Technological Forecasting & Social Change 156 (2020) 120038

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

A network analytic method for measuring patent thickets: A case of FCEV technology

Xiaodong Yuan[a,*], Xiaotao Li[b]

[a] School of Management, Huazhong University of Science and Technology (HUST), P.R. China
[b] School of Literature, Law and Economics, Wuhan University of Science and Technology (WUST), P.R. China

ARTICLE INFO

Keywords:
Patent thickets
Network analytic method
Triad census
Data-driven social role analysis

ABSTRACT

Patent thickets may hinder technology innovation by preventing manufacturers from access to given technology fields. How to prove the existence of patent thickets or how to hack through patent thickets in complex technology areas has attracted a great concern in academia since Shapiro (2000) put forward the theory of patent thickets. This needs a valid and simple method to measure patent thickets. The existing methods are hard to further explore who own complementary patents in patent thickets. The paper proposes a novel method of combining triad census and data-driven social role analysis to measure patent thickets. Taking fuel cell electric vehicle (FCEV) patents at the USPTO and EPO as examples, the paper demonstrates the proposed method is valid and feasible in practice. Both the density of patent thickets and key patent holders in patent thickets can be accurately detected. It can help downstream manufacturers to make a decision whether they should enter given technologies or who could be potential licensors in patent thickets. Researchers or company managers can use this method to measure patent thickets over time from a microscopic perspective.

1. Introduction

The phenomenon of patent thickets has drawn wide attention from scholars and policymakers during the past two decades. From the perspective of economic theory, the generic issue inherent in patent thickets is the complements problem that originally proposed by Cournot in 1838 (Shapiro, 2000). If a patent is independent but should work with the other patent, these patents will form complementary relation. Each complementary patent should be got a license before it is used to products. In given technologies, the complements problem will arise if there are a large number of complementary patents but owned by numerous patent holders. Worse yet, some patentees may assert royalties significantly above their patent economic contribution in this context, which is known as the patent hold-up problem (Lemley and Shapiro, 2007). Moreover, the stacking effect will increase royalties in patent fragmented markets, which is called the royalty stacking problem (Lemley and Shapiro, 2007). The simultaneous occurrence of these problems leads to that downstream manufactories fail to get licenses to make their products. As a result, many complementary patents will be underused. Heller and Eisenberg (1998) discovered the blocking effect of patent thickets on biomedical technology, and then proposed the famous hypothesis patent "tragedy of the anticommons".

So far, the most popular definition of the patent thicket is a large overlap of patent rights requiring that firms seeking to commercialize new technology obtain licenses from multiple patent holders (Shapiro, 2000). From the perspective of cumulative innovation, numerous incremental inventions are developed mainly on the basis of previous patents or "prior art". The rush for patent applications will lead to overcrowded patent thickets, which may deter cumulative innovation and continued product development (Tullis, 2012). The increasing transaction cost of navigating patent thickets results in that a large number of patents are underused but numerous manufacturers are entirely prevented from accessing given technologies (Hall, 2013). Therefore, patent thickets will hinder technology innovation where new patented products depend on thousands of patents owned by numerous patent holders. If downstream manufacturers fail to trek through patent thickets, a patent tragedy of the anti-commons could arise (Zimmeren and Verbeure, 2006).

To address the issues, it is very important to measure patent thickets in given technology fields (Bondibene, 2012). At least three methods have been proposed in the existing literature. First, Ziedonis (2004) proposed the fragmentation index built on the Herfindahl concentration index to calculate the fragmentation of patent rights at the firm level. The function of the fragmentation index is to measure whether patents are concentrated in the hands of a few companies or dispersed among numerous firms. Second, Graevenitz et al. (2011) proposed the triples

* Corresponding author.
E-mail address: yuanxd@hust.edu.cn (X. Yuan).

https://doi.org/10.1016/j.techfore.2020.120038
Received 13 February 2019; Received in revised form 7 February 2020; Accepted 24 March 2020

Table 1
Key scholarly works on the definitions of a patent thicket.

Types	Code	Definitions	Author
Strategic option	1	An innovator possesses a great number of patents in given technological fields.	Teece (1986)
		An organization files a lot of adjacent patents.	Somaya (2003)
	2	Companies endeavor to file numerous patent applications with lower-quality.	Hegde et al. (2009)
The fragmentation of patent rights	3	A series of complementary patents are owned by multiple patent holders.	(Heller and Eisenberg, 1998)
	4	A dense web of overlapping patent rights	Shapiro (2000)
		Patent horizontal overlap and patent vertical overlap	(Burk and Lemley, 2003)

indicator to measure patent thicket across technology fields. A triple is composed of three firms, in which each firm has forward citation limiting claims on the patents owned by the other two firms (Fischer and Ringler, 2015). Forward citations refer to the citations that a prior patent receives from subsequent patents. Third, Clarkson (2005) proposed a derived equation to identify patent thickets by calculating the outdegree or indegree of citation networks. The common feature of the methods is that the measurement of patent thickets builds on forward citations. Each prior patent in the forward citations is regarded as the complementation of subsequent patents (Entezarkheir, 2017).

However, there are some limitations of the proposed methods in the previous literature. Creating a package license or patent pool can reduce much royalty rates compared with independent negotiation (Shapiro, 2000), which may be the most effective method of weeding out patent thickets (Ayres and Parchomovsky, 2007). A valid approach should achieve two purposes: measuring patent thicket density and detecting key patentees in patent thickets. The existing methods cannot detect key patent holders in patent thickets, though the density of patent thickets can be measured to some extent. For instance, the method of triples indicator fails to disclose key patentees with complementary patents in patent thickets. There is a need to develop a new method to detect patent thickets in given technology fields.

To fill this gap, the paper proposes a modified approach of measuring patent thickets by utilizing Social Network Analysis (SNA). Patent holders who are in patent thickets have the blocking relations if they cannot reach licensing agreements. SNA is a useful tool for studying various social relations. To further explore patent thickets, the paper proposes a novel method of combining triad census and data-driven social role analysis. Triads, configurations of triples of actors and relations between them (Faust, 2008), can be used to measure the density of patent thickets and depict the evolution of patent thickets over time. Moreover, data-driven social role analysis is helpful in detecting who could be potential licensors in patent thickets. Overall, the new method proposed by this paper is very suitable for measuring patent thickets in given technology fields instead of across technology areas.

The rest of this paper is organized as follows. Following the brief introduction, we have a short review of the literature on the definition of patent thickets and the methods of measuring patent thickets. In Section 3, a new method of measuring a patent thicket is introduced, and the application of such a method is presented in Section 4. Finally, a conclusion is drawn after discussions conducted in Section 5.

2. Literature review

The literature surveyed in this section is divided into two parts: the study on patent thicket theory and literature on measuring methods of patent thickets.

2.1. The theory of patent thickets

In the context of cumulative innovation, each inventor generally stands on “the shoulders of giants”. When complementary patents are owned by a few patent holders, the complements problem can be resolved by cross-license (Grindley and Teece, 1997). With the trend of patent explosion (Hall, 2004), however, the increasing complementary patents are dispersed among many patentees in given technology fields. Such a phenomenon is called patent thicket by Shapiro (2000).

When more than three patentees are concerned in a patent licensing negotiation, the transaction cost will sharply increase due to the hold-up problem and royalty stacking. As a result, a package license is hard to be reached due to patent thickets. In biomedical research, Heller and Eisenberg (1998) firstly proposed the hypothesis of anti-commons tragedy, in which a large number of complementary patents were underused but downstream manufacturers cannot make patented products.

Patent thickets become barriers for downstream manufacturers to entry in complex technology areas (Hall, 2013). Cockburn et al. (2010) provided evidence that patent thickets had a negative influence on the performance of product innovation when companies engaged in-licensing. Therefore, patent thickets may prevent commercial development of innovative devices, and block new research into follow-on inventions (Heller and Eisenberg, 1998).

