Finding the Signal in the Noise of Patent Citations: How to Focus on Relevance for Strategic Advantage

Derek Smith

“ We must not always judge of the generality of the opinion by the noise of the acclamation.”

Edmund Burke (1729–1797)
Statesman, political theorist, and philosopher

Patent evaluations based on prior art citation data are business methodologies that can reveal hidden relationships between a patent and the associated prior art citations. These hidden relationships can further identify firms, actors and, technology and can identify strategic business relationships and opportunities. However, a fundamental problem with existing methodologies for discovering relevant relationships in citation data is noise. Citation noise obscures relevant relationships and impedes the potential value of these methodologies. This article reviews the literature on prior art citations as they relate to citation-based patent evaluation methodologies. A framework is presented to overcome citation noise and reveal relevant relationships to evaluate knowledge, technology, and invention properties, thereby opening up strategic and competitive advantages to the evaluating firm. The article provides guidance to executives and managers to help improve strategic decision making through relevant citation-based patent evaluations.

Introduction

Any decision to obtain a patent requires careful and strategic business consideration because patents are very expensive and they do not always generate the anticipated business value. Value may be limited in a business sense due to a firms capabilities or misalignment to the strategic vision. The overriding challenge is therefore to create as much strategic or competitive advantage as possible from each patent.

The value of a patent can be assessed in a number of ways, many of which focus on the relationship of the patent to its citations of older technology information, or "prior art". A citation-based patent evaluation can provide strategic or competitive advantage if it can reveal relevant, hidden relationships and an assessment of: i) knowledge flow, ii) technology properties, or iii) invention properties.

Patents and their prior art citations enable this obscure form of strategic or competitive advantage based on the underlying business and technology relationships that may be identified from the relevant prior art citations. Extracting this form of strategic advantage requires revealing the relevant relationships that can provide insight pertinent to the strategic or competitive advantage of a firm.

Extant research on prior art citations, their applicability, and limitations to patent evaluation methodologies has focused on individual, distinct applications. However, many authors (e.g., Hu & Jaffe, 2003; Gay et al., 2005; Criscuolo & Verspagen, 2008) also identify citation noise as a problem that limits the effectiveness of these methodologies.

Citation noise occurs due to timing in the patenting process, the point in time when the prior art references are identified and the amended state of the patent claims at that point in time. Citation noise also occurs due to different perspectives and comprehension concerning the technology and invention. Prior art references may be provided by many actors such as the
Finding the Signal in the Noise of Patent Citations: How to Focus on Relevance

Derek Smith

inventor with the patent application, a searcher during a patent office search phase, an examiner during a patent examination phase, or an interested third party after publication of the application. Inventor-supplied citations are typically addressed before filing a patent application and are typically not relevant at the time of filing the patent application. Searcher-supplied citations and third party supplied citations may arrive later, after amendments. They may not be relevant due to the preceding amendments that redefine the invention. Together, all citations that have been identified form the raw citation data associated with the patent. The differing perspectives and comprehensions are resolved during prosecution of the patent by the examiner and patent agent. During this process, the relevant prior art citations are identified from the raw citation data based on the patent claims or amended state of the patent claims and reaching agreement on the final form of the patent claims.

Citation noise arises from prior art citations in the raw citation data that are not relevant to the patent, and it can be substantial. As Jaffe and colleagues (2000) observed: "There is, however, a large amount of noise in citations data; it appears that something like one-half of all citations do not correspond to any perceived communication, or even necessarily to perceptible technological relationship between the inventions." Unfortunately, existing citation-based patent evaluation methodologies lack an effective way of filtering out this noise. If there was a means of discovering the relevant prior art citations, or filtering out the citation noise, the effectiveness of the various evaluation methodologies could be improved.

The author of the present article previously described how citation information, business information, and patent information might be combined in a patent evaluation to reveal insight by identifying four key business considerations: strategic partners, strategic innovation, strategic invention, and strategic linkages (Smith, 2014a). In this article, the author describes how citation information may be filtered to discover relevant citations from the noise, thereby revealing insight from three key perspectives: knowledge flow, technology relationships, and relative degree of invention. A framework is also provided that illustrates how discovering the relevant prior art citations from the noise can lead to clear outcomes, thereby assisting executives and managers with the strategic assessment of opportunities concerning an individual patent or portfolio of patents.

