

A Framework for Analyzing the Underlying Inventions That Drive Technical Improvements in a Specific Technological Field

Christopher L. Benson¹ & Christopher L. Magee²

¹ Department of Mechanical Engineering, Massachusetts Institute of Technology, MA, USA

² Engineering Systems Division and Department of Mechanical Engineering, Massachusetts Institute of Technology, MA, USA

Correspondence: Christopher L. Benson, Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. Tel: 1-720-628-6763. E-mail: cbenson@mit.edu

Received: March 27, 2012

Accepted: April 27, 2012

Published: May 1, 2012

doi:10.5539/emr.v1n1p2

URL: <http://dx.doi.org/10.5539/emr.v1n1p2>

The research is supported by the SUTD/MIT International Design Center.

Abstract

This paper presents a repeatable method to identify, describe and analyze the important inventions in an industry. The method involves qualitative and quantitative methods of analysis of patents to understand the inventions that have contributed to the growth of an industry. We present a process for mapping the inventions by their importance in an attempt to understand the set of inventions that contribute most significantly to the development of an industry. As a demonstration of the method, a set of 2484 patented inventions in the solar PV industry was selected as a case study. The case study results show that the conceptual basis for the method is useful in understanding technological change in an industry. For example, the analysis identifies important PV patents but none that are discontinuously important. These findings support the probable utility of the framework and have some preliminary broader implications to our understanding of technological development.

Keywords: technological planning, technological breakthroughs, importance of individual patents

1. Introduction and Importance of Technological Innovation

The development of new technologies has enabled mankind to grow into the civilization that exists today. Industries of all kinds have created unprecedented new capabilities that would not have been possible without the advent of new technological understanding and inventions. Attempting to understand how the future is created is a daunting challenge; our approach for doing so is to examine the past in order to understand how we got here today. Many people believe that radical or breakthrough inventions lie at the center of wealth creation (Ahuja and Lampert, 2001). This paper presents a new method for the analysis of technological change in an industry and allows for some examination of breakthrough inventions.

2. Literature Review

The field of technological change has gained attention in recent years due to the more apparent results of what happens when a large organization does not innovate effectively and becomes overtaken by newer organizations with new technologies and practices. New firms replacing old ones is not new, but has been accelerated in recent years and thus made much more apparent, especially in rapidly developing fields such as information technology. Clayton Christensen (1997) has analyzed aspects of this field with his work on disruptive technologies and innovations. While his work is one of the most familiar to general audiences, there are many other theories and ideas around how technologies are developed including Utterback's (1996) theory of dominant designs and the theory by Henderson and Clark (1990) that describes multiple modes of innovation including incremental, architectural and radical innovation. Existing methods of analysis can be usefully supplemented by objective analysis of the actual technological change in an industry.

2.1 Methods of Measuring Technological Change

One of the ways that people have attempted to understand technological development is to list a set of developments that are considered innovations. Table 1 shows an example of this, with a list of important innovations throughout the last 300 years as given by Girifalco (1991).

Table 1. Example list of innovations throughout history (Girifalco, 1991)

Innovation	Year
Seed Drill	1731
Watt Engine	1776
Rifle	1824
DC Motor	1872
Airplane	1903
Methyl Methacrylate	1935
Transistor	1947
Personal Computer	1974

From examining this list, it is clear that there are issues in deciding what constitutes an important invention. These issues stem partially from ambiguity about the *level* of innovation to be considered. For example, the transistor is shown in the list and is widely considered a breakthrough technology, at the same time so is the integrated circuit (an important way of utilizing and manufacturing transistors) and so has the personal computer (also shown in the list and which depends directly on integrated circuits). It is also reasonable to consider the entire field of information technology (which relies on personal computers and many other technologies) the most important technological breakthrough of the latter half of the 20th century.

The ambiguity is further compounded by the fact that the described technological improvements can be collectively combined over varying time periods. In fact, this is often done to simplify communication about developments within a field. For example, the initial invention of the transistor was completed in a much shorter time than all of the ensuing and continuing changes in transistors. The same is true relative to the initial invention of the integrated circuit and the modern computer, which is only one aspect of the development of information technology.

It is interesting to note that many design changes designated as significant improvements or breakthroughs appear to identify individual points of improvement. Thus, the list in Table 1 would integrate well with the widely accepted conjecture that technological breakthroughs drive the overall improvement of a technology. The importance of breakthroughs is supported by Kaplan (1999), who states that

'substantial growth over the long horizon requires discontinuous innovation' (Kaplan, 1999).

A similar point of view is given by Ahuja and Lampert (2001) as they state that breakthrough inventions

'serve as the basis of new technological trajectories and paradigms and are an important part of the process of creative destruction in which extant techniques and approaches are replaced by new technologies and products'.

The idea that a small set of technological changes account for a large portion of the overall improvement has been echoed many times in regards to different aspects of technology.

With the apparently widespread acceptance of these related notions, it is important to be able to define a repeatable method to determine technologies that can be considered as breakthroughs. Currently, the definitions found in the literature are difficult to use when broadly analyzing breakthrough technologies across time and technical fields. Some of the definitions that are provided tend to be most useful when considering technologies post-hoc, such as the definition provided by Sahal (1981)

'Major technological innovations represent technical advance so significant that no increase in scale, efficiency, or design can make older technologies competitive with the new technology'

Definitions like this are difficult to use on a repeatable basis due to the reliance on counterfactual analysis, post-hoc data, and particularly the size of the technical advances brought by a technology and on the lack of a means to assess the size of that change. In addition, the term "technology" in these definitions is unclear as to how much hierarchical (and time) consolidation is inherently assumed. The terms "breakthrough" and "radical" often imply singular inventions but no care is exercised in clarifying this in the definitions. Other definitions tend to be intentionally more broad, such as the one due to Mascitelli (2000)

'Discontinuous innovation, for example, typically refers to products that involve significant new technologies and are aimed at a market that is unfamiliar with the product class'

Terms such as 'significant new technology' cannot easily be applied to diverse situations objectively. This overall lack of clarity about abstractions over time and level makes it very difficult to unambiguously describe the technical changes or inventions that contribute to the progress in a technical field, and thus to understand at an

actionable level how to utilize the concept of breakthroughs. It is therefore not surprising that lists such as these have been criticized as potential sources of data due to their subjectively selected nature, lack of repeatability and reliance on a priori assumptions (Yu and Hang, 2010; Wang et al, 2010).