However, there is a heated argument about what are patent thickets in theory. The definition of patent thickets has become more complex over time. The ambiguity of the term “patent thickets” raises a voice of questioning the existence of patent thickets in practice. The report from a committee in the U.S. National Research Council of the National Academies argued there was no substantial evidence for the existence of patent thickets in genetics (Barnett, 2016). This report shows both tracking and measuring patent thickets are very important. How to understand patent thickets may influence the method option of measuring patent thickets. Indeed, there are at least four definitions in the existing literature as shown in Table 1, which can be categorized into two types: strategic option or patent right fragmentation.

The first type of definition tries to understand patent thickets from the patent strategic perspective. In fact, the origin of the term “patent thicket” can be traced back to Teece (1986), who pointed out an innovator should apply for a series of adjacent patents based on his original patent. Since a single patent may hardly confer strong protection for promoting innovation, innovators need to possess an impenetrable thicket of patents. A patent thicket set up by innovators can successfully foil the attempt of competitors at inventing around the original patent (Somaya, 2003). From the perspective of business strategy, a patent thicket can be defined as a large number of related patents are owned by a patent holder (Teece, 1986). Moreover, Hegde et al. (2009) expanded the concept of “patent thicket” by applying the game theory. Companies endeavor to file numerous patent applications with lower-quality seeking to accumulate a patent thicket. In this context, a patent thicket can be understood as a large number of patents with low quality, which are usually vague or likely invalid, but little commercial value. Patent applicants play games with patent examiners in national patent offices and competitors in markets.

The second type attempts to define patent thicket as a phenomenon, rather than a strategic option. (Heller and Eisenberg, 1998) defined a patent thicket as a series of complementary patents owned by different patent holders in given technologies. To make and produce useful products, manufacturers should access to numerous complementary patents. Moreover, Shapiro (2000) introduced a patent thicket as an unintentionally dense web of overlapping patent rights that firms have to hack their ways through in order to actually commercialize new inventions. The key feature of this conception is that numerous patents are separately owned by many patentees. The overlap between patents include two special forms: horizontal overlap and vertical overlap. A horizontal overlap refers to numerous complementary patents are largely adjacent to one another, whereas a vertical overlap means a large number of patents are related through cumulative innovation (Burk and Lemley, 2003).

The main difference between these definitions is whether numerous patents are owned by many patentees. Both Teece and Hegde observed patent thickets from the perspective of enterprise patent strategy and emphasized a single organization owned a great number of patents to prevent competitors from imitation. However, Heller and Shapiro considered patent thickets as a phenomenon, and the attribute of patent right fragmentation was paid attention especially. It is worth noting that the definition proposed by Shapiro (2000) has been widely adopted in academia. Therefore, this paper adopts the definition of patent thicket introduced by Shapiro and highlights the following characteristics of a patent thicket: 1) a great number of patents exist in a given technology field; 2) these patents are complementing each other; 3) these patents are owned by more than at least three patentees.

2.2. The methods of measuring patent thickets

A growing literature has focused on measuring patent thickets and discussing the negative influences of patent thickets on technology innovation in recent years. However, the detection of patent thickets is limited by lacking reliable methods. An ideal measurement should analyze a series of patents in detail, and then identify whether patent rights are overlapping in given technology fields. Unfortunately, it is unavailable to directly measure patent thickets in practice.

The indirect evidence contained in patent citations can be used to measure patent thickets indeed. The state of the prior art is regarded as relevant for the patentability of an invention (Graevenitz et al., 2011). In this context, the indicators of identifying patent thickets generally build on patent citation data generated in the process of patent examination. So far, at least three methods have been proposed to detect patent thickets in the existing literature.

The first approach is to take the fragmentation index as a proxy of measuring patent thickets. Ziedonis (2004) proposed the fragmentation index to calculate the fragmentation of patent rights on the firm level. The fragmentation index originates from a normalized Herfindahl concentration index, which is often utilized for measuring the competitive level in a market. From a firm level, fragmentation index can show whether patents in given technology fields are concentrated in a few companies or dispersed among a great number of patent holders (Farrell et al., 2008). Moreover, Cockburn et al. (2010) extended the use of the fragmentation index as a proxy of measuring patent thickets to the technology level and found patent fragmentation had a negative influence on innovative performance. In general, the fragmentation index is very suitable for studying the effect of patent fragmentation from an empirical perspective. However, the fragmentation index is hard to measure patent thickets on a large scale due to the difficulty of collecting data. For instance, Ziedonis (2004) took 67 U.S. semiconductor firms as samples for calculating the fragmentation index. Moreover, the measurement result of using the fragmentation index may be imprecise. To meet a minimum size of patents owned by sample firms, Cockburn et al. (2010) had to establish a threshold value before measuring patent thickets. Many patentees without numerous patents will be filtered out due to the set threshold value. For instance, some firms that owned less than 100 patents are excluded as samples. Thus, the result of measuring patent thickets may be inaccurate if only large companies are taken as samples.

The second method is the triples indicator. Graevenitz et al. (2011) proposed that the number of “triples” could be a proxy of patent thickets on the basis of European patent critical citation data. In the process of patent examination, patent examiners in the European Patent Office (EPO) will reference prior art documents related to the novelty or inventive step of the claimed invention as critical patent citations. A critical reference is supposed to be a blocking patent in patent thickets. If a patent owned by a company has received a critical reference from the other firm, the cited reference would be regarded as a blocking patent. As a result, these two companies will block each other. A triple can be defined as a group of three patentees where each firm has critical patent citation limiting claims on patent documents of each of the other two firms (Graevenitz et al., 2011). If a great number of triples are detected, it suggests the patent thicket is dense in given technology fields. To demonstrate the effectiveness of the triples indicator, Graevenitz et al. (2011) illustrated that information technology and semiconductors had denser patent thickets than the other technology areas between 1980 and 2003. Futhermore, Fischer and Ringler (2015) modified the initial triples indicator by ignoring the distinction between critical and non-critical citations and found the density of patent thickets was different across patent systems. It implies that patent systems may influence the density of patent thickets in given technology fields.

The third method is derived from measuring citation network density by using a network analytic technique. Clarkson (2005) proposed the outdegree or indegree of citation networks can be used to calculate patent thicket density. Utilizing such this method, Zingg and M (2018) measured the patent thicket density in Europe. However, the outdegree or indegree of citation networks can merely reflect bilateral ties between related patents, rather than trilateral relationships among patentees. The key feature of patent thickets is that complementary patents are owned by different patent holders. This approach cannot accurately measure the density of patent thickets.

Overall, the triples indicator offers a simple way to measure patent thickets from a technical perspective, which is widely applied in the existing literature. Fischer and Henkel (2012) utilized the triples indicator to study whether patent trolls focused on acquiring patents in given technology fields with dense patent thickets. Moreover, Hall (2013) provided evidence that patent thickets in telecommunications, audiovisual technology, and computer technology were growing by using the triples indicator. Moreover, Graevenitz et al. (2013) used the triples indicator to prove patent thickets were more serious in complex technologies than discrete technologies in Europe.

However, the triples indicator cannot reveal the structure of patent thickets. Indeed, patent thickets are very complicated. A valid method should disclose key patent holders in patent thickets, which is helpful for manufacturers to navigate patent thickets. To close this gap, the paper seeks to combine triad census and data-driven social role analysis to detect patent thickets. In essence, measuring patent thickets is to reveal some special relationships between patent holders. SNA is a computational tool for us to understand important phenomena depending on networks (Kim and Song, 2013). A social role of community member means a qualitative description of a member's position within a social network, and the reasons it chooses to interact with others (Doran, 2015). In particular, data-driven social role analysis is a promising approach to discovering social roles in large-scale social networks. Large numbers of patents are accumulated in given technology fields over time. Thus, there is a need for adopting the method of data-driven social role analysis to detect patent thickets by analyzing thousands of patents.