The remainder of this article is organized as follows. The first section reviews the literature concerning prior art citations, citation noise, and the interdependency between prior art citations and patents. The second section describes a patent evaluation framework based on relevant citations (i.e., those discovered from the noise of citation data), which is accompanied by three example scenarios. A final section concludes the article.

A Review of Prior Art Citations and Citation Noise

Citation-based patent evaluation methodologies enable the assessment of patent metadata to reveal hidden details and relationships that can provide strategic insight. These methodologies focus on the raw data – prior art citations – that are used for different types of measures or proxies that generally relate to properties concerning knowledge, technology, or invention.

Prior art citations identify technology that was known, or the "state of the art", prior to the invention of the patented technology. These citations may create a legal link and relevant relationship between the prior art and the patented technology. When a prior art citation refers to an existing, older patent, this technology coupling or link is made explicit with the patent through what is called a "cited citation". However, the older patent also inherits a citation forward in time to the new patent, and this is called a "citing citation". Thus, a given patent may have cited citations (i.e., referring backwards in time to prior art) and the patent may also be listed as a citing citation (i.e., referring forwards in time to newer patents). Together, both types of links can create a network of insightful interrelationships between citations and patents, but only if they can be discovered from the background noise commonly found in citation data.

The objective of this literature review is to examine the current state of knowledge in prior art citations as they relate to citation-based patent evaluation methodologies. The relevant literature was located using a broad keyword search of scholarly journals in the Business Source Complete database (tinyurl.com/22teqy) with respect to prior art citations. The abstracts and introductions of the articles were examined closely with a focus on prior art patent citations. This step resulted in a list of 11 relevant articles and a book related to the topic of prior art citations. Ten of the articles are based on empirical research; one article is a literature review. Many of the authors identified a long-standing problem with prior art patent citations: citation noise.
Finding the Signal in the Noise of Patent Citations: How to Focus on Relevance

Derek Smith

Prior art citations
A patent evaluation methodology can use prior art citations in a number of different ways. For example, they can be used to track a flow or exchange of knowledge between the patent and the prior art citations (Bacchiocchi & Montobbio, 2010; Criscuolo & Verspagen, 2008; Hu & Jaffe, 2003), especially if they are cited citations (Hu et al., 2012). When the flow of knowledge is traced with secondary data associated with each prior art citation, the patent can be used to evaluate other aspects such as firms, actors, regions, and countries, and it can be used to identify technology segments associated with the knowledge (Bacchiocchi & Montobbio, 2010; Hu & Jaffe, 2003). Criscuolo and Verspagen (2008) found that European patents typically have fewer prior art citations than patents from the United States. They also found that prior art citations provided by an patent examiner tend to anticipate the claims having a strong relationship between the patent and prior art citations whereas prior art citations provided by the inventor tend to have a much weaker relationship.

Prior art citations can be used to evaluate the technical value or significance of an invention and indicate the origins of the flow or exchange of knowledge. Prior art citations also reveal a network of technological communities based on the links in a citation network. Cited citations reflect codified cumulative knowledge. Secondary data with citations can reveal business intelligence concerning actors, geographic proximity, and the exchange of knowledge between actors (Gay & Le Bas, 2005).

Gay and colleagues (2005) found that prior art citations can be used to evaluate the technical value of an invention based on a citation rate, especially if they are citing citations (Hu et al., 2012). Trajanberg (1990) found that prior art citations can be an indicator to the importance or value of a patent. The frequency of a citation in technological fields also suggests value (Gay et al., 2005; Jaffee & Trajanberg, 2002). A long-standing cited citation that is frequently cited in other technological fields suggests a broader, more valuable patent, or a core technology required by many other technologies. Similarly, a high number of citing citations can be used to identify a high potential or breakthrough invention, in part because breakthrough inventions are built upon past knowledge that would be reflected in the number of cited citations (Kelley et al., 2013). A higher number of citations indicates a higher value for appropriating an invention (Nikulainen et al., 2008).