2.2 Using patents as a proxy for inventions

A direct way to eliminate subjectivity over time is to utilize patents as the observed unit of development (Abraham and Moitra, 2001). The patent database covers most fields of invention and it has information on nearly every invention in the last 200 years (Markides, 2006). Furthermore, the patent databases are accessible and include important metadata such as citations that can be used to gain more information about the patents and how they relate to technical progress as a whole.

Patent citation analysis has been used for assessing the importance of individual patents by Manuel Trajtenberg (1990), and has been used to validate the data sets of other studies in the field of technical innovation (Dahlin et al, 2004). Trajtenberg showed that

'patent citations may be indicative of the value of innovations and, if so, that they may hold the key to unlock the wealth of information contained in patent data'

Much of the analysis in this paper is built on the ideas of Trajtenberg et al. (1997) that the importance of patents as well as other useful measures can be determined by analyzing the citation data. We are not aware of any previous study that utilizes patent citation analysis to identify and analyze the nature of the most important patents in a given technological field but it appears to be a viable way to build a framework for such a purpose.

3. Framework for Analyzing Technological Change

This section will explore a repeatable methodology that is both quantitative and qualitative in its approach, and lends itself to both a deep and broad understanding of many of the most important inventions in a field. A graphical representation of this method can be found in Figure 1, the specifics of each of the steps shown in the figure will be discussed throughout the remainder of this section.

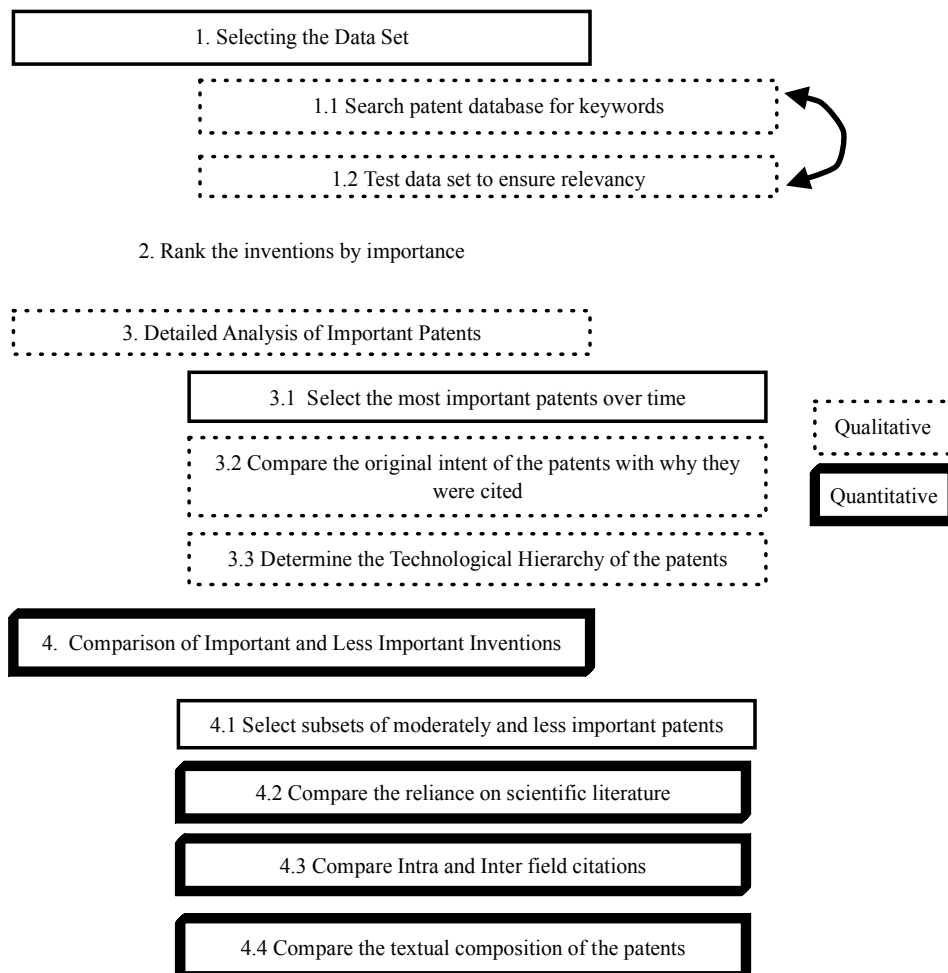


Figure 1. Summary of qualitative and quantitative framework

3.1 Selecting the Data Set

The first step (step 1 in Figure 1) in the methodology is to populate a data set that is focused narrowly on the field of interest, but broad enough to capture as many inventions related to the field as possible. In addition to the patents that are directly relevant to the growth of a technical field, we will be interested in understanding “spillover” of developments from other fields to the field of interest (Hur and Watanabe, 2001; Kwang and Chihiro, 2001; Griliches, 1992). Thus, we must specify the primary field of interest and differentiate it from others.