3. The proposed method

From a theoretical perspective, patent thickets can be detected if some key features of citation networks are identified. First, there are numerous complemented patents in complex technologies. Technology areas can be divided into complex and discrete technologies. A complex technology refers to a new product or process is comprised of numerous patentable elements (Graevenitz et al., 2011). Patent thickets stem from the classic economic theory of complements. Patent thickets will arise when the complementary patents for given technologies or particular products are controlled by more than three patentees. These complemented patents are essential for new product development. If downstream manufacturers fail to get licenses from multiple patentees, complemented patents for given products will become blocking patents that prevent downstream manufacturers from entering markets. Thus, Shapiro (2000) called those complemented patents as essential patents or blocking patents, rather than substitute or rival patents. For research and development, the blocking or complementary relation is not always but often nothing more than a citation (Shapiro, 2000). In this context, the prior patents in the forward citations can be interpreted as complementary or blocking relations.

However, the rules of patent citation implemented by patent offices across countries or regions are significantly different. For instance, an examiner of the EPO will add prior art documents, which can be marked as either critical types or the other citation types, to a patent in the examination processes. Since the critical references are critical for the patent's novelty and inventive step, von Graevenitz et al. (2011) interpreted a critical reference as blocking relation. Nevertheless, an applicant has to list prior art himself at the USPTO. The differentiation between critical references and the other is not available for the USPTO examination process. In this regard, Fischer and Ringler (2015) interpreted all citations of the patents as blocking patents and demonstrated the modified method of triples indicator was valid. The purpose of this paper is to explore a novel method of measuring patent thickets across countries. To maximize the application of our proposed method, the paper adopts the interpretation of patent citation conducted by Fischer and Ringler (2015) and regards the cited patents in the forward citations as the complementary patents for subsequent patents.

Second, complementary patents are owned by different patent holders. In essence, the complements problem inherent in patent thickets is how to efficiently aggregate with numerous complementary patents owned by different companies (Shapiro, 2000). Patent thickets will not emerge if complementary patents are absent or concentrated in less three patentees. On the contrary, patent thickets will arise when complementary patents are dispersed among more than three companies. That is, patentees who own complementary patents in a given technology is fragmented.

An effective approach of accurately detecting patent thickets should meet two fundamental requirements: measuring patent thicket density and revealing key patent holders in patent thickets. The citation relations between patents help us identify complementary patents, while the information on assignees reflects a fact who own complementary patents. Thus, patent thickets can be composed of two networks: a patent citation network and a patentee network. Appendix 1 presents how to construct a new network by combining a citation network with a patentee network. For instance, patent "a" cites patent "b" and "c" but "a" is cited by patent "j", which forms a directed citation network. In addition, firm "A" owns three patents: a, x and y, which develop an undirected patentee network. The new directed patentee network can reflect the citation relations between patent holders. Inspired by Graevenitz et al. (2011), we utilize the technology of the triad census in SNA to explore tripartite relations among patent holders.

In general, a triad consists of a triple of nodes and the state of the arcs between them in a directed network. Fig. 1 presents the sixteen possible isomorphism classes of triads with the standard label. Each kind of triad is assigned three numbers or a letter to characterize its structure. In the standard labeling, three figures respectively mean the number of mutual dyads, asymmetric dyad, and null dyad. Moreover, the letter is applied to differentiate triads that are assigned the same 3-digit label, in which the letter "D" refers to the arrow points downwards, and "U" points upwards, while "C" forms a cycle, and "T" means transitive (Snir and Ravid, 2016). For instance, the triad 120D contains 1 mutual dyad and 2 asymmetric dyads that point downwards, while the triad 300 consists of three mutual dyads without asymmetric and null dyads.

Triad census is very suitable for investigating the structural properties of patentee networks. Triadic patterns can provide the basic sociological insights of network structure. Some fundamental graph properties can be derived from counting triads for given networks. Particularly, the triadic census can be used to not only characterize common triadic tendencies but also link triadic patterns to theoretically network structures, such as clusterability, structural balance, etc. (Faust, 2008). For instance, the structure of the triad 16–300 is balanced. In the triple, three patentees have a strong incentive to reach a licensing agreement in the case of complementarity. However, one patent holder may predominate in the intransitive triads, such as 11–201. As a result, patent licensing agreements are very difficult to be reached in intransitive triads than balanced triads.

In this paper, we adopt the modified method of triples indicator proposed by Fischer and Ringler (2015) to measure patent thickets. Fig. 2 presents the procedure of identifying patentee triples. If three patentees separately own patents that are cited each other, the three patentees compose a triple, namely 16–300 triad.

The paper proposes a new method of combining data-driven social role analysis and triad census to measure patent thickets. Fig. 3 depicts the key steps on how to measure the density and structure of patent thickets in given technology fields.

In step 1, the related patent data, including patent citations and patent assignees, are retrieved from the patent database. Many databases are available for retrieving patents. For instance, the USPTO and EPO database has been widely used in the existing literature due to rich patent citation data. In addition, the Derwent innovations index (DII) database has collected a large volume of patent documents in the world since 1963 (Luan et al., 2013).

Step 2 is to construct a directed patentee network by combining citation data with patentee data. See Appendix 1 for a detailed description of how to transform a citation network into a directed patentee network.

Then, sixteen kinds of triads in patentee networks are calculated in Step 3. According to the balance theory in SNA, all triads can be divided into balanced triads and intransitive triads. A network tends to be a structural balance if the proportion of balanced triads is becoming larger over time (Yuxi et al., 2012). Otherwise, the network is likely to be unbalanced.

In Step 4 the issue is posed, "Have detected 16–300 triad in the directed patentee networks? If the answer to the issue posed in Step 4 is "no", the approach will continue to Step 5, where the other technology candidates are considered. Otherwise, the method directly accesses to Step 6.

To reveal the roles of key patentees in networks, Step 6 conducts a data-driven social role analysis on patentee triple networks. In light of the similarity of features, a social role can be defined as a group that the algorithm places actors into (Doran, 2015). Either distributions of out-degree and in-degree or k-means clustering can be used to identify the role of actors in networks. Finally, the results of measuring patent thickets are interpreted and reported in step 7.

4. Application of the proposed method

To illustrate the validity of the proposed approach, the paper takes fuel cell electric vehicle (FCEV) as an example to measure patent thickets. The main reasons are as below. First, patent thickets are often triggered by the combination of complex technologies and numerous patent applications (Shapiro, 2000). The generic technologies of FCEV are composed of battery technology, motor technology, and powertrain system integration (Zhuravleva and Panichev, 2016), while special technologies involve fuel cells and equipment of storing hydrogen (Wesseling et al., 2014). A great number of patent applications have been accumulated in the complex technology field of FCEV.

Second, FCEV is concerned by governments as a kind of green technology (Barbieri, 2015). With the increasing concerns about the

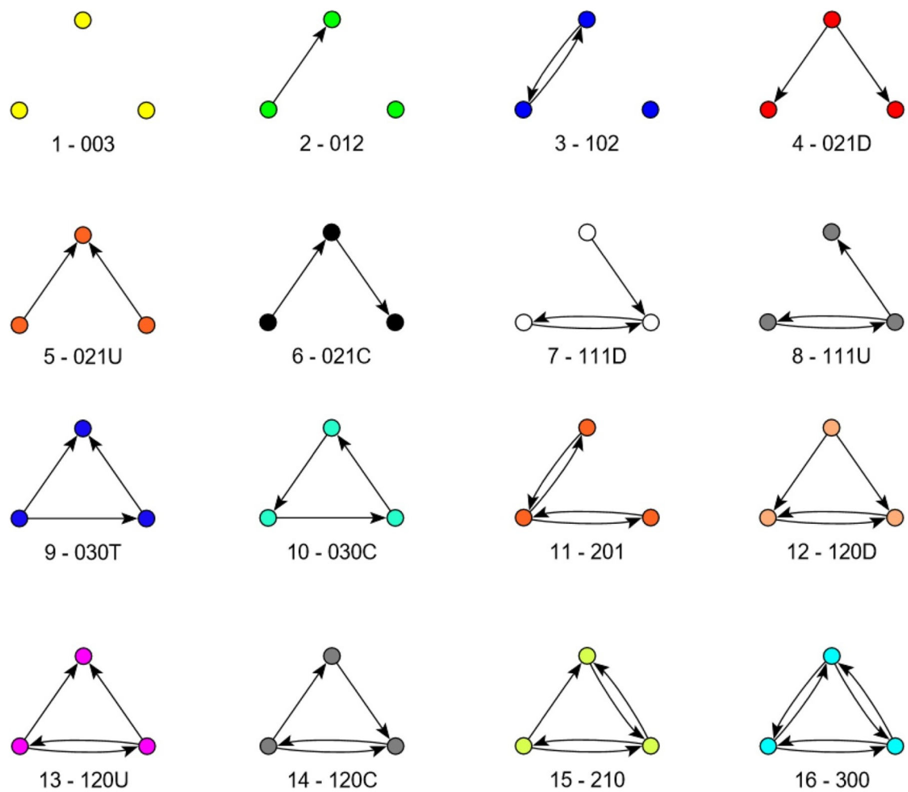


Fig.. 1. The sixteen possible kinds of triads.

