



# The Emerging Clusters Project

Anthony Breitzman Ph.D.

Presented at Evaluation 08  
Denver, Colorado  
November 8, 2008

# Overview

---

- In this talk I will present some ideas and highlights from the Emerging Technological Clusters project sponsored by the U.S. Department of Commerce's Technology Administration. Final report at <http://www.ntis.gov/pdf/Report-EmergingClusters.pdf>
- The project aimed to develop and validate a methodological tool based on patents, citations, co-citations, and clustering of patents in order to identify emerging, high-risk, early-stage, technologically innovative activities.

# Motivation for Analysis

---

- Policy makers, economists, and regional development planners need a repeatable, mechanical, objective method for identifying emerging technology trends
- In addition to determining emerging trends, it is also necessary to ascertain where new technologies are being developed, and by whom
- This project describes a novel technique for identifying and mapping clusters of important, emerging patented technologies

# • What We Know About Patents and Citation Analysis

---

- Patents build on earlier patents
  - (AKA prior art, backward citations, references)
- Citation analysis counts the number of citations (forward citations) received by a patent
- Most patents are only cited a few times in their lifetimes and very few are cited a lot
- Frequently cited patents typically represent key technologies that have led to many subsequent innovations
- High citation is correlated with various measures of success
  - awards, licensing revenue, increases in sales and profits, etc.
- Patents that build on highly cited patents are themselves cited more than average as they age

# Weaknesses of Citation Analysis

---

- Citation Analysis is a good evaluation tool.
- It can be used to compare impact of individual labs, companies, industries, countries, etc.
- However, citations take time to accumulate, making citation analysis a poor tool for identifying emerging technologies.

# Hot Patents Solution

---

- To identify emerging technologies we need a forward looking indicator
- We wish to concentrate on recently (no older than 18 months) issued patents
- About 275,000 patents issued in last 20 months; need a filter to identify small fraction of important patents
- Citation analysis provides a filter, but can't be used on recent patents
- Solution is to identify recent patents that build upon “technology leaps” rather than “incremental improvements”

# Key Steps (1 of 2)

---

## 1. Identify Hot-Patents

- Weed out “citation-classics” – older key technologies that continue to be refined incrementally. (e.g. LCD Panels, Ink-Jet technology, transdermal patches, etc.)
- Hot-Patents receive the bulk of their citations from recent patents.
- A Hot-Patent can be 1 to 25 years old
- Has to have 10+ recent citations, and the proportion of recent cites to total cites must be proportional to its age. (Older patents need 25% of their cites as recent to be hot patents; newly issued patents need a much higher proportion)

## 2. Cluster Hot-Patents

- Hot-Patents will necessarily have many citing patents, making them great candidates for co-citation clustering (diagram 2 slides ahead)