Van Zeebroeck (2011) found that the network associated with prior art citations illustrates linkages between upstream and downstream technology knowledge that may be evaluated. The upstream prior art citations suggest a more valuable invention and patent.

Trajanberg and colleagues (1997) found that cited prior art citations relate to the basicness of an invention and citing prior art citations relate to the appropriability of the invention. A patent with a higher count of citing citations is relatively more important than a patent with a lower count of citing citations.

The prior art citation literature may be further synthesized into three groups of properties that relate to citation-based patent evaluations, as shown in Table 1. The first property relates to knowledge surrounding the invention, where the prior art citations are generally applied to reveal the relationship with knowledge over a period of time between the date of the patent and the dates of the prior art citations. The knowledge properties can identify firms and actors associated with a particular patented technology and are beneficial when seeking to identify strategic relationships around a patented technology.

The second property relates to the technology, where the prior art citations can be applied to evaluate the technology described and claimed in the corresponding patent. The technology property can identify technological value of a patent and be beneficial when seeking, or offering, venture capital based on a particular patent or group of patents. This property works both ways and should be made relevant to focus on real technological value.

The third property relates to the invention, where the prior art citations are applied to evaluate the degree of originality. The invention property can identify invention value disposed within a patent and can be beneficial when managing a patent portfolio or considering expenses. This property can distinguish inventions with a high degree of originality from inventions with a lower degree of originality.

Citation noise
The literature review revealed citation noise as a long-standing problem. It remains as a missing piece in our knowledge and understanding of patent evaluations based upon prior art citations. Citation noise is a problem that appears in the literature as early as 1990 (Trajanberg, 1990) and continues to be a known problem referenced by many authors at least up to 2013 with a paper by Kelly, Ali, and Zahra (2013).
Finding the Signal in the Noise of Patent Citations: How to Focus on Relevance

*Derek Smith*

**Table 1.** The applicability of prior art citations to patent evaluations

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Applicability of Prior Art Citations</th>
</tr>
</thead>
</table>
| **Knowledge Properties**            | • Can be a proxy to track knowledge flows between patents and measure the intensity of knowledge flows. Secondary patent data provides an indication of regions, countries, and specific technologies (Bacchiocchi & Montobbio, 2010).  
  • Are a proxy for knowledge flow and create a paper trail of the flow of knowledge (Criscuolo & Verspagen, 2008).  
  • Provide the linkages between upstream and downstream technology knowledge. Downstream, they reveal previous research activity; upstream, they suggest value of inventions (van Zeelbroeck, 2011).  
  • Can be an indicator to the importance of value of a patent. A frequency or count of citations is applied to calculate a weighted count of patents. Citations that cross technology areas are indicative of knowledge spillover (Trajtenberg, 1990). |
| **Technology Properties**           | • May be a proxy to assess the technical value of inventions, indicate the origins and flow of codified knowledge, or lead to a measure of technical significance. Patent citations can reveal the network of technological communities. Cited citations contain codified knowledge and are a proxy for cumulative knowledge. Cited and citing citations are a proxy for knowledge flow. Secondary information reveals business intelligence concerning the actors associated with the patent citation, geographic proximity, and knowledge exchange between actors. Citing patents represent knowledge flows of new ideas and knowledge (Gay & Le Bas, 2005).  
  • Can be a proxy to the technological value of inventions based on the citation rate or frequency of citation in other technological fields (Gay et al., 2005)  
  • Are a proxy for knowledge flows between patents. With secondary patent data such as the inventor name and country, can be a proxy for knowledge flows between countries (Hu & Jaffe, 2003).  
  • Can be a proxy to technological value when citing and a proxy to knowledge flow when cited against the target patent. Cited patents represent knowledge flows of old or known knowledge (Hu et al., 2005). |
| **Invention properties**            | • Make it possible to trace the links between patents and citations and the flow of technological knowledge between the citations and patent. Citations may be an indicator of invention importance for the degree of generality or originality. Citations illustrate how technological knowledge combines to produce new knowledge over time. A highly cited patent suggests importance. Citations may also be used as a proxy for technological impact and, when used to weight a count of patents, can lead to more meaningful information (Jaffee & Trajtenberg, 2002).  
  • Can be a proxy for high potential or breakthrough inventions, based on the highest number of citing citations. Cited citations can be a proxy for intermediate value of a technology (Kelley et al., 2013).  
  • Cited citations are a proxy for the present economic value of a company, and they position the patent with respect to the basicness and originality of the patent. A relatively low count suggests low basicness and technological significance. Citing citations are a proxy for technological significance, indicating a potential for higher economic return. More citing citations suggest higher value for appropriating and applying the invention (Nikulainen et al., 2008).  
  • Aspects of basicness and appropriability exist in the relationships between cited and citing patent citations. Basic relates to each innovation building upon a body of existing knowledge before it and appropriability relates to benefiting from the earlier patented innovation. Citations can therefore identify and trace the citation relationships. A patent is important if it is highly citing, and a patent is basic if it is highly cited (Trajtenberg et al., 1997). |
Finding the Signal in the Noise of Patent Citations: How to Focus on Relevance
Derek Smith