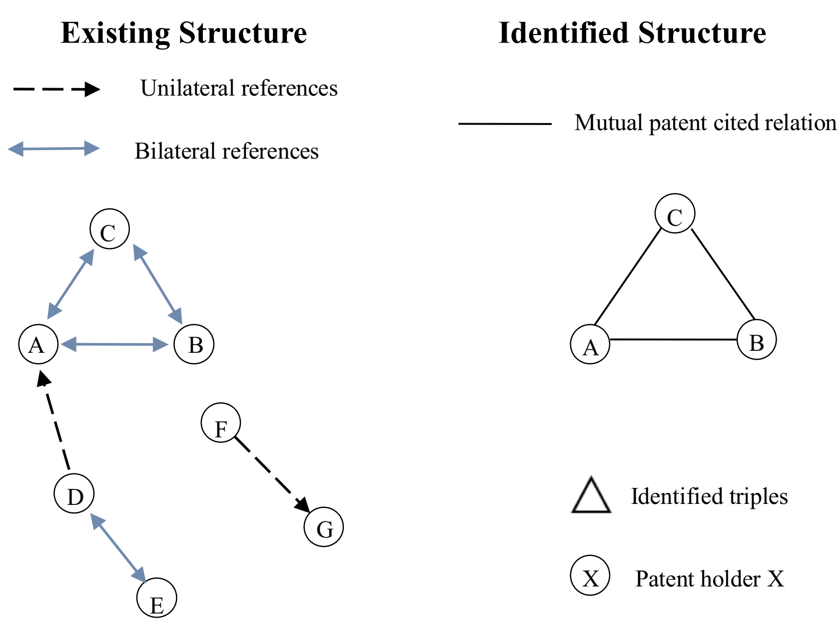


Fig.. 2. The procedure of identifying patentee triple.

environmental influence of internal combustion engine vehicles, the adoption of electric vehicles has been stimulated by environmental innovation policies. Barbieri (2016) found that both fuel prices and taxes might induce the development of low emission vehicles. Both battery electric vehicles (BEV) and FCEV are promising ways to reduce carbon emissions (Han et al., 2017). However, FCEV may be more environmentally friendly because it merely produces water and heat without harmful emissions. Taking London Taxi as an example, Baptista et al. (2011) demonstrated the vehicles configurated fuel cell power had lower energy consumption and less CO2 emission than BEVs. Offer et al. (2010) suggested the government should support FCEV and BEV because the best option might be the integration of FCEV with BEV.

Third, the pattern of developing FCEV provides an opportunity to

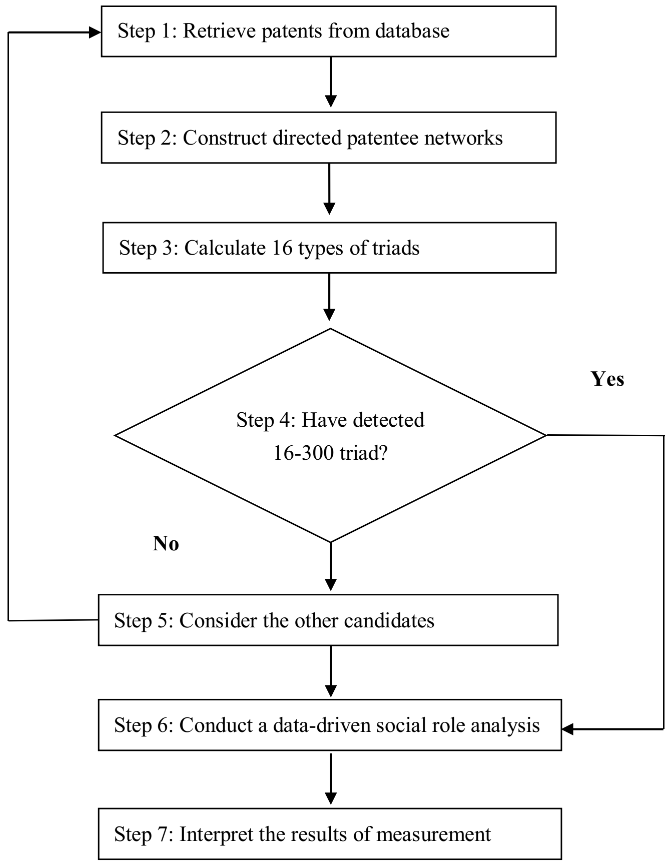


Fig. 3. The procedure of identifying patent thickets.

observe the evolution of patent thickets across stages. FCEV experienced a wave of rapid development during the period of the late 1990s through the mid-2000s, but the wave was not sustained (Wesseling et al., 2014). Additionally, the structure of innovators who participated in technology innovation is different between FCEVs and BEVs. While smaller companies and new entrants play an important role in developing BEV technology, large automobile manufacturers dominate the development of FCEV technology (Wesseling et al., 2014). On the other hand, the fuel cell technology in hybrid and electric vehicles is still in an emerging stage (Nagula, 2016), whereas the technological competition in BEVs is more intense than FCEVs. If we can successfully detect patent thickets in the field of FCEV technology, patent thickets should have arisen in given technology fields where competition is more intense.

4.1. Patent data retrieval

The first step is to retrieve patents regarding FCEV from the patent database under the guidance of the retrieval strategy. We adopt Thomson Reuters' Derwent innovations index (DII) to retrieve the related patents. As a famous comprehensive database, DII has collected a large volume of patent documents in the world since 1963 and updates every week about 25,000 patent documents published by more than 40 patent offices and 45,000 patent citation documents from six important offices (Luan et al., 2013). DII is very suitable for measuring patent thickets in the FCEV technology field. First, patent citations are given that allow measurement of patent thickets. Second, there is a precise record of patent families. Some applications or patents that are related to each other but filed in several countries or regions are generally known as a patent family (OECD, 2010,). It is possible for us to measuring patent thickets across countries or regions. Lastly, a four-letter code is assigned to each applicant by experts manually standardizing individual or firm names (Thomson Reuters, 2009). Indeed, it is crucial to accurately identify patent holders in the process of measuring patent thickets. Experts of Thomson Reuters check and identify firm names including their subsidiaries and related firms. Patents will be assigned to a parent firm if it owns a majority share of another company. This allows us to accurately identify patentees who owned relevant patents in patent thickets.