## Key Steps (2 of 2)

---

### 3. Next-Generation Clusters

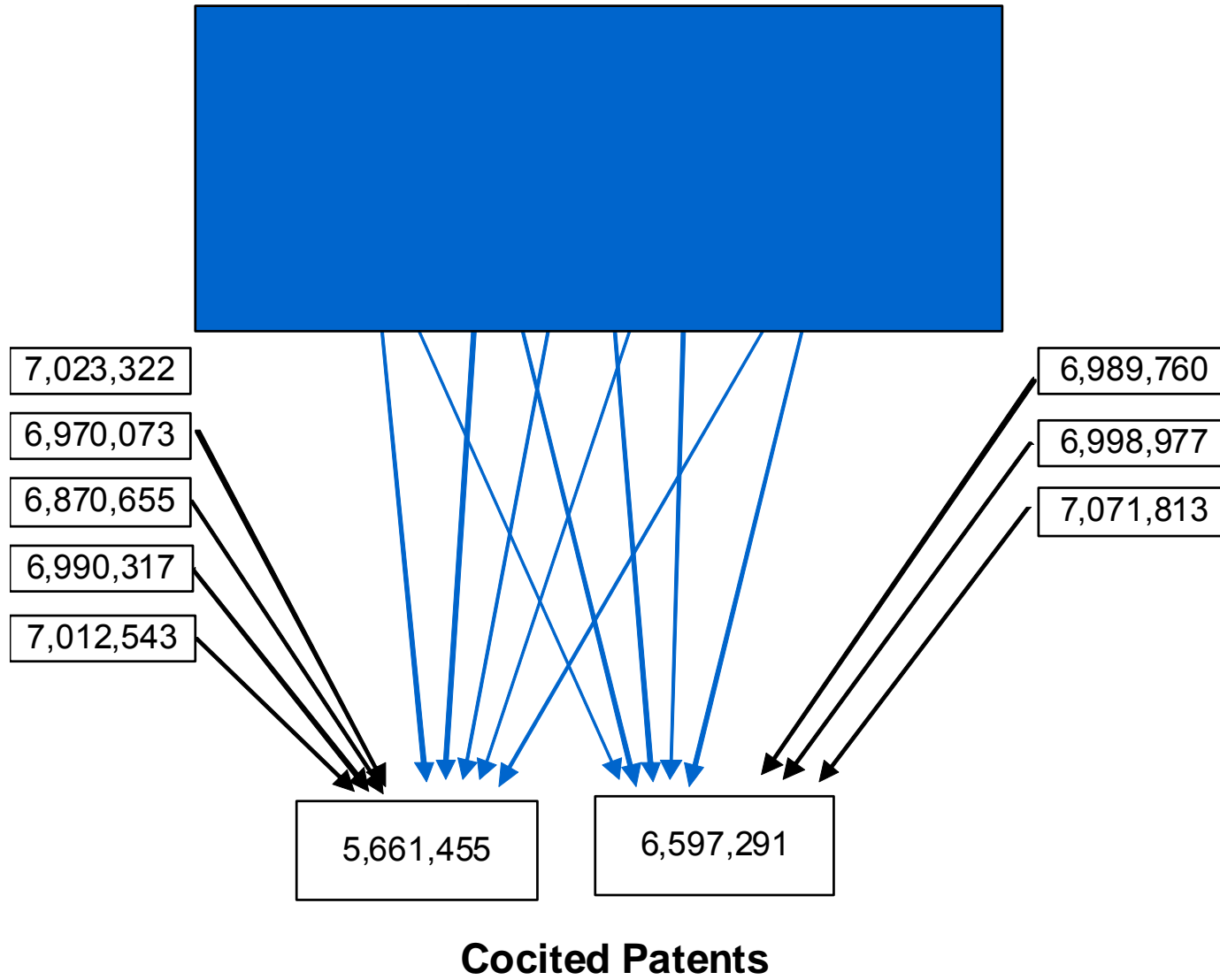
- Hot-Patent Clusters are interesting, but old technology
- Hot-Patent Clusters give rise to Next-Generation Clusters, which consist of recent patents that use Hot-Patents as prior art

### 4. Scoring and Ranking of Clusters

- A Scoring method based on indicators such as public-sector presence, science intensity, and diversity of prior art are used to rank clusters
- Top 50 US Patent Clusters and Top 50 EP/WIPO Clusters examined in more detail



# Co-Citation Provides a Natural Clustering Mechanism for Cited Patents



## More on Next-Generation Clusters

---

- Represent current patent activity by many entities centered around like technologies
  - e.g. short range radar object locator hot-patent gives rise to next generation patents related to touchless faucets, occupant detection systems, sealed weapon systems etc.
- Contain patents issued within last 20 months
- Represent about 20% of all patents issued in the last 20 months

# The Emerging Clusters Project Models

---

- Identified Next-Generation Clusters from 3 Time Periods: 1998, 2002, 2006.
- The 2006 set was built using the then current database. The 1998 and 2002 sets were built using only what would have been available in 1998 and 2002.
- The 1998 and 2002 sets were used for validation and to build a predictive model.
- The model was applied to the 2006 set to rank the top 100 clusters that were most likely to contain emerging technologies.