Citation noise occurs as a natural result of timing and different perspectives and comprehension during the process to obtain a patent. Some jurisdictions have a higher degree of noise in the citation data when compared to other jurisdictions (Criscuolo & Verspagen, 2008; van Zeebroeck, 2011; Gay & Le Bas, 2005). Bacchiocchi and Montobbio (2010) found a home bias effect, where local citations are more likely to be included than citations from other jurisdictions. A home bias effect depends on different legal and patent examination procedures that generate the representative patent citations. They also found that patent citations are localized at the European Patent Office, with fewer citations per patent than observed at the United States Patent & Trademark Office. This occurs because of differences between the jurisdictions. The United States Patent & Trademark Office places a duty to disclose prior art references upon each applicant in contrast to the European Patent Office, which has more such duties to disclose. The duty to disclose prior art and the consequences for not disclosing prior art tend to drive up the citation count and corresponding citation noise.

Patents may contain some amount of noise in the associated citation data (Criscuolo & Verspagen, 2008; Gay & Le Bas, 2005; Gay et al., 2005; Hu & Jaffe, 2003; Hu et al., 2012; Jaffee & Trajtenberg, 2002; van Zeebroeck, 2011) or the citation data may not capture any flow of knowledge between the patent and prior art citation (Hu et al., 2012). The amount of noise in the citation data is a factor of the total number of prior art references supplied by interested actors; the number of differing perspectives, miscomprehensions between the technology described in the prior art references and the technology covered by the patent claims; and the timing in the process. Raw or simple counts of citations may underestimate the actual number of citations (van Zeebroeck, 2011) and may not be suitable across different technology fields (Hu et al., 2012).

Citation noise therefore presents a major challenge to the effectiveness of patent evaluation methodologies, which may therefore be poor indicators of the economic value of patents (Gay & Le Bas, 2005). When the noise is caused by citation lag and citation variability, an evaluation concerning technology value may result in a diversity of technology value (Gay and colleagues, 2005). A timing issue causes this type of noise when a patent is evaluated with respect to the number of citing (future) citations. Citation noise may further impact evaluations relating to knowledge flow if many citations are not associated with knowledge flow (Hu & Jaffe, 2003). Knowledge flow occurs when a patent is building upon or improving older technology and may be identified when there is a direct link between the claims of the patent and relevant citations.

Citation-based patent evaluation methodologies
Based on the literature review, the patent evaluation methodologies that use prior art citations are summarized and illustrated in Figure 1. Depending on the business strategy of the evaluator, a patent and its associated noisy prior art citation data may be evaluated through a lens of three different properties – knowledge, technology, and invention – that can yield three distinct potential outcomes. However, any potential evaluation using current methodologies is impeded and limited by the citation noise within the citation data and propagates through to each of the three potential outcomes. Citation noise results in an obscured or unreliable potential outcome.