Fischer and Ringler (2015) provided evidence that the density of patent thickets was different across patent systems. Since the regulation of patent citations differs, we suppose that patent thickets in the FCEV technology field could be different between the USPTO and EPO. Therefore, we focus on measuring patent thickets based on applications

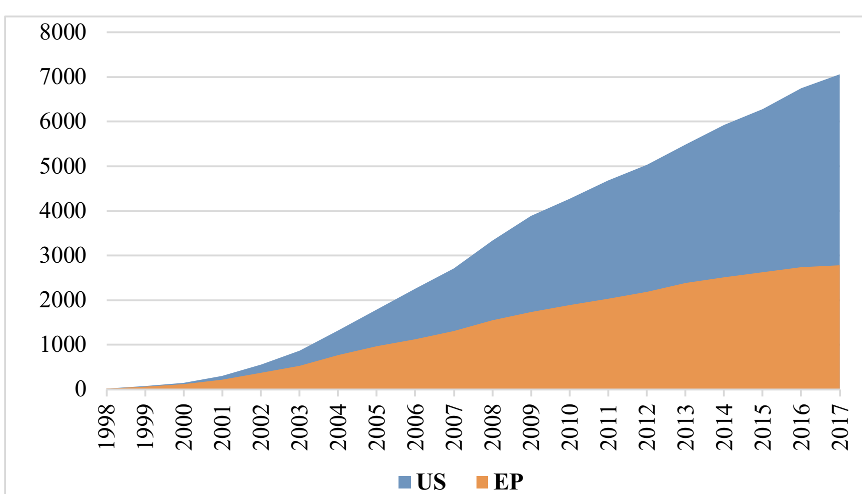


Fig. 4. The accumulated number of applications at the USPTO and EPO. Legend: US = applications at the USPTO, EP = applications at the EPO.

Table 2
The triadic pattern of patentee networks in the U.S. and Europe (2009–2014).

Type	Number of triads (US)	Number of triads (EP)	Model
3 - 102	2,601,984	65,371	Balance
16 - 300	**568**	**30**	**Balance**
1 - 003	17,146,459,749	186,147,005	Clusterability
4 - 021D	402,056	30,762	Ranked clusters
5 - 021U	1,474,776	27,404	Ranked clusters
9 - 030T	30,071	1338	Ranked clusters
12- 120D	9783	261	Ranked clusters
13- 120U	5867	507	Ranked clusters
2 - 012	107,367,152	2,768,211	Transitivity
14- 120C	2164	85	Hierarchical clusters
15 - 210	3349	142	Hierarchical clusters
6 - 021C	882,055	45,953	Intransitive pattern
7- 111D	297,185	7991	Intransitive pattern
8- 111U	132,373	8517	Intransitive pattern
10- 030C	537	19	Intransitive pattern
11- 201	20,471	648	Intransitive pattern

filed at USPTO and EPO.

In this study, we adopt a search strategy of combining keywords and patent technical classifications. The specific keywords in the title of patent documents provide a high number of patents that are closely related to FCEV technology, while IPC classes are used to differentiate distinct technologies and further exclude irrelevant patents (Aghion et al., 2016). This paper follows the search strategy for FCEV technology proposed by Borgstedt et al. (2017), which is IP = (B60W-010/28 OR B60L-011/18 OR H01M-008*) AND TS = (vehicle* OR car OR automobile*) AND TS = (fuel cell*). The time span was set to 1998 –2017. In total, the dataset consists of 7058 applications at the USPTO and 2776 at the EPO, which are shown in Fig. 4. The patent records were downloaded on 30, Dec. 2018. Additionally, the software Python is used to extract the necessary information, such as assigners, patent citations and filing dates, etc., from patent records.

4.2. Triad censuses of patent holders

In step 2, we construct directed patentee networks by the combination of citation data and patentee information. Utilizing the software Python and Pajek, we extracted the basic information from patent records and then constructed two directed patentee networks: the U.S. and Europe. Step 3 is to calculate 16 types of triads in patentee networks and analyze whether the structure of networks is balanced. To clearly display the results, the interval time between filing data and citation had better for more than five years. Table 2 describes the results of triad censuses in the U.S. and Europe between 2009 and 2014.

It is worth noting that the triad 16–300 is equal to the “triple” called by Graevenitz et al. (2011). The algorithm of measuring patent thicket density is to calculate the number of triples where patentees own patents mutually cited. The number of triples, the triad 16–300, is 568 in the U.S., while only 30 in Europe between 2009 and 2014. However, a great number of intransitive triads are also detected, such as 6–021C, 7–111D, 8–111 U, 10–030C and 11–201.

Repeating the above steps, we depict the evolution of patent thickets in the U.S. and Europe, as shown in Fig. 5. According to the view of Wesseling et al. (2014), the wave of developing FCEV continued until the mid-2000s. It is interesting the number of triad 16–300 peaked at 1365 in the U.S. during the period of 2002–2007, while reached 60 in Europe between 2000 and 2005. As the wave of FCEV development was interrupted, the number of the triad 16–300 declined gradually. To simplify, we adopt peak time and low period to conduct the next analysis.

4.3. Data-driven social role analysis

In Step 4, we conducted a data-driven social role analysis on patentee triple networks. To simplify, we adopt the method of calculating out-degree or in-degree to identify actor roles in networks. Since the triad 16–300 is balanced, out-degree is indeed equal to in-degree in patentee triple networks. Fig. 6 depicts the distribution of patent holders in the patentee triple networks over time.

Table 3 reports the comparative results of network characteristics between the U.S. and Europe. In the U.S., the peak time of patent thickets was the period of 2002–2007, in with the number of the triad 16–300 reached 1365. In this period, 253 patent holders were detected in the networks, in which the density of networks was 0.027 but the average degree was 13.94. However, the number of the triad 16–300 dropped to 568 in a low period of 2009–2014. There were 166 patentees in the networks in total, while the average degree was 11.25.

By contrast, the number of the triad 16–300 peaked at 60 in Europe

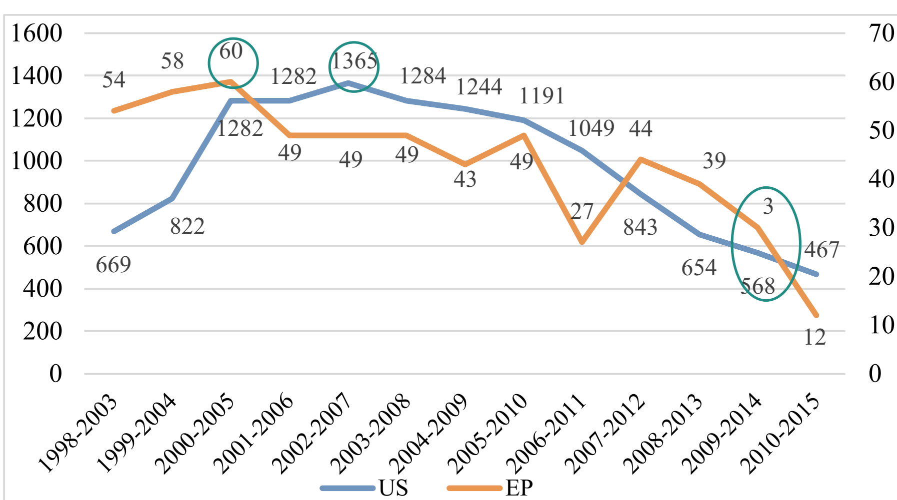


Fig. 5. The evolution of patent thickets in the U.S. and Europe.