In order to do this, appropriate keywords must be used for the initial patent search, (step 1.1) so as to provide as complete as possible data set from which to analyze the field. Several combinations of key words in the field of interest were explored and were used as search terms within the title, abstract and claims in the US Patent database of issued US patents. This allows a comparison of the size and relevancy of the returned sets. (Note 1)

Step 1.2 in Figure 1 involves testing the returned data sets for relevancy to the field. This can be done by multiplying the total number of patents in the returned data set (N) by the percentage of relevant patents (P) in a returned set, which will result in the total number of relevant patents (R).

$$P \times N = R$$

One notes that the larger the data set (N) is, the more likely it is to have a higher absolute number of total relevant patents (R), up until the limit of searching the entire U.S. Patent Database, which would include all of the relevant patents, but only a very small percentage in the set would be relevant. It is desirable to select a data set with a relatively high percentage of relevant patents (P), as this will make the rest of the process more repeatable and accurate. The goal in this process is to maximize R, while keeping P as high as possible.

In order to find the percentage of relevant patents (P), the patents in the data sets must be sampled and read to determine if they are relevant. The relevance of each patent set can be determined by reading the title and abstract of three sets of 50 patents of varying forward citation count in each of the possible data sets and recording the percentage of relevant patents.

Once search terms and thus a final data set has been chosen, sample sets of 150 patents of varying forward citation count are chosen to ensure that the final data set is relevant to the specific technical field. Steps 1.1 and 1.2 combined are an iterative process, and must be repeated until a data set sufficiently representative of the field of interest has been found.

3.2 Ranking the Inventions by Importance

Once the data set is selected, one can map the inventions in a technical field over time utilizing citation analysis (step 2 in Figure 1).

As mentioned in section 2.2, the forward citation count can be used to represent the importance of a patent. Thus, one can rank the patents by their forward citation count as reflecting their relative importance. It is then possible to graph the rank order of importance of the patents against their overall forward citation count in order to see the distribution of importance of inventions throughout the time period. This graph can be used as a map of important inventions throughout the entire history of the specified field.

3.3 Analyzing the Most Important Inventions over Time

In step 3.1, one selects the three most cited patents in the field of interest from each decade to represent the most important patents in the field. Using this sampling method, it is possible to understand the most important inventions throughout the history of a technical field, which helps mitigate the issues of increasing patenting rates and citation truncation (Note 2). This sub-set of ‘Very Important’ patents is analyzed qualitatively to understand how they impacted the overall technical field.

Step 3.2 determines the specific ways these highly cited patents improved on the previous state of the art. In section 3.2a, the prior art section of each patent is examined for statements about shortfalls of existing technological approaches and for specific statements about what is superior in the method described in the patent over technologies that existed at the time of issue. This step provides an understanding of what the inventor(s) viewed as the major contributions of the invention. Next, insight into the technical improvement provided by a patent as recorded by future patents is determined by analysis of the in-text citations from patents that cite these important patents. This summary of the ex-post interpretation of the technical improvement of the patent gives us a more clear understanding of how future inventors built upon the ideas from this invention. The ex-post summary can be compared with the ex-ante summary of the technical improvement to shed light on how the intentions of the inventor compare with the future impact of the invention. The method is built upon the concept that citations of a

patent reflect its importance to the field but if the citations are for reasons not related to the inventor's claims, this assumption might be questioned.

Step 3.2b uses the ex-ante and ex-post summaries and more detailed study of the patents to determine the technological hierarchy levels of the 'very important' patents. A method of characterizing the technological hierarchy of an invention was described by Magee (2010) that places inventions into one of 6 technological levels in a hierarchy ranging from *Materials and Process Improvement* which is the lowest level in the hierarchy through *Component Redesign, etc.* up to *System Phenomenon Change* where the actual scientific phenomenon in the technology is affected. Table 2 summarizes the technical hierarchy states that have been adapted from Magee (2010) for use in this methodology. Such a hierarchy analysis has not previously been performed from patent information so the test case is needed to see if it can be done.

Table 2. Levels of technological hierarchy along with example technological changes for information transmission adapted from Magee (2010)

Hierarchy	Example in information transmission	Level
Materials/Process Improvement	Coatings on glass fibers	1
Materials/Process Substitution	Glass fibers replacing metallic conductors	2
Component Redesign	Optical 'solitons'	3
System Redesign	Optical amplification	4
System Operation Change	TCP/IP; wavelength division multiplexing	5
System Phenomenon Change	Wireless vs. wired transmission	6

3.4 Performing Comparative Analysis of Important and Less Important Patents

Step 4 in Figure 1 provides a broader comparison of the important patents to other sub-sets of patents in order to examine how important inventions differ from the rest. In a similar way that the sub-set of 'Very Important' patents was selected, a sub-set of 'Moderately Important' and 'Less Important' patents is created in Step 4.1 by selecting patents with the median citation count and a citation count of one, respectively. Unlike the data set as a whole, all non-PV patents were removed from these three subsets to assure relevance of the analysis to the solar PV field.

Step 4.2 is a comparison of the reliance on scientific research by the very important, moderately important and less-important patents. Patents often cite non patent literature (NPL), which in almost all cases are basic scientific literature. The NPL citation rate is indicative of the importance on new science in the field and also compared across the three sub-sets of patents to understand their reliance on scientific literature.

An analysis possible from the *backward* citations of patents is to identify the patent categories that the citing patents built upon for *their* inventions. This is possible because each patent is required to have a category, which can be used as a proxy for a technical field. Examples of these categories range from 'Structures' to 'Batteries, Thermoelectric and Photoelectric'. In step 4.3, the primary patent classes of the citations of the entire patent set are analyzed in order to understand which patent classes are built upon by the inventions in the field of interest. The top 10 most cited fields from the overall patent set are determined and show what specific areas the field of interest builds upon. A comparison can then be made between the most cited classes of the citations of highly cited patents and the overall patent set. In order to achieve a more accurate representation of the cited fields, a larger sample size of the 250 most highly cited patents is recommended to represent the important patents in this comparison.