**Figure 1.** Three views of citation-based patent evaluations and potential outcomes influenced by citation noise

www.timreview.ca
Finding the Signal in the Noise of Patent Citations: How to Focus on Relevance

Derek Smith

Interdependency between prior art citations and patents
Citation noise was identified as a common problem in the literature as it relates to distinguishing relevant prior art citations from the prior art citations that are not relevant to the patent. Therefore, a second stream of literature consisting of six articles was located using a broad keyword search of scholarly journals in the Business Source Complete database with respect to interdependency between prior art citations and patents. Interdependency is a legal relationship that may or may not form between a prior art citation and the patent as part of the patent examination process.

The mere presence or listing of a prior art citation in citation data does not provide any indication of the relevance or lack of relevance of the prior art citation (Abbott Laboratories Inc. v. Baxter Pharmaceutical Products Inc., 2003). The patent examination process (i.e., prosecution) may result in a formal rejection based on a prior art citation and an amendment to the patent claim (Berger et al., 2011). Patent applications are examined using the prior art citations. A prior art citation, or combination of prior art citations, is relevant when an examiner accepts an amendment to overcome a rejection. Amendment may or may not occur and the application may be granted or be rejected based on the legal arguments or amendments to the pending claims (Kica & Groenendijk, 2011). The author of this article further found that the interdependency between a patent and prior art citation could discover relevant citations and eliminate the noise in citation data (Smith, 2014b).

In the United States, a formal rejection may be based on novelty or obviousness; in Europe, a formal rejection may be based on novelty or inventiveness (Laub, 2006). Other jurisdictions are similar. The prosecution process and the formal rejections are documented in the form of a patent prosecution history file. This file is available from a patent office and contains documents that codify formal rejections and amendments accepted by an examiner (Cotropia et al., 2013).

A Framework for Relevant Citation-Based Patent Evaluations
A framework for relevant citation-based patent evaluations is proposed from the ideas identified from the first literature stream concerning prior art citations and citation noise in combination with the ideas from the second literature stream concerning patent and citation interdependency. As shown in Figure 2, the framework synthesizes these ideas while incorporating the associated linkages between the three properties, as viewed through a citation-data lens. The linkages first relate to a flow of noisy citation data. The noisy citation data is passed through the citation relevancy filter to discover relevant citations from the noise. Relevant citations are then evaluated with respect to knowledge properties, technology properties, or invention properties. This framework results in a clear potential outcome based on the evaluated property.

The framework presented in Figure 2 builds on Figure 1 with the addition of a citation-relevancy filter. When applied to a list of prior art citations included in the raw citation data, this filter identifies the citation as relevant if it has been cited by an examiner and resulted in an amendment to the patent claim or prevented grant of the patent. Thus, the filter is an evaluation tool for reviewing the patent prosecution history file and associated documents concerning the patent to be evaluated. Without the filter, the potential outcomes are obscured by the citation noise (Figure 1); with the filter, the path is clear and the likelihood of a successful, more reliable outcome is greater (Figure 2).

![Figure 2. A framework for relevant citation-based patent evaluations](image-url)

www.timreview.ca
Finding the Signal in the Noise of Patent Citations: How to Focus on Relevance

Derek Smith

In the subsections below, three example scenarios illustrate potential applications of the framework from Figure 2 for relevant citation-based patent evaluations.

**Scenario 1: Knowledge properties**
Consider a patent holder that has limited resources and skills to create technology based on their patent. One option is to identify other firms or actors to possibly collaborate with to create the patented technology.

In this scenario, a patent evaluation with a focus on the aspect of knowledge properties is beneficial to identify other firms or actors with similar or advanced skills relating to the patented technology. Such an evaluation could use the framework identified here to discover relevant citations from the associated patent citations, as follows:

1. List each prior art citation from the raw citation data associated with the patent.
2. Access the patent office website that granted the patent and locate the patent prosecution history file.
3. Scan the documents disposed with the prosecution history file to find: i) each prior art reference that was cited by an examiner and ii) any amendments to the patent claims based on each cited reference.
4. Finally, note whether or not the examiner accepted the amendment to the patent claims.

A relevant citation is identified for each prior art citation that was i) listed as a prior art reference to the patent and ii) was cited as a prior art reference by the examiner, and for which an amendment to the patent claims was entered and accepted by the examiner.

