The Science of Innovation

Joseph Engler

Outline

1. What is Innovation
2. Theories of Innovation
3. Evolving Innovation (CAS)
4. Collaborative Innovation

What is Innovation

An innovation is invention + exploitation
-Karim Hirji

Continuous innovation occurs largely because a few key executives have a broad vision of what their organizations can accomplish for the world and lead their enterprises toward it.
- James Brian Quinn

Invention, converting an idea into a new process or product.
Innovation, extending the utilization of a product or process.
- Abol Ardalan

Innovation is an iterative process aimed at the creation of new products, processes, knowledge or services by use of new or existing scientific knowledge.
-Andrew Kassik

Innovation Definition

• New Product, Service or Methodology

• Must be accepted by the marketplace

• Can be built upon existing innovations (Incremental Innovation)

• Can be built using relatively few current innovations in a very unique manner (Disruptive Innovation)
Inventions not Innovations

Patent database contains items that were never accepted in the marketplace.

Inventions not Innovations

Some patents may not represent true innovation

Theories of Innovation

Altshuller’s Theory (TRIZ)

Ulwick’s Theory

Von Hippel’s Theory

Zhang’s Theory

Altshuller’s Theory (TRIZ)

Innovation proceeds in a systematic manner building upon and adding to the body of inventive knowledge.
Theories of Innovation

Ulwick’s Theory
Innovation is a process of engineering pushed to the consumer

Von Hippel’s Theory
Innovation is a constant feedback loop in which the customer input is valued and drives change in the innovation

Zhang’s Theory
Innovation is the functional integration around a core project that is strategic to the corporate setting.

Systems View of Innovation

- The system of innovation proceeds in an evolutionary manner, building upon previous individual innovations.
- The system of innovation should be studied in the context of biological systems.
- Complex Adaptive Systems best serve the study of the evolution of the innovation system as inspired by biological systems.
- The system of innovation can be modeled utilizing agent-based modeling in the context of Complex Adaptive Systems.
- Collaboration between innovating entities can increase innovation success in rapidly changing market domains.
Evolvable Innovation

- How can innovation be evolved?
  - Start with an initial product, service or methodology idea.
  - Perform web mining and other research on the domain of the idea.
  - Collect a set of requirements for the idea to become an innovation by utilizing data mining techniques on the web mined information.
  - Utilize Genetic Programming to generate a population of highly fit ideas, some of which would meet the requirements of innovation.

- Telephone
  - Initially a single purpose device utilizing a wired grid through poles and companies.
  - Combined with wireless communication technology.

- Cell Phones
  - Initially a single purpose device for voice communication.
  - Combined with Email to enhance communication abilities.
  - Text messaging.
  - Smart Phones.
  - Blackberry.

Evolution in the System of Innovation

- Utilize Genetic Programming to generate new individual innovations.
  - Initial population consists of existing individual innovations.
  - Each individual innovation is represented as a tree with each node of the tree representing a distinct property of the individual innovation.

- Define the fitness function for the GP based upon innovation requirements mined from the internet.
  - User reviews.
  - Expert reviews.
  - Patent trends.
  - Market analysis.

- Utilize multi-parent crossover instead of single point crossover in the GP to achieve the limited number of epistatic interactions.

Multi-parent Crossover

Multi-parent crossover illustrated with $n = 3$ parents of a genetic algorithm whose genotype is represented as bit strings.
Evolution in the System of Innovation

- Algorithm:
  1. Define fitness function based upon requirements mined from the web (strong enough definition to account for phenotypic contradictions in multi-parent crossover)
  2. Generate initial population from existing individual innovations
  3. Perform multi-parent crossover on population
  4. Perform mutation with low mutation rate on population
  5. Evaluate fitness of individual innovations
  6. Perform Tournament Selection for recombination
  7. Goto step 3 for i number of iterations

- It is best to limit the number of iterations in individual innovation evolution to ensure the new individual innovation does not attempt to span too large a space on the fitness landscape thus decreasing its chances of market success.