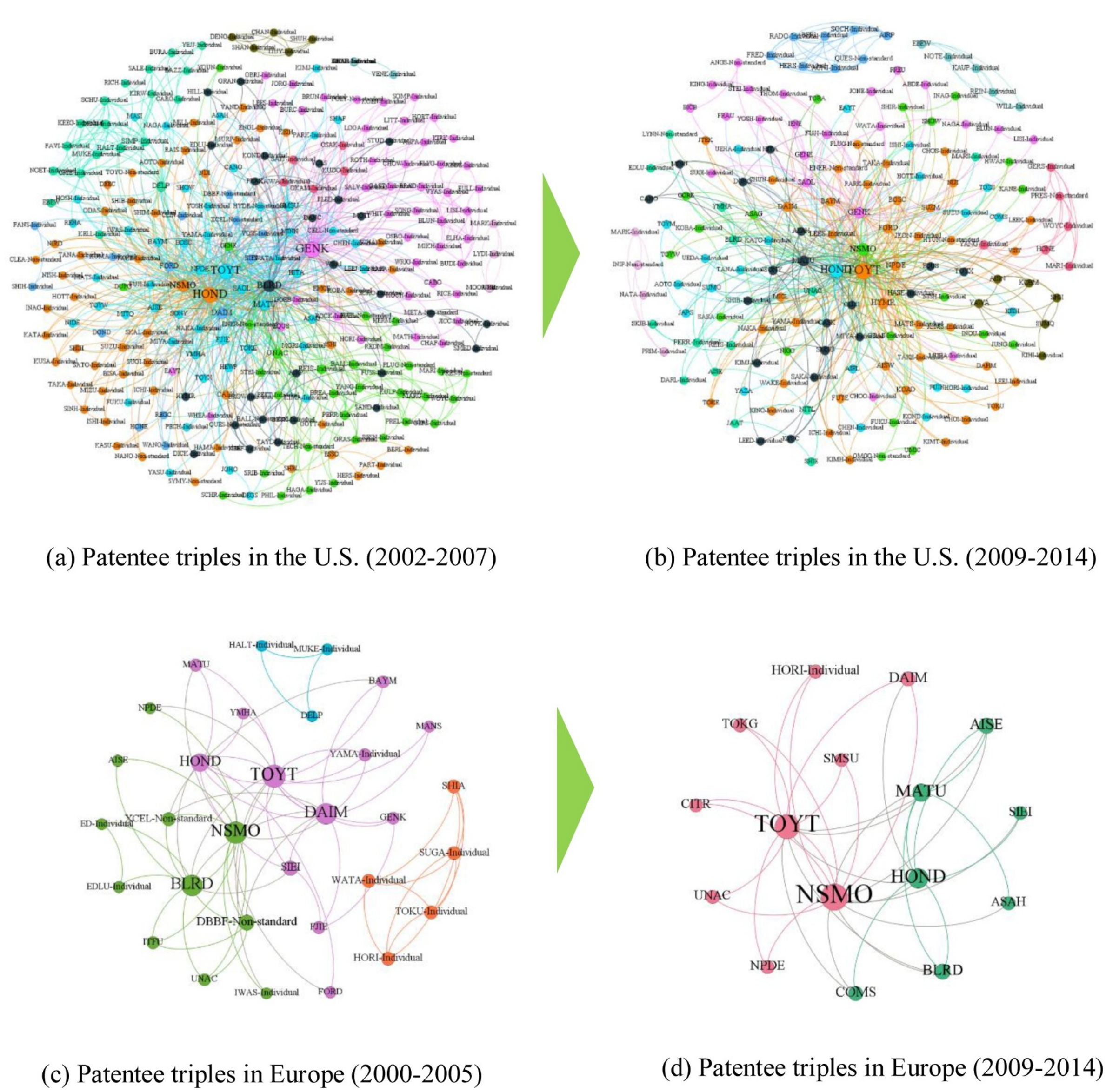


Fig. 6. Patentee triples in the U.S. and Europe at different stages.

Table 3
Description of network characteristics in different periods.

	Periods	Number of triples	Number of vertexes	Total number of lines	Network density	Average degree
US	2002–2007	1365	253	1763	0.028	13.93
	2009–2014	568	166	964	0.035	11.61
EU	2000–2005	60	31	139	0.149	8.97
	2009–2014	30	16	72	0.300	9.00

during the period of 2000–2005. There were 31 patent holders in the networks, while the density of networks was 0.149 but the average degree was 8.97. However, the number of the triad 16–300 was only 30 during the low period of 2009–2014. In this period, 15 patentees were detected in the networks but the density of networks was 0.300.

In general, the degree centrality of a given node, which is defined as the number of links incident upon the given node, is used to detect important actors in a special network. However, degree centrality is often normalized by the maximum possible degree in order to compare different networks. Table 4 reports the top 11 patent holders based on degree centrality and normalized degree centrality.

It is interesting that most patent holders who occupy important positions in networks are the same across periods. In the U.S., nine out of eleven firms are the same during the period of 2002–2007 and 2009–2014. However, the order of sorting patentees slightly changes. For instance, General Motors Corp (GENK) was the top company in the period of 2002–2007, followed by the Japanese automotive manufacturers Honda Motor Company (HOND) and Toyota Motor Corporation (TOYT). A few years later, Toyota Motor Corporation listed the top one, then followed by Honda Motor Company (HOND) and Nissan Motor Co., Ltd (NSMO).

A similar case emerged in Europe. In the peak period of 2000–2005, the top three companies were Toyota Motor Corporation (TOYT), Nissan Moto Co., Ltd (NSMO) and Ballard Power Systems Inc (BLRD), while Nissan Moto Co., Ltd (NSMO) listed on the top during the low period of 2009–2014. However, seven out of eleven firms are the same in peak time and low period. It is notable that six companies listed in the top 11 patent holders are the same across the networks.

4.4. Interpret results of measurement

The last step is to interpret the results of measuring patent thickets.

Table 4
Distribution of top patent holders in the U.S. and Europe.

The U.S. (2002–2007)			The U.S. (2009–2014)			Europe (2000–2005)			Europe (2009–2014)		
Firms	**Degree**	**Normalized Degree**	**Firms**	**Degree**	**Normalized degree**	**Firms**	**Degree**	**Normalized degree**	**Firms**	**Degree**	**Normalized degree**
GENK	116	0.460	TOYT	107	0.648	TOYT	14	0.467	NSMO	15	1.000
HOND	110	0.437	HOND	70	0.424	NSMO	13	0.433	TOYT	13	0.867
TOYT	98	0.389	NSMO	52	0.315	BLRD	12	0.400	HOND	8	0.533
BLRD	64	0.254	GENK	35	0.212	DAIM	12	0.400	MATU	6	0.400
NSMO	56	0.222	MATU	33	0.200	HOND	9	0.300	AISE	4	0.267
DAIM	40	0.159	HYMR	29	0.176	DBBF	6	0.200	BLRD	4	0.267
MATU	39	0.155	DAIM	18	0.109	SIEI	5	0.167	DAIM	3	0.200
UNAC	30	0.119	FORD	14	0.085	SHIA	4	0.133	COMS	3	0.200
DELP	25	0.099	NPDE	14	0.085	XCEL	4	0.133	NPDE	2	0.133
FORD	24	0.095	BLRD	13	0.079	UNAC	3	0.100	UNAC	2	0.133
NPDE	23	0.091	SMSU	13	0.079	MATU	3	0.100	SMSU	2	0.133

Overall, the number of the triad 16–300 is declining gradually after the peak in the U.S. and Europe. This suggests that the density of patent thickets does not become denser but is weakening. The primary reason may be patent applications regarding FCEV technology do not keep a rapid growth. In light of Wesseling et al. (2014), the wave of developing FCEV technology interrupted in the mid-2000s. Thus, the evolution of patent thickets coincides with the wave of technological development. However, great numbers of intransitive triads are detected across the networks in the U.S. and Europe. From the perspective of network structure, these triads are unbalanced. It implies that some patent holders have predomination over the others. If a new wave of technology development is triggered, patent thickets in FCEV technology will become denser.

Additionally, the important patentees have been detected by using data-driven social role analysis. We find the key patent holders across the networks are large companies by sorting the degree centrality. Many famous vehicle manufacturers hold the dominant positions in patent thickets, such as Toyota, Honda, General Motors, and Mercedes-Benz. This conclusion coincides with the view of Wesseling et al. (2014), in which large automobile manufacturers dominate the development of FCEV technology. It suggests that the important patent holders, which occupy dominant positions in patent thickets, may be potential licensors for downstream manufacturers who seek to hack through patent thickets.

5. Discussion

The results show the proposed method is an efficient and convenient way of measuring patent thickets.

5.1. Comparing with the previous approaches

To confirm the validity of our method, we compare it with the previous approaches. First, Ziedonis (2004) proposed the “fragmentation index” to measure patent thickets. The arithmetic is based on the number of backward citations in a company's patent portfolio. Suppose there are N companies that have patent applications in a given technology field j, in which company i has referred to patents held by company k. The formula for calculating the fragmentation index is shown below.

$$\text{Fragementation} = \frac{1}{N}\sum_{i=1}^{N}\left\{\left(1-\sum_{k=1}^{k}\left(\frac{references_{ijk}}{references_{ij}}\right)^{2}\right)\left(\frac{references_{ij}}{references_{ij}-1}\right)\right\} \tag{1}$$

In a given period, References ij records the number of references in a company i’s applications in technology j, while Referencesijk means the total number of references in a company i’s applications that refer to patents held by company k in technology j.

