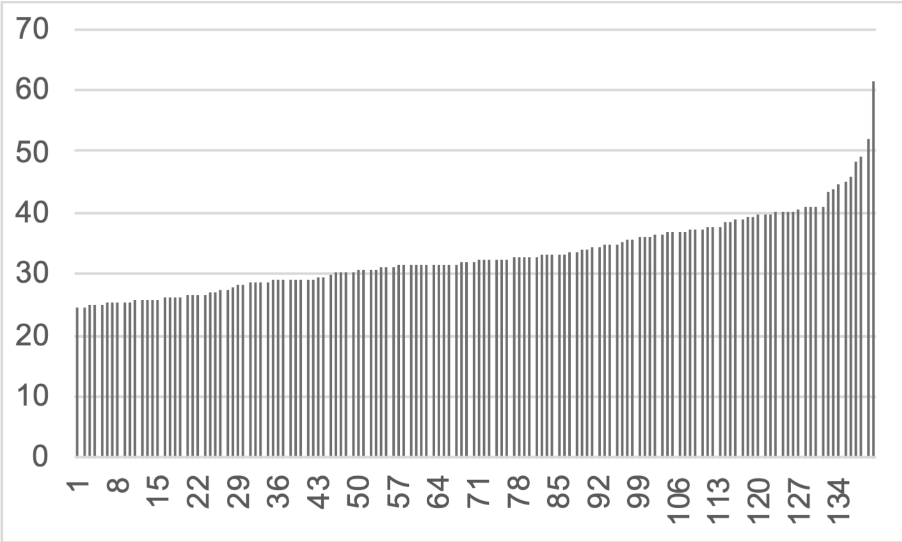
Source: Own elaboration on CrunchBase (2018)

Figure 5: Frequency distribution for all firms in the network of strategic proximity



Source: Own elaboration on CrunchBase (2018)

Figure 6: Frequency distribution for M&A firms in the network of strategic proximity



Source: Own elaboration on CrunchBase

Table 1: Number of clean-tech firms per year

Year	Number of firms	Year	Number of firms
Before 2000	1620	2010	266
2001	95	2011	257
2002	98	2012	255
2003	90	2013	277
2004	119	2014	247
2005	152	2015	200
2006	188	2016	90
2007	252	2017	43
2008	260	2018	1
2009	272		

Source: Own elaboration on CrunchBase (2018)

Table 2: Number of clean-tech firms per country

Country	Number of firms	Country	Number of firms
USA	2462	SWE	61
GBR	403	ITA	51
CAN	372	DNK	39
FRA	123	IRL	39

DEU	118	FIN	34
IND	114	ZAF	33
AUS	103	CHE	28
CHN	97	NOR	24
NLD	78	BEL	22
ESP	69	SGP	21
ISR	67	JPN	20
		Other countries	404

Source: Own elaboration on CrunchBase (2018)

Table 3: Number of clean-tech firms per sectoral and technological areas (category group)

Industry_category Number _group	In %	Industry_category Number _group	In %	Industry_category Number _group	In %
sustainability	26.5%	biotechnology	0.9%	navigation and mapping	0.2%
energy	21.4%	consumer goods	0.8%	sports	0.2%
natural resources	12.8%	food and beverage	0.8%	administrative services	0.2%
science and engineering	4.7%	design	0.8%	content and publishing	0.2%
hardware	4.2%	health care	0.7%	government and military	0.2%
manufacturing	3.7%	mobile	0.7%	privacy and security	0.2%
software	3.1%	professional services	0.6%	video	0.1%
consumer electronics	3.0%	artificial intelligence	0.5%	gaming	0.1%
data and analytics	1.7%	media and entertainment	0.4%	advertising	0.1%
real estate	1.6%	community and lifestyle	0.3%	platforms	0.1%
commerce and shopping	1.5%	lending and investments	0.3%	payments	0.0%
internet services	1.4%	sales and marketing	0.3%	messaging and telecommunications	0.0%
transportation	1.4%	apps	0.3%	events	0.0%

information technology	1.4%	education	0.2%	music and audio	0.0%
agriculture and farming	1.0%	clothing and apparel	0.2%		
financial services	0.9%	travel and tourism	0.2%		

Source: Own elaboration on CrunchBase (2018)