In Step 4.4, the text of the patents is parsed and analyzed in an attempt to better understand the make-up of the very important patents as compared to the moderate and less important counterparts. Analysis of the length of abstract (in number of characters), the length of the prior art section (in number of character) and the number of embodiments of invention mentioned in the patent are performed.

4. Results of Case Study

This Section will demonstrate the use of the framework for understanding the development of a technical field using the field of Solar PV as a case study. This allows us to test the assumptions underneath the framework and to assess the usefulness of the framework. Solar PV is a very acceptable case study for the research because it

has developed substantially over time, the time period in which it has developed is an acceptable length, and because the field relies on patents to protect new inventions. As is shown in Figure 2, the development of the field of solar PV has been substantial, as the capability to produce useful energy grew from 0.014W_p/\$ in 1976 to 0.358W_p/\$ in 2009 (EIA, 2010; Harmon, 2000). Following the framework given in Figure 1 we will examine technological change in the solar PV field over the past half-century.

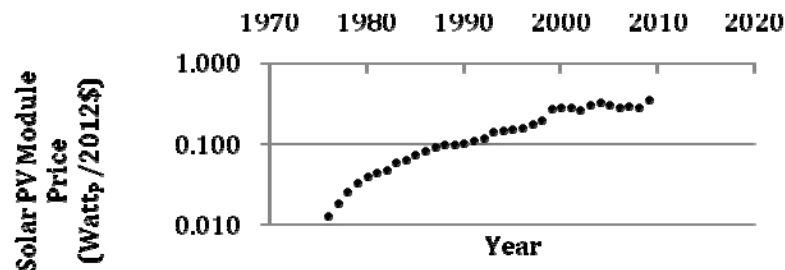


Figure 2. The exponential growth of kWhr/\$ for solar PV from 1977-2010 is demonstrated by the logarithmic plot.

4.1 Selecting the Solar PV Data Set

Step 1 is to create the data set from which to analyze the contributions of inventions in the solar PV field. In order to do this, we explored several combinations of key words in the solar PV field (solar, cell, module, silicon, photovoltaic, electricity) as search terms in the title, abstract and claims in the US Patent database of issued US patents and compared the size and relevancy of the returned sets.

A search of the term 'Solar Electricity' returned almost 7000 results. The determination of relevance to the solar PV field was done by reading the patents and discounting patents such as 'Solar powered smart card with integrated display and keypad' (5777903) and 'solid state electronic camera including thin film matrix of photosensors' (4788594) which are not directly related to the PV field.

A sampling of six selections of 50 patents from the search results of 'Solar Electricity' contained only 38% patents directly relevant to the solar PV field, and the sampling included many entirely irrelevant patents such as PN 5892900 'Systems and methods for secure transaction management and electronic rights protection.' The low occurrence of solar PV patents as well as the significant presence in this set of completely unrelated patents make the patents from the search for 'Solar Electricity' not appropriate for the analysis of the solar PV field.

A more specific search for the term 'Photovoltaic Electricity' between 1970 and 2010 returned a data set of 2484 patents. Five selections of 150 patents across the data set (750 in total) were analyzed and indicated that 62% of the patents in the set were related to the solar PV field, which was judged to provide an acceptable representation of the field of interest for this case study. The final data set was not cleaned of the discovered non-PV patents to allow for repeatability since not all 2484 patents in the set were searched in detail.

4.2 Ranking the Inventions by Importance

Figure 3 shows the distribution of the most cited to the least cited patents in the data set of 2484 solar PV patents as is prescribed in Step 2 of the framework. There is clear evidence of patents that are cited much more highly than others, there are 57 patents that are cited at least 50 times each, inventions that can be considered important to the technological development of solar PV. There are also many patents that are moderately important, with over 600 patents being cited between 10 and 20 times. Finally, there is a clear tail with 522 patents being cited zero times, an indicator of less important patents. Figure 3 shows that despite the enormous difference between the most highly cited and the least cited patents, there is no discontinuity in citation rate within the entire set of patents –the distribution is relatively continuous.

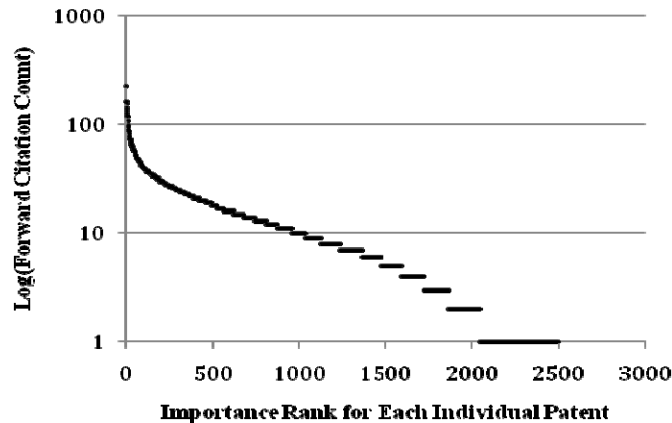


Figure 3. A plot of the citations received by each patent against the rank of the patent (the patent with the most citations is shown at ranking position 1)

4.3 Detailed Analysis of Important Patents

Step 3.1 in the framework calls for a selection of important patents over time in the field of interest. For the solar PV case study, we selected 3 patents from each of the 4 decades in the lifetime of the solar PV field (1971-81, 82-91, 92-01, 02-11) to results in a selection of 12 ‘Very important’ patents that are shown in Table 3.