- New individual innovations may be predicted based upon the use of this methodology
Industry Examples Support CAS Theory of the System of Innovation

- Telephone
  - Initially a single purpose device utilizing a wired grid through poles and companies.
  - Combined with wireless communication technology
    - Cell Phones
      - Initially a single purpose device for voice communication
      - Combined with Email to enhance communication abilities
        - Text messaging
        - Smart Phones
        - Blackberry
  - Limited number of epistatic interactions at each generation

- Computer Printer (consumer market)
  - Originally dot matrix with ribbons of ink and a rotating head to "type" characters
  - Combined with Xerox type copying machines to replace "typing" heads.
    - Laser and Inject Printers
      - Originally still just a printer
      - Combined with the Facsimile machine
        - Multi-purpose printers
        - Today it is possible to print, fax, print photos, scan and send emails from a home printer
  - Again limited number of epistatic interactions at each generation

Collaborative Innovation

- Often referred to as “Open Innovation”
- Innovation occurring from the combination of ideas from various entities
- Utilizing an existing innovation in a new manner with minor changes
- Working with the most authoritative person(s) in a given domain to create new innovations

Collaborative Authoritativness

- Focused Web Mining for Papers/Articles
- Creation of Authoritativness Matrix
- Apply Authoritativeness Metric
- Document Clustering
Focused Web Mining

- Standard Web Crawling Methodologies
- Download of Query Specific Files
- Forms the repository from which Text Mining takes place

Authoritativeness Matrix

- Authors names are parsed from the documents
- Cited references are parsed from the documents
- The publication date is parsed from the document

Authoritativeness Matrix

<table>
<thead>
<tr>
<th>Author</th>
<th>P</th>
<th>D</th>
<th>K</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kusiak A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2006</td>
</tr>
<tr>
<td>Lin G</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2004</td>
</tr>
<tr>
<td>Stokic D</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1999</td>
</tr>
</tbody>
</table>

Parsing of Names

- Heuristics
  - Name should be first non-empty line after title
- Regular Expressions
  - `\w+\s\w\[\]`  
  - `\w+\s\w\[.\]`  
- Names Database with Dice Coefficient
Dice Coefficient

- Create Bigrams of the two words being compared.
  - Night = \{ni, ig, gh, ht\} = X
  - Nacht = \{na, ac, ch, ht\} = Y

- Calculate Similarity

\[
Dice_{coef} = \frac{2(X \cap Y)}{|X| + |Y|}
\]

Authoritativeness Metric

- Scan Authoritativeness Matrix
  - Create Hash of Authors (row)
  - Create Hash of Referenced Authors (column)
  - Create Hash of Average Age of Document for each author

- Authors Hash (rows) measures Out-Links
- Reference Hash (columns) measures In-Links

Authoritativeness Metric Cont.

- Calculate the initial authoritativeness for each author and referenced author.

\[
A_i = ln(\lambda^t(out_i) + in_i)
\]

- \(out_i\) is the number of out-links
- \(in_i\) is the number of in-links
- \(\lambda\) is a user defined weight parameter of document age in [0,1]
- \(t\) is the average age of the document for author \(i\)

Authoritativeness Boosting

- Similar to PageRank Algorithm
- Iterative Approach

- If an authoritative author references a paper, the in-link to that reference is increased
- In-Links of less authoritative authors pose no detriment
In-Link Boosting

- Calculate the mean of the in-links
  \[ \overline{\text{ein}} = \frac{\sum_{j=1}^{N} e^{\text{ein}}_j}{N} \]

- Update the Authoritativeness Metric
  \[ \text{in}^{i'}_j = \sum_{j=1}^{N} \left\{ \begin{array}{ll}
  e^{\text{ini}}_j - \overline{\text{ein}} & \text{if } e^{\text{ini}}_j > \overline{\text{ein}} \\
  1 & \text{if } e^{\text{ini}}_j \leq \overline{\text{ein}}
\end{array} \right. \]

Determining Authoritativeness

- Order the authors by Authoritativeness
- Select Top K authors
- Cluster the documents
- Find cluster closest to current issue
- Find most authoritative k authors for that issue
  - Authors that are authoritative overall may not be authoritative on specific topic
  - Possible application of Apriori Principal

Experimental Results

Determine the most authoritative person in the field of Genetic Algorithms.

- 945 Articles on Genetic Algorithms
- Unclustered to determine overall authority
- 30 iterations of In-Link Boosting
- \( \lambda \) set to 1 to not discount older authoritativeness

<table>
<thead>
<tr>
<th>Author</th>
<th>Original Authoritativeness</th>
<th>Boosted Authoritativeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. H. Holland</td>
<td>5.278</td>
<td>6.2122</td>
</tr>
<tr>
<td>J. R. Koza</td>
<td>4.8828</td>
<td>5.8105</td>
</tr>
<tr>
<td>F. H. Bennett</td>
<td>4.5849</td>
<td>5.1350</td>
</tr>
<tr>
<td>L. Altenberg</td>
<td>4.3438</td>
<td>4.4306</td>
</tr>
<tr>
<td>D. Andre</td>
<td>3.9512</td>
<td>4.0105</td>
</tr>
</tbody>
</table>

Note the minimal boost of the last two authors.
<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
</table>