## First Validation

---

- We took a known set of emerging, high-risk patents related to NIST's Advanced Technology Program (ATP) and found that they were twice as likely to be found in the next-generation clusters as in the general population of patents
- This suggests that the Next Generation Clusters contain high-risk, emerging technologies in a higher concentration than the patent system as a whole
- Also suggests that ATP dollars are likely having a broad impact beyond individual award recipients

# Twice as Many ATP Patents are Found in Next Generation Clusters as Expected

Time Period	# Patents	# Next Gen Patents	% Next Gen Patents	# ATP Patents	Actual # ATP Patents in NG	% ATP Patents found in NG	Expected # ATP Patents in NG	Actual / Expected #NG Pats	# ATP Projects with Patents	Actual # ATP Projects with NG Pats	Expected # ATP Projects in NG	Actual / Expected # ATP Proj. in NG
Jan. 1997- Aug. 1998	208905	43223	21%	64	28	44%	13	<b>2.11</b>	42	21	9	<b>2.42</b>
Jan. 2001 - Aug. 2002	275175	66216	24%	92	43	47%	22	<b>1.94</b>	51	26	12	<b>2.12</b>

## Second Validation

---

- The second validation is a classic “back-test.” The 1998 and 2002 next-generation clusters were constructed as they would have been in 1998 and 2002
- However now that time has passed we can test these older clusters using current evaluation methods including citation analysis
- We find that patents that the top 10% emerging clusters that would have been selected in 1998 and 2002 are now cited on average 50% higher than typical patents from 1998 and 2002 in the same technology classes
- This suggests, that patents that are selected in the current next generation should be cited 50% higher than their peers in 5 years.

## Scoring and Ranking of Clusters (1 of 2)

- The first three steps of the method generate thousands of next generation clusters
- Needed a way to rank those that were most likely to contain emerging, high-risk technologies
- Using the 1998 and 2002 sets containing ATP patents, we were able to identify parameters that emerging clusters are likely to contain:
  - **Public Sector Participation** – Clusters with ATP patents are more likely to contain patents from universities and government labs (even though ATP participants are restricted to commercial enterprises)
  - **Science Linkage** – Clusters with ATP patents tend to use more scientific papers for prior art (even when normalized by technology class) suggesting they are building upon new ideas as well as existing technologies
  - **Originality** – Clusters with ATP patents have patents that tend to build on prior art from a number of different technologies, suggesting they are putting multiple ideas together to build something new, rather than creating incremental improvements to existing technologies

## • Scoring and Ranking of Clusters (2 of 2)

---

- A Scoring Equation was built using these parameters so that each cluster could be ranked as to whether it was most likely to contain emerging technology
- Note that unlike citation measures, these parameters are available as soon as a patent is published
- Thus a 2008 set of emerging clusters containing patents through October 2008, could be built and ranked today



# Key Results From Top 100 Emerging Clusters

---

- US Inventors create 50% more emerging technologies than expected
  - US Inventors produce 51% of all US granted patents, but 73% of patents in the emerging clusters of US patents
- US lead is large
  - US inventors produce 73% of emerging clusters (in US patent set), second place Japanese produce 13%
  - US inventors produce 52% of emerging clusters (in EP/WIPO patent set), second place Germans produce 16%

## Key Results (II)

---

- Japanese result is surprising
  - Japanese invent 21.2% of US granted patents, but only 13.4% of patents in US emerging clusters
  - Japanese invent 17.4% of EP/WO patents, but only 6.1% of patents in Emerging EP/WO clusters

## Key Results (III)

---

- Information Technology (IT) is Increasingly Important in the US
  - 60% of patents in Emerging US Clusters are related to semiconductors, computers, or communications.
  - 90% of Top 100 Emerging US Clusters have at least one IT patent.
  - These include clusters related to engines, biotechnology, medical devices. In short almost every technology cluster has some IT related component.
- In contrast, 60% of patents in EP/WO Emerging clusters are life science related
  - This does not suggest that the US is lagging in life science
  - A large fraction of the EP/WO life science patents come from US inventors and companies