Table 3. Summary of the results from Step 3 of the framework for the 12 ‘Very Important’ Solar PV patents

Patent Number - Patent Title (year)	Frequency #	Improvement on Prior Technologies from Prior Art section	In Text/ Total Citat	Summary of in-text citations
A 6441298 - Surface-plasmon enhanced photovoltaic device (2002)	3	Current devices are not stretchable and do not conform to the shape of surfaces used in many electrical devices.	0/56	No In-Text Citations
B 6878871 - Nanostructure and nanocomposite based compositions and photovoltaic devices (2005)	2	Current photovoltaic devices or cells employ thin layers of semiconductor material, e.g., crystalline silicon, gallium arsenide, or the like, incorporating a p-n junction to convert solar energy to direct current - their efficiency has been somewhat limited.	7/57	Photovoltaic devices having thin layer structures that include inorganic nanostructures
C 6340788 - Multijunction photovoltaic cells and panels using a silicon or silicon-germanium active substrate cell for space and terrestrial applications (2002)	1	For the multiple-cell PV device, efficiency is limited by the requirement of low resistance interfaces between the individual cells	2/42	Multiple graded buffer layers placed between upper and lower subcells to provide maximum flexibility in subcell lattice constant and band gap
D 5747967 - Apparatus and method for maximizing power delivered by a photovoltaic array (1998)	4	There is a need for a peak power tracker for a photovoltaic array that operates the photovoltaic array at peak power output for significant periods of time	1/67	No In-Text Citations
E 5350644 - Photovoltaic cells (1994)	1	It is important that the titanium dioxide film be coated with a photosensitizer which harvests light in the wavelength domain where the sun emits light.	16/64	Graetzel et al reported solid-state dye-sensitized mesoporous TiO ₂ solar cells with up to 33% photon to electron conversion efficiencies.
F 5092939 - Photovoltaic roof and method of making same (1992)	3	What is needed is an economical system for constructing a roof structure that uses the flexibility of thin film photovoltaic materials	5/62	A roof structure comprising panels in which a photovoltaic layers has been incorporated
G 4698234 - Vapor deposition of semiconductor material (1987)	2	Applicants' low band gap (1.45eV) semiconductor alloy material is of the same high quality (with respect to the number of defect states in the band gap thereof and the photoconductivity thereof) as the best 1.75	0/162	No In-Text Citations

			eVmaterial heretofore produced (and far superior to any low band gap semiconductor alloy material heretofore produced).		
H	4775425 - P and n-type microcrystalline semiconductor alloy material including band gap widening elements, devices utilizing same (1988)	1	A method of fabricating wide band gap microcrystalline n-type silicon alloy material to which have been added a band gap widening element.	3/161	Microcrystalline silicon semiconductors used in p-type or n-type semiconductor layers in solar cells of a pin structure use an amorphous i-type semiconductor layer
I	4309225 - Method of crystallizing amorphous material with a moving energy beam (1982)	2	Method does not require that the semiconductor material be heated to a temperature above the melting point of the crystalline material to achieve crystallization. It is in fact preferred to carry this method out completely as a solid phase transformation.	4/155	Technique for improving the crystallinity of semiconductor films by laser annealing and doping the films.
J	4021323 - Solar energy conversion (1977)	5	The invention is particularly concerned with the efficient conversion and storage of solar energy, and with the efficient use of materials for that purpose.	22/66	Semiconductor elements are also placed in contact with an electrolytic fluid. This results in a solar energy converter (semiconductor module) where electrolysis of the electrolyte takes place when illuminated with sunlight.
K	4234352 - Thermophotovoltaic converter and cell for use therein (1980)	4	Photons having less energy are either not absorbed or are absorbed as heat, and the excess energy of photons having more than 1.1 electron volts energy creates heat. These and other losses limit the efficiency of photovoltaic cells in directly converting solar radiation to electricity to the order of 10-20%.	14/60	A cell described therein includes a back surface reflector for reflecting incoming below band gap energy radiation back out of the cell.
L	4281208 - Photovoltaic device and method of manufacturing thereof (1981)	2	In order to be light transmissive, the film needs to be thinner; however, it is extremely difficult to control the film thickness of such a film and to attain good reproducibility.	15/67	A photovoltaic device which comprises a TCO front electrode, a semiconductor layer, and a back electrode stacked on a transparent insulative substrate such as a glass plate is described

4.3.1 Ex-ante Understanding of Technological Improvement

In Step 3.2a of the framework, the technological improvement as understood ex-ante and ex-post is determined and compared. The technical intent of the inventor is determined by reading selected sections of the very important patents. The fourth column in Table 3 describes the specific way these highly cited patents improved on the previous state of the art. Although the statements of improvement were clearer in some patents than others, overall it was possible to describe the technological change inherent in the invention as believed by the inventor at the time of invention.

For example, patent number 5747967 (row **D**) cites the following problem in the 'Prior Art' section of the patent:

'Once the optimal operating conditions have been determined, the systems intentionally change the operating conditions to non-optimal operating conditions and restart the process of determining the optimal operating conditions. Therefore, the disclosed systems do not provide for extended or sustained operation at the optimal operating conditions once the optimal operating conditions have been determined, which causes a loss in efficiency of the disclosed systems and reduces the power delivered by the photovoltaic device to a battery or load. Consequently, in spite of the well-developed state of solar array and photovoltaic array technology, there is still a need for a peak power tracker for a photovoltaic array that is simple to construct, operates the photovoltaic array at peak power output for significant periods of time, and allows the electric energy and electric power produced by the photovoltaic array to be stored in a battery.' (Bold type added)

This section can be reduced down to a condensed form of the need for the invention. This simplified need for the invention can then be combined with components from the abstract and/or the 'Summary of Invention' section of the patent. In this case, the abstract provides a clear summary of how the invention addresses that problem:

'A method and apparatus for maximizing the electric power output of a photovoltaic array connected to a battery where the voltage across the photovoltaic array is adjusted through a range of voltages to find the voltage across the photovoltaic array that maximizes the electric power generated by the photovoltaic array and then is held constant for a period of time. After the period of time has elapsed, the electric voltage across the photovoltaic array is again adjusted through a range of voltages and the process is

repeated. The electric energy and the electric power generated by the photovoltaic array is delivered to the battery which stores the electric energy and the electric power for later delivery to a load.' (bold type added)

These two statements were combined to provide a concise interpretation of the technical improvements disclosed in the patent as understood at the time of issue of the patent. The 4th column of Table 3 shows such summaries for all 12 of the very important patents.