Relevant citations for this scenario relate to a flow or exchange of knowledge between each relevant prior art citation and the patent, as well as the linkages to the citations. For each relevant prior art citation identified, the evaluator should locate the document and the secondary information associated with the document to reveal at least the firm and actor associated with the document. This step identifies strategic firms and actors that have experience with the same or similar technology to that of the patented technology.

**Scenario 2: Technology properties**
Consider a firm that has interest in accessing the technological value from the perspective of the importance of the technology to other technologies as a required building block. An example is the relationship of a new battery technology and the importance to many other portable electronic devices that require this battery technology.

In this scenario, an evaluation might occur before seeking venture capital based on the patent. Technological value can be evaluated based on the number of relevant citing citations: the higher the number of relevant citing citations, the higher the relative technological value of the patent.

The same process described in Scenario 1 can be used to discover relevant citations. But, here, the total number of relevant citing citations associated with the patent should be counted. The patent may then be compared to other similar patents to evaluate the relative value. If there are no relevant citing citations, the patent may not be valuable from a technology perspective or is still in an immature state.

**Scenario 3: Invention properties**
Consider a firm that wishes to assess the originality (or basicness) of its patent. In this case, it is the number of cited citations that is important: the higher the number of relevant cited citations, the more basic the invention. The purpose of this evaluation could be to rate and rank multiple inventions within a patent portfolio.

Again, the process begins with discovering the relevant citations as described in Scenario 1. However, in this scenario, the evaluator should count the total number of relevant cited citations associated with the patent. The patent may then be compared to other similar patents to evaluate the degree of originality or basicness. If there are no cited citations, the invention is not basic, it is an invention with a high degree of originality.

**Conclusion**
Existing citation-based patent evaluation methodologies are limited due to citation noise. A patent evaluation based on relevant citations eliminates citation noise and identifies relevant hidden relationships and insight in support of identifying potential strategic relationships and in conducting technology and invention assessments. A patent evaluation based on relevant citations can also provide strategic and competitive advantage to a firm, especially when aligned with the present and future needs of the business.

Future research should focus on testing the framework for relevant citation-based patent evaluations. Other
Finding the Signal in the Noise of Patent Citations: How to Focus on Relevance

Derek Smith

citation-based patent evaluation methodologies that include patent and business information could also be examined to build upon the framework.

In practice, executives and managers should proactively conduct a relevant citation-based patent evaluation as part of a patent portfolio management activity and routinely assess strategic opportunities.

About the Author

Derek Smith is the founder and principal of Magneto Innovation Management, an intellectual property consulting firm that assists entrepreneurs and small businesses with difficult intellectual property issues. He is a registered patent agent in both Canada and the United States, and he has over 20 years of experience working as an intellectual property management consultant and patent agent for IBM Canada, Bell Canada, and Husky Injection Molding Systems where he was Director of Global Intellectual Property. Prior to entering the field of intellectual property, he was an advisory engineer at IBM Canada where he was involved in a variety of leading-edge software development projects. Derek holds an MASc degree in Technology Innovation Management from Carleton University in Ottawa, Canada, for which he was awarded a Senate Medal for Outstanding Academic Achievement. His thesis is "A New Methodology For Citation Dependent Patent Evaluations". Derek also holds a BEng degree in Systems and Computer Engineering, also from Carleton University.

References


http://dx.doi.org/10.1111/j.1467-9442.2010.01614.x

http://dx.doi.org/10.1016/j.respol.2011.07.004


http://dx.doi.org/10.1016/j.respol.2013.01.003

http://dx.doi.org/10.1007/s10880-003-00730

http://dx.doi.org/10.1007/s10880-003-007329

http://dx.doi.org/10.1016/S0167-7187(03)00035-3

http://dx.doi.org/10.1002/asi.22632


http://dx.doi.org/10.1257/aer.90.2.215

http://dx.doi.org/10.1111/jpim.12055

http://dx.doi.org/10.1080/13511610.2011.571405
Finding the Signal in the Noise of Patent Citations: How to Focus on Relevance

Derek Smith


Keywords: intellectual property, patent evaluation, patent evaluation limitations, prior art citations, citation-based patent evaluation, citation noise, literature review