It is very suitable for small samples in given technology fields, but not fit the large scale network analysis. For the FCEV technology, for instance, 4697 patent holders are detected in the U.S. citation network, while 1044 patentees in Europe between 2009 and 2014. With so much data, it is hard to use the fragmentation index to measure patent thickets on a large scale. second, the principle of network density approach proposed by Clarkson (2005) is to calculate the network density in given citation networks. Suppose there are g actors in directed citation networks, where x_{ij} records a tie from actor i to j. The arithmetic is shown as below.

$$\text{Density} = \frac{\sum_{i=1}^{g}\sum_{j=1,i\neq j}^{g}x_{ij}}{g(g-1)} \tag{2}$$

We calculated the network density on the basis of the same data. In the period of 2009–2014, the network density is 0.00114 while there are 711 patents with 546 lines in Europe. By contrast, the network density in the U.S. is 0.000079 while there are 33,842 ties between

Table 5
The comparison of our method with previous approaches.

Methods	Measuring patent thickets	Detecting key patentees
Fragmentation index	Suitable for small samples on the firm level	–
Patent citation network density	Calculate citation network density on a tech. level	–
Triples indicator	Measure patentee triples on a macro level	–
The proposed method	Triad censuses on a patentee level	Data-driven social role analysis on a micro level

20,682 patents. The results show that the number of ties between patents in the U.S. is far more than in Europe, while the network density in the U.S. is less than in Europe. The density of a given network is negatively correlated with the network scale. With the increase of network scale, the network density is decreasing and vice versa. This implies the patent thickets in the U.S. and Europe are not dense when network scales are considered. The results achieved by our method is in line with the outcomes calculated by the method of citation network density.

As our method is inspired by the triples indicator proposed by von Graevenitz et al. (2011) and (Fischer and Ringler, 2015), our analysis result on the triad 16–300 is also in line with the conclusion drawn by using the method of triples indicator.

Table 5 shows the comparison of the proposed method with previous approaches. The proposed method is very suitable for detecting patent thickets in given technology fields on a micro-level. There are two important features of the proposed method.

Our method can measure patent thicket density in a more accurate and comprehensive way. The rationale for using the triad census is that the bargain cost of patent licensing in triads is more than bilateral relations due to the complements problem. Triads can be divided into two types: balanced and intransitive triads. In a balanced triad, patent holders could have an incentive to reach licensing agreements. The triad 300 represents mutual relations between three parties who are much willing to reach license agreements. By contrast, one patentee may have a strong predomination over the others in intransitive triads, which leads to the complements problem is hard to be resolved. Thus, calculating sixteen triads can help to more accurately and comprehensively measure patent thickets.

More importantly, our method can detect key patentees in given patent thickets. Patent thickets can be divided into “actual patent thickets” and “potential patent thickets” (Farrell et al., 2008). A series of problems could arise when “actual patent thickets” formed. Thus, a valid measuring method should detect key patentees who may be potential licensors in “potential patent thickets” or prevent the emergence of “actual patent thickets”. The proposed method uses data-driven social role analysis technology to identify key patent holders in patent thickets, who would be potential patent licensors. Downstream manufacturers may hack through patent thickets if they get licenses from key patent holders.

5.2. Implications for management

The approach proposed by this paper can help companies further detect patent thickets and seek potential patent licensors in order to hack through patent thickets. Our research contributes to the ongoing debate on how to measure patent thickets in given technology fields.

Downstream manufacturers, who plan to enter given technology fields, can use the proposed method to detect whether patent thickets exist in given technology areas. Both the density and structure of patent thickets may change over time. Triad censuses can reveal the density of patent thickets, which can be utilized to forecast the evolutional trend of patent thickets. Downstream companies can enter the given technology and develop new products if the density of patent thickets is not dense. Otherwise, manufacturers should consider how to get patent licensing. For instance, the density of patent thickets has weakened since the mid-2000s in FCEV technology. The blocking effect of patent thickets on green technology innovation is influenced by the density of patent thickets. The evolution of patent thickets shows that nowadays it may be a good opportunity for new entrants to enter the FCEV technology.

Moreover, the proposed method can be also used to identify the potential licensors in the given technology. Data-driven social role analysis can reveal who has competitive advantages in patent thickets. Patent holders who occupy dominant positions in patent thickets tend to reach cross-licensing in order to trek through patent thickets. Moreover, our method can also help downstream manufacturers without the complemented patents know who are potential licensors. For instance, many large automobile companies, who hold the dominant positions in patent thickets, maybe the potential licensors, such as Toyota, Honda, General Motors, and Mercedes-Benz, for downstream manufacturers in the FCEV sector.

6. Conclusion

How to prove the existence of patent thickets in complex technology areas has attracted a great concern in academia since Shapiro (2000) put forward the theory of patent thickets. A valid and simple method should have two fundamental functions: measuring patent thickets and revealing key patentees in patent thickets. The existing methods can measure the density of patent thickets to some extent but cannot identify important patent holders in patent thickets. To close this gap, the paper proposes a novel approach combining triad census and data-driven social role analysis to measure patent thickets. Taking FCEV patents at the USPTO and EPO as examples, the paper demonstrates the proposed method is valid and feasible in practice. Both patent thicket density and key patent holders in patent thickets can be accurately detected. This can help companies make a decision whether they should enter a given complex technology area or who could be potential licensors in patent thickets. Therefore, researchers or company managers can use the proposed approach to comprehensively measure patent thickets over time from a microscopic perspective.

There are also some limitations to our proposed method. So far, all the methods of measuring patent thickets depend on patent citation data. There is no exception for our method. However, patent citation rules between the USPTO and EPO are very different. In the U.S., applicants have to list prior art by themself, whereas only patent examiners can add references to patents at the EPO (Fischer and Ringler, 2015). As a result, the average number of citations in each patent at the USPTO is more than the other countries or regions. Indeed, many national patent offices in the world adopt the latter mode, namely references are added by patent examiners at the patent office. The conclusion that the patent thicket of FCEV technology in the U.S. is denser than Europe may partly attribute to the difference of citation rules. To reduce this bias, it may be more appropriate for exploring patent thickets in one given country or region in practice. Additionally, the proposed method is very suitable for measuring patent thickets on a micro-level, rather than a macroscopic analysis across technology areas.

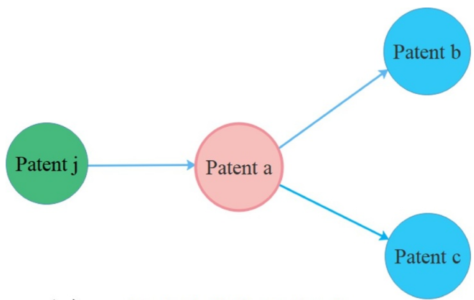
CRediT authorship contribution statement

Xiaodong Yuan: Conceptualization, Methodology, Software, Writing - review & editing. **Xiaotao Li:** Data curation, Writing - original draft.

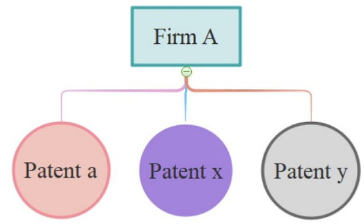
Acknowledgment

This paper was supported by the National Natural Science Foundation of China (Project No. 71774058).