There is a need for a peak power tracker for a photovoltaic array that operates the photovoltaic array at peak power output for significant periods of time (achieved by an apparatus that optimizes voltages to operate at peak power)

4.3.2 Ex-post Understanding of Technological Improvement

In order to understand how future inventors used the knowledge divulged in the patents, analysis of the in-text citations of the 'very important' patents is performed. The fifth column in Table 3 shows the number of in-text citations for each of the 12 'very important' patents, with approximately 11% of the citations being in-text. Of the 12 'very important' patents three were not cited in-text at all, conversely, patent number 4021323 (J) was cited in-text 22 times out of 67 forward citations. The sixth column of Table 3 shows a summary of the in-text citations that a patent received.

All 16 in-text citations that referred to patent number 5350644 (E) mentioned the dye sensitized solar cell that was claimed in the patent. In general almost all of the in-text citations were consistent with each other, meaning that the patents were cited for almost the exact same thing. There were only 2 patents that had clear separations between the citing patents. Patent number 4234352 (K) had citations both for its back scatter reflector and for the more general idea of a concentrator solar cell. Patent number 4281208 (L) had an even more divergent set of citing patents, with 8 patents citing the transparent conductive layer while the other 7 patents cited its potential use as an anti-counterfeiting tool. For the purposes of this study, the patents that had diverging sets of in-text citations used the citations that were more specific to the solar PV field for the summary of in-text citations. For example, patent number 5350644 (Row E) was cited in-text by 16 different patents. Two example in-text citations are shown below:

CITED BY 6649824 - Graetzel et al. (J. Am. Chem. Soc. 115(1993) 6382, U.S. Pat. No. 5,350,644) also reported that performance as high as that of a silicon solar cell was achieved by improving dye and a semiconductor electrode. There, a ruthenium type coloring agent was used as dye and an anatase type porous titanium oxide (TiO₂) was used as a semiconductor electrode.

CITED BY 6929970 - A wet type solar cell having a porous film of dye-sensitized titanium dioxide semiconductor particles as a work electrode was expected to surpass an amorphous silicon solar cell in conversion efficiency and cost. These fundamental techniques were disclosed in 1991 by Graetzel et al. in Nature, volume 353, pages 737-740 and in U.S. Pat. Nos. 4,927,721, 5,350,644 and JP-A 05-504023. Graetzel et al reported solid-state dye-sensitized mesoporous TiO₂ solar cells with up to 33% photon to electron conversion efficiencies.

These in-text citations were condensed into a qualitative and concise statement:

Graetzel et al reported solid-state dye-sensitized mesoporous TiO₂ solar cells with up to 33% photon to electron conversion efficiencies.

The summaries of technical improvement that are reflective of the original inventor are compared with the summaries of the in-text citations that are indicative of the acknowledged contributions of the patents. Of the 9 patents that were cited in-text, 6 show a nearly verbatim correlation between the technical improvement as stated at the time of invention and the future in-text citations. Patent number 5350644 (E) shows an example where the technical improvement is the addition of a dye to the titanium dioxide film that is used in solar cells to help absorb light in the visible spectrum, which is consistent with the citations that refer to the use of dye sensitized titanium dioxide cells in order to increase efficiency. The remaining three patents (4775425 (H), 4309225(I), 4234352(K)) also have consistency between claims and citations, but require a deeper reading to discover the relationship. For example, 4309225 (I) is cited in text about the laser annealing of a semiconductor film, which is a clear indicator of a solid state transition (not melting or phase change) that is mentioned in the original patent as an improvement over the prior art. In summary, the most important solar PV patents over the past 40 years are cited for the developments that their inventors believed were the basic for the innovation in the patent.

4.3.3 Technological Hierarchies of Solar PV Patents

Following Step 3.3 in Figure 1, the technological hierarchy levels of the 12 ‘very important’ patents were determined; these are summarized in the 2nd column of Table 3. For example, patent number 6878871 (Row B) replaces a polymer layer with a layer containing inorganic nanostructures, giving it a technological hierarchy level of 2: Material Substitution. Our first observation is that assessing technological hierarchy levels from patents is relatively straightforward through careful qualitative analysis.

Of the 12 ‘very important’ patents, 7 were either hierarchy levels 1 or 2, and none of the important patents involve changes at (the most “radical” change) level 6. The highest level change was observed in patent number 5747967 (Row D), which is a change at the system operation level with the addition of a component that allows the solar pv system to operate at the peak power for more of the time thus generating more electricity. The level of the technical hierarchy of the important inventions can also be examined for temporal trends -in the solar PV case, no change in the pattern over time is apparent.

4.4 Comparison of Important and Less Important Inventions

Step 4.1 involves selecting subsets of patents that are then compared using a variety of measures. Three sets of 12 patents each were selected, varying by number of forward citations, Table 4 shows the statistical analysis of the citation data of the three sub-sets. These sets are samples across time, with 3 patents in each set coming from each of the last 4 decades (1971-1981, 82-91, 92-2001, 02-11).

Table 4. The mean, variation and range of the forward citations of the 3 patent subsets

Type	N	Mean	Stdev	Min/Max
Very Important	12	86	45	57/163
Moderately Important	12	15	1.8	11/17
Less Important	12	2	0	2/2

Step 4.2 of the process analyzes the reliance on scientific literature in the sub-sets of patents, which is done by analyzing the non-patent literature citations of the patents. For instance, patent number 6441298 cites 50 other patents and has 6 citations of non-patent literature, all of which are scientific journal articles. In particular, patent 6441298 cites articles such as “Theory of Diffraction by Small Holes” (Beth, 1944) in *Physical Review* and “Extraordinary optical transmission through sub-wavelength hole arrays” (Ebbesen et al., 1998) in *Nature*. We follow the prior work by Trajtenberg et al (1997) in inferring that a patent with more NPL citations is more likely to be based on new scientific findings as compared to patents that have fewer NPL citations.