## Key Results (IV)

---

- Much of the US Emerging technology patents are invented on the west coast
  - 22% of the patents in the US Emerging Clusters were invented on the coast between San Francisco and Los Angeles
  - About 1/3 of the patents were invented in 3 states (CA, OR, WA) that account for only 16% of the US Population

## Some Limitations

---

- The method does have limitations:
  - Not all innovations are patented - some are kept as trade secrets
  - The method is a statistical scoring method, and inevitably some emerging technologies will be missed, and some high scoring clusters will turn out to be mundane technology
  - Patenting varies widely by technology – e.g. retail related innovations are not patented widely, whereas Information Technology (IT) patents exist at a rate disproportionate to the level of IT company revenues
- However, the great advantage of using the patent system for identifying emerging clusters (which outweighs these limitations) is that the patent system is the single largest collection of innovations in existence
- Further, a geographical location is identified for every inventor, making innovation studies at a regional level possible

## Example Cluster #6- 3D Memory

---

- Matrix Semiconductor developed three-dimensional non-volatile memory architectures ~ 2005
- Traditional flash memory is based on single layer crystalline silicon substrates, so storage devices using this technology are essentially two dimensional
- Matrix developed three-dimensional stackable memory devices, which provided more power and took up less space
- In 2006, Sandisk paid Matrix \$250 million to acquire its 3D technology
- Sharp, Sony, Micron, and Infineon are all building upon Matrix's 3D technology
- Sleeper technology. Underlying hot-patent is 30 years old. #3,886,577 Filament-type memory semiconductor device and method of making the same. Largely ignored and then 50 citations since 1999.

## Example: Testing ATP Outreach

---

- Inventor addresses can be used to identify where emerging technologies are being developed, how this has changed over time, etc. One application is to see if ATP outreach corresponds to emerging cluster locations
- In general, ATP outreach was hitting the main areas
  - Top 10 Next-Gen regions contain 52% of the patents and 43% of the ATP applications
  - Top 20 Regions contain 70% of patents, 60% of ATP Apps
  - Top 30 Regions Contain 78% of patents, 69% of ATP Apps
- Some regions were more successful at winning ATP awards than others. Among regions with 10+ ATP awards,
  - Albany, NY applied 120 times and received awards 30% of the time
  - San Jose and Detroit received awards 18% and 22% of the time
  - Atlanta, Los Angeles, and Washington, are less successful with only an 11% hit rate
- Areas where ATP could have increased its outreach effort include: Austin, TX; Dallas, TX; Seattle, WA; Boise, ID; Burlington, VT

## Example: Small Businesses and Emerging Technologies

---

- Because of costs and challenges involved in obtaining patents, only about 8% of all US patents are owned by small firms (fewer than 500 employees)
- In a project for the Small Business Administration (SBA) we showed that 24% of patents in the emerging clusters are owned by small firms (3 times as many as expected)
- This suggests small firms generate an outsize proportion of emerging technology relative to their overall level of invention



## Conclusions

- Patents have been accepted by researchers as backward looking tools to study innovation as well as current looking tools to measure impact. We've developed and validated a tool that uses patents to look forward
- Identifying emerging technologies is inherently difficult, but we've developed an algorithmic, objective, repeatable method to identify likely candidates
- We barely touched on this, but patents identify locations as well as technologies making geographic studies of emerging technologies feasible
- US inventors and firms are surprisingly strong in generating emerging technologies. ATP participants and small firms are stronger than expected as well

- 
- 
- 

## Contact Information

---

### **1790 Analytics LLC**

130 Haddon Avenue  
Haddonfield, NJ 08033  
Tel: 856-216-1790

### **Anthony Breitzman**

[abreitzman@1790analytics.com](mailto:abreitzman@1790analytics.com)