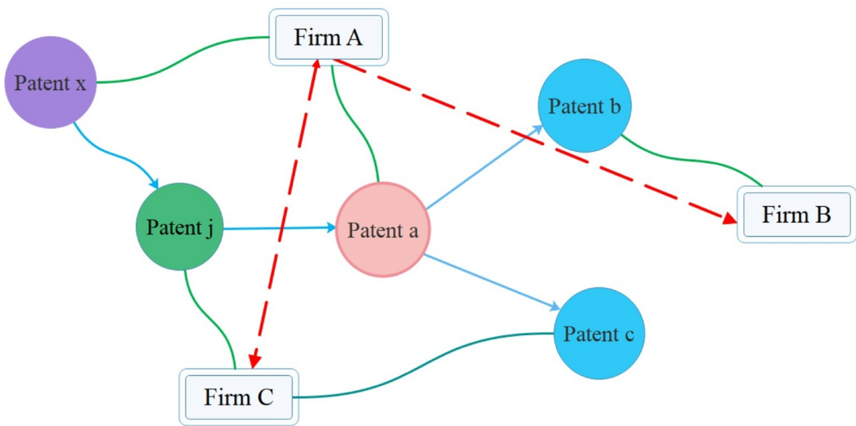
Appendix 1. How to construct a new directed patentee network



(a) A patent citation network



(b) A patentee network



(c) A new network by combining a citation network with a patentee network

References

Aghion, P., Dechezlepretre, A., Hemous, D., Martin, R., Van Reenen, J., 2016. Carbon taxes, path dependency, and directed technical change: evidence from the auto industry. J. Polit. Econ. 124, 1–51.
Ayres, I., Parchomovsky, G., 2007. 'Tradable patent rights'. Stanf. Law Rev. 60, 863–894.
Baptista, P., Ribau, J., Bravo, J., Silva, C., Adcock, P., Kells, A., 2011. Fuel cell hybrid taxi life cycle analysis. Energy Policy 39, 4683–4691.
Barbieri, N., 2015. 'Investigating the impacts of technological position and European environmental regulation on green automotive patent activity'. Ecol. Econ. 117, 140–152.
Barbieri, N., 2016. Fuel prices and the invention crowding out effect: releasing the automotive industry from its dependence on fossil fuel. Technol. Forecast. Soc. Change 111, 222–234.
Barnett, J.M., 2016. 'ARE there really patent thickets?'. Regulation 39, 14–17.
Bondibene, 2012. 'A study of patent thickets'. Niesr Discuss. Papers.
Borgstedt, P., Neyer, B., Schewe, G., 2017. Paving the road to electric vehicles - A patent analysis of the automotive supply industry. J. Clean. Prod. 167, 75–87.
Burk, D.L, Lemley, M.A, 2003. Policy levers in patent law. Virginia Law Review 89 (7), 1575–1696.
Clarkson, G., 2005. 'Patent informatics for patent thicket detection: a network analytic approach for measuring the density of patent space'. Acad. Manag.
Cockburn, I.M., Macgarvie, M.J., Muller, E., 2010. Patent thickets, licensing and innovative performance. Ind. Corp. Change 19, 899–925.
Doran, D. (2015) 'On the discovery of social roles in large scale social systems', social network analysis and mining, 5, 18.
Entezarkheir, M., 2017. Patent thickets, defensive patenting, and induced R&D: an empirical analysis of the costs and potential benefits of fragmentation in patent ownership. Empir. Econ. 52, 599–634.
Farrell, J., Jaffe, .A..B.., Lerner, J, Stern, S, 2008. 'Intellectual property as a bargaining environment'. Innovation Policy and the Economy 9. Univ Chicago Press, Chicago.
Faust, K., 2008. 'Triadic configurations in limited choice sociometric networks: empirical and theoretical results'. Soc. Netw. 30, 273–282.
Fischer, T., Henkel, J., 2012. 'Patent trolls on markets for technology – An empirical analysis of NPEs' patent acquisitions'. Res. Policy 41, 1519–1533.
Fischer, T., Ringler, P., 2015. 'The coincidence of patent thickets—A comparative analysis'. Technovation 38, 42–49.
Graevenitz, G., Wagner, S., Harhoff, D., 2013. 'Incidence and growth of patent thickets: the impact of technological opportunities and complexity'. J.Ind. Econ. 61, 521–563.
Grindley, P.C., Teece, D.J., 1997. Managing intellectual capital: licensing and cross-licensing in semiconductors and electronics. Calif. Manag. Rev. 39, 8–41.
Hall, 2004. 'Exploring the patent explosion'. J. Technol. Transf. 30, 35–48.
Hall, E., 2013. 'Technology entry in the presence of patent'. Work. Papers.
Han, L., Wang, S., Zhao, D., Li, J., 2017. The intention to adopt electric vehicles: driven by functional and non-functional values. Transp. Res. Part A 103, 185–197.
Hegde, D., Mowery, D.C., Graham, S.J.H., 2009. Pioneering inventors or thicket builders: which us firms use continuations in patenting? Manag. Sci. 55, 1214–1226.
Heller, Eisenberg, 1998. Can patents deter innovation the anticommons in biomedical research. Science 280, 698–701.
Kim, H., Song, J., 2013. Social network analysis of patent infringement lawsuits'. Technol. Forecast. Soc. Change 80, 944–955.
Lemley, M.A., Shapiro, C., 2007. Patent holdup and royalty stacking. Tex. Law Rev. 85, 1991–2049.
Luan, C.J., Liu, Z.Y., Wang, X.W., 2013. 'Divergence and convergence: technology-relatedness evolution in solar energy industry'. Scientometrics 97, 461–475.
Nagula, M., 2016. 'Forecasting of fuel cell technology in hybrid and electric vehicles using Gompertz growth curve'. J.Stat. Manag. Syst. 19, 73–88.
Oecd, 2010. OECD Patent Statistics Manual'. M2PressWIRE.
Offer, G.J., Howey, D., Contestabile, M., Clague, R., Brandon, N.P., 2010. Comparative analysis of battery electric, hydrogen fuel cell and hybrid vehicles in a future sustainable road transport system. Energy Policy 38, 24–29.
Shapiro, 2000. Navigating the patent thicket: cross licenses, patent pools, and standard setting. Innov. Policy Econ. 1, 119–150.
Snir, R., Ravid, G., 2016. Global nanotechnology regulatory governance from a network analysis perspective. Regul Gov. 10, 314–334.
Somaya, D., 2003. 'Strategic determinants of decisions not to settle patent litigation'. Strat. Manag. J. 24, 17–38.
Teece, 1986. 'Implications for integration, collaboration, licensing and public policy'. Res. Policy 15, 285–305.
Tullis, T.K., 2012. 'Current intellectual property issues in nanotechnology'. Nanotechnol. Rev. 1, 189–205.
Zimmeren, Van, Verbeure, 2006. A clearing house for diagnostic testing: the solution to ensure access to and use of patented genetic inventions? Bull. World Health Organ. 84, 352–359.

Von Graevenitz, Wagner, Harhoff, 2011. 'How to measure patent thickets—a novel approach'. Econ. Lett. 111, 6–9.
Wesseling, Faber, Hekkert, 2014. 'How competitive forces sustain electric vehicle development'. Technol. Forecast. Soc. Change 81, 154–164.
Yuxi, H., Doroud, M. & Wu, S.F. (2012) 'On a triadic approach to connect microstructural properties to social macrostructural patterns', 353–59.
Zhuravleva, N.A., Panichev, A.U., 2016. Megatrends of global transportation systems development: the Russian segment, Zilina, Univ Zilina. Fac. Op. Econ. Transp. Commun.
Ziedonis, R.H., 2004. Don't fence me in: fragmented markets for technology and the patent acquisition strategies of firms. Manag. Sci. 50, 804–820.
Zingg, R., M, F., 2018. The nanotechnology patent thicket revisited. J. Nanopart. Res. 20, 267.

Xiaodong Yuan got his Ph.D. Huazhong University of Science and Technology, China. Now, he is a professor at School of Management, Huazhong University of Science and Technology, China. His-research interests cover technology forcast, patent analysis and patent management. In recent years, he focuses on studying patent thickets across technology areas.

Xiaotao Li is an associate professor at School of Literature, Law and Economics, Wuhan University of Science and Technology, China. Her research interests include technology innovation, patent analysis and patent policy. In recent years, she focuses on researching electric vehicles technology change and evolution.