Table 5 shows that NPL citations for the 3 sub-sets of 12 patents vary between an average of 5% NPL citations and 29% of citations to the scientific literature. In order to determine if the three sets of data are statistically different, a Single Factor ANOVA tests was run on the three sets of NPL citation ratios. The P-value of the ANOVA test was .226, much higher than the P-value of .05 that would be required to indicate with 95% confidence that the 3 sets of data are statistically different. Although the ratios vary, the three samples are not statistically significantly different, suggesting that important patents do not rely more on NPL or scientific findings than patents of moderate or low importance.

Table 5. Comparison between the NPL citations of the three sub-sets of patents with outliers removed

Very Important	NPL/ Backward Citations	NPL Ratio	Mod Important	NPL/ Backward Citations	NPL Ratio	Least Important	NPL/ Backward Citations	NPL Ratio
6441298	6/41	15%	6924164	3/179	2%	6700057	1/8	13%
6878871	2/52	4%	6936761	2/24	8%	6472296	0/3	0%
6340788	5/14	36%	6531653	0/8	0%	6452090	1/6	17%
5747967	0/8	0%	5391236	0/6	0%	5479043	6/10	60%
5350644	5/9	56%	6118572	0/4	0%	5458695	0/4	0%
5092939	0/6	0%	5821597	5/16	31%	5401331	6/10	60%

4698234	0/0	0%	4612409	1/9	11%	4707561	0/6	0%
4775425	2/6	33%	5019177	0/8	0%	4412091	0/1	0%
4309225	0/11	0%	4658086	0/4	0%	4348545	0/2	0%
4021323	2/8	25%	4177083	0/5	0%	4228315	5/13	38%
4234352	4/7	57%	4254546	0/3	0%	4188239	1/2	50%
4281208	1/8	13%	3982265	4/12	33%	4156310	0/4	0%
Average	2.3/14.2	16%	Average	1.2/23.2	5%	Average	1.75.8	29%

Step 4.3 of the framework describes analysis of intra and inter field citations within the field of solar PV. The primary patent classes of the citations made in the important patents were analyzed in order to understand which fields impacted these important patents. Table 6 shows the distribution of the 20 most highly cited fields by the overall set of 2484 patents. A plurality of the patents in the overall list of 2484 were classified in the 'Batteries, Thermoelectric and Photoelectric', which was expected, as the field of interest is solar photovoltaics. In order to achieve a more accurate representation of the cited fields, a larger sample size was used to represent the important patents in this comparison as shown in the final column of Table 6.

Table 6. Comparison of top 10 most highly cited patent classes between the overall set of 2484 patents and the 250 most highly cited patents

Patent Class Name	All 2484 Patents	Top 250 Patents
Batteries: thermoelectric and photoelectric	11%	20%
Semiconductor device manufacturing: process	5%	6%
Radiant energy	4%	6%
Active solid-state devices (e.g., transistors, solid-state diodes)	4%	5%
Stoves and furnaces	3%	1%
Communications: electrical	3%	2%
Illumination	2%	1%
Optical: systems and elements	2%	4%
Electricity: battery or capacitor charging or discharging	2%	1%
Stock material or miscellaneous articles	2%	1%

The most cited field for both the 250 important patents and the overall set was 'Batteries: thermoelectric and photoelectric', the same field in which most of the patents are classified. It is interesting to note that the 250 important patents cite the 'Batteries...' class at a rate 2 times that of the full list of patents. Beyond this, the top 4 most cited classes are the same for the overall set of patents and the 250 important patents. When the complete lists of citation rate by class are compared using a t-test, the results show that the two distributions are not significantly different, with a P-value of 0.9, very far from the P-value of 0.05 needed to claim significantly different distributions with 95% confidence.

The textual structure of the very important and less important sub-sets of patents is compared in Step 4.4 of the framework and is summarized in Table 7. A comparison of the difference in length of abstract (in number of characters) between the important and moderately important patents shows that the less important patents actually have slightly longer abstract lengths than the important Type 1 patents. This difference is potentially statistically significant, with a P-value of 0.08. Ultimately the data does not establish that more important inventions have longer or shorter abstracts. The embodiment counts of the very important and less important patents were compared. The two data sets are not statistically different, with a P-value of .13. There is not evidence in this study that more important patents have a different number of embodiments of invention than other patents. Finally, the length of the text in the prior art sections was compared, showing no significant difference, with a P-value of 0.14. The data does not indicate that the more important patents have a significantly shorter or longer description of prior inventions. Ultimately, the structural text analysis did not show a significant difference between the sub-sets of patents for this case study.

Table 7. Other comparisons between the 3 sub-sets of patents

Very Important	Abstract Length	Embodiment Counts	Prior Art Lengths	Less important	Abstract Length	Embodiment Counts	Prior Art Lengths
6441298	723	0	2705	6924164	663	6	5285
6878871	743	38	2878	6936761	906	3	6395
6340788	461	11	6071	6531653	835	1	6724
5747967	723	0	6834	5391236	732	1	2025
5350644	395	0	2001	6118572	870	2	4795
5092939	883	3	2533	5821597	1016	0	7524
4698234	452	0	87442	4612409	449	0	4290
4775425	298	7	21798	5019177	491	0	9747
4309225	224	0	5148	4658086	578	1	4171
4021323	552	0	N/A	4177083	464	0	1155
4234352	759	1	3241	4254546	458	0	4098
4281208	668	2	6341	3982265	895	1	1495
Average	573.4	5.2	13362	Average	696.4	1.3	4808

7. Conclusions

This paper presents a framework for understanding the contributions of individual inventions in a technical field over time. This framework attempts to improve upon the current state of the art by providing a qualitative and quantitative method that is rigorous, repeatable and useful.

The methodology described in this paper was demonstrated using the field of Solar PV as a case study. The underlying assumption of the framework and supporting tools were supported by the case study results. In particular, it was shown that an objective determination of the basic claimed improvement due to a specific patent by analysis of the prior art and claims sections of the patent was consistent with the use of the patent by later inventors. This result is an important confirmation of Trajtenberg's suggestion that patent citation data is meaningful. We also showed in the case study that it was feasible to construct a useable data set. Even with this imperfect data set, it was possible to gain an understanding of the technological development of Solar PVs. The results also showed that a hierarchy construct can be applied to patent analysis as assumed in the framework with small uncertainty.

By examining just this one case, several potentially interesting results concerning technological development were found. Perhaps most importantly no discontinuously important inventions are found. Thus, the methodology is shown to be operational and have utility in discerning the nature of technological change in a field.

7.1 Opportunities for Future Work

While this research provides a starting point, there is still a significant amount of work to be done in understanding how inventions contribute to the increases in the overall technical capability of a field. First, this methodology should be performed on more case studies of different technological fields to examine its applicability to other technical areas. One finding from further cases might be deeper understanding of the difference between important inventions and less important ones. In this report, we found many results to show that they were not different, which, while interesting, could be significantly aided by more clear differentiated factors between the two.

References

- Abraham, B. P., & Moitra, S. D. (2001). Innovation assessment through patent analysis. *Technovation*, 21(4), 245-252. [http://dx.doi.org/10.1016/S0166-4972\(00\)00040-7](http://dx.doi.org/10.1016/S0166-4972(00)00040-7)
- Ahuja, G., & Lampert, C. M. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, 22(6), 521-543. <http://dx.doi.org/10.1002/smj.176>

- Christensen, C. (1997). *The innovator's dilemma: When new technologies cause great firms to fail* (Chapter 1). Cambridge, MA: Harvard Business School Press.
- Dahlin, K., Taylor, M., & Fichman, M. (2004). Today's edisons or weekend hobbyists: Technical merit and success of inventions by independent inventors. *Research Policy*, 33(8), 1167-1183. <http://dx.doi.org/10.1016/j.respol.2004.06.003>
- Ebbesen, T. W., Lezec, H., Ghaemi, H., Thio, T., & Wolff, P. (1998). Extraordinary optical transmission through sub-wavelength hole arrays. *Nature*, 391(6668), 667-669. <http://dx.doi.org/10.1038/35570>
- Girifalco, L. A. (1991). *Dynamics of technological change*. New York: Van Nostrand Reinhold.
- Griliches, Z. (1992). The search for R&D spillovers. *Scandinavian Journal of Economics*, 94, S29-S47. <http://dx.doi.org/10.2307/3440244>
- Harmon, C. (2000). Experience Curves of Photovoltaic Technology, IR00-014, IIASA, Laxenburg, Austria.
- Henderson, R. M., & Clark, K. B. (1990). Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, 9-30. <http://dx.doi.org/10.2307/2393549>
- Hur, K. I., & Watanabe, C. (2001). Unintentional technology spillover between two sectors: Kinetic approach. *Technovation*, 21(4), 227-235. [http://dx.doi.org/10.1016/S0166-4972\(00\)00037-7](http://dx.doi.org/10.1016/S0166-4972(00)00037-7)
- Kaplan, S. M. (1999). Discontinuous innovation and the growth paradox. *Strategy & Leadership*, 27(2), 16. <http://dx.doi.org/10.1108/eb054631>
- Magee, C. L. (2010). The role of materials innovation in overall technological development. *Journal of Management*, 62(3), 20-24.
- Mascitelli, R. (2000). From experience: Harnessing tacit knowledge to achieve breakthrough innovation. *Journal of Product Innovation Management*, 17(3), 179-193. [http://dx.doi.org/10.1016/S0737-6782\(00\)00038-2](http://dx.doi.org/10.1016/S0737-6782(00)00038-2)
- Markides, C. (2006). Disruptive innovation: In need of better theory. *Journal of Product Innovation Management*, 23(1), 19-25. <http://dx.doi.org/10.1111/j.1540-5885.2005.00177.x>
- Sahal, D. (1981). *Patterns of technological innovation*. Reading, MA: Addison-Wesley Pub. Co.
- Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *The Rand Journal of Economics*, 172-187. <http://dx.doi.org/10.2307/2555502>
- Trajtenberg, M., Henderson, R., & Jaffe, A. (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, 5(1), 19-50. <http://dx.doi.org/10.1080/10438599700000006>
- United States Department of Energy. (2009). Solar energy technologies program FY 2008 annual report.
- United States Energy Information Agency. (2011). Annual energy review 2010.
- Utterback, J. M. (1996). *Mastering the dynamics of innovation*. Boston, MA: Harvard Business Press.
- Wang, C., Chin, Y., & Tzeng, G. (2010). Mining the R&D innovation performance processes for high-tech firms based on rough set theory. *Technovation*, 30(7-8), 447-458. <http://dx.doi.org/10.1016/j.technovation.2009.11.001>
- Yu, D., & Hang, C. C. (2010) A reflective review of disruptive innovation theory. *International Journal of Management Reviews*, 12(4), 435-452. <http://dx.doi.org/10.1111/j.1468-2370.2009.00272.x>

Notes

Note 1. Exploration of search terms is most effectively done on an online database. For our initial tests, we used a website called patsnap.com

Note 2. A concern is the changing rate of innovation and patenting over time, as the number of inventions patented per year rises over time and thus may make recent inventions seem more important (Dahlin et al, 2004).