The Network Value of Products

Gal Oestreicher-Singer
(Tel Aviv University)

Joint work with
Eyal Carmi (Google, NY)
Barak Libai (IDC, Israel)
Liron Sivan (TAU, Israel)
Ohad Yassin (TAU, Israel)
The Network of Products
Consider the online interface which consumers face in eCommerce sites.
This book “recommends” the purchase of other books.

Customers Who Bought This Item Also Bought

- Linked: How Everything Is Connected to... by Albert-Laszlo Barabasi
  - Rating: 5
  - Price: $10.88

- Six Degrees: The Science of a Connected Age by Duncan J. Watts
  - Rating: 5
  - Price: $10.89

- Networks, Crowds, and Markets: Reasoning About a... by David Easley
  - Rating: 5
  - Price: $41.47

- Bursts: The Hidden Pattern Behind Every... by Albert-Laszlo Barabasi
  - Rating: 5
  - Price: $10.78

- Social Network Analysis (Quantitative Ap... by Professor David Knoke
  - Rating: 5
  - Price: $14.23
And other books recommend this book.
Actually, this can be seen as a “product network”
This is the “aisle structure” of electronic commerce
Summary of previous results

Visible networks and the amplification of demand correlations

• The visible presence of links between complementary products *triples*, on average, measured demand complementarity
• The influence of visible networks is *higher for more popular* products, is higher *from* more popular products and is higher for more recent products

(Management Science, forthcoming)

Visible networks and demand/revenue distributions

• An increase in the influence of product recommendation networks is associated with *evening out* of both demand and revenue fractions across products (leading to a more pronounced “long tail”)
• Network properties affect the likelihood that consumers notice links more frequently (and hence “spread” demand more evenly)

(MIS Quarterly, forthcoming)
Today’s research questions

• Can we suggest a relatively simple algorithm, to calculate the network value of products in large-scale product networks?
• If applied – will such an algorithm provide insights as to the contribution of different sales levels (e.g., the “Long Tail”) of products to the firm?

Why should we care?

● **Help online sellers**
  ● to recognize which products to offer (or to stop offering) to customers,
  ● which products to promote.

● **Retailer-Manufacturer aspects**
  ● Most product networks managed by online retailers include multiple items made by different manufacturers.
  ● Retailers and manufacturers may have **different interests**
Inter-product effects do not only exist outside the web

Attribution of “Value” in a product network

\[
\text{Revenue (u)} = \text{Intrinsic Value (u)} + \text{Incoming Value (u)}
\]

The “independent” value of a product in this system

The contribution \textbf{from} other products

\[
\text{Network Value (u)} = \text{Intrinsic Value (u)} + \text{Outgoing Value (u)}
\]

The “full” value of a product to the seller

The contribution \textbf{to} other products
Here we would like

- To suggest a relatively simple algorithm, in the spirit of Google’s PageRank, to calculate the network value of products in large scale product networks

- Apply it to the Amazon network (and also to other sellers), demonstrating how this method helps us better understand the contribution of different sales levels (e.g., the “Long Tail”) of products to the firm.
A product network value model

*Impressions*(ν) is the number of people visiting product ν’s page.

α_{ν→u} is the *Recommendation Conversion Rate* (RCR)
- The probability that a link exposure will turn into a purchase
- In the case of books in Amazon, this probability is a combination of the probability that a link will be clicked on, and the probability that the visit to the next page will result in a purchase due to the link.
The Incoming Value of the product: The revenue portion of the product that is attributed back to the network is:

\[
\text{Incoming Value}(u) = \sum_{v \in \text{In}(u)} \alpha_{v \rightarrow u} \cdot \text{impressions}(v) \cdot P(u)
\]

Where \( P(u) \) is the price of product \( u \).

The Intrinsic Value of the product: The revenue portion of the product that is not attributed back to the network is:

\[
\text{Intrinsic Value}(u) = \text{Revenue}(u) - \text{Incoming Value}(u)
\]

\[
= P(u) \cdot \left( Q(u) - \sum_{v \in \text{In}(u)} \alpha_{v \rightarrow u} \cdot \text{impressions}(v) \right)
\]
The Outgoing Value of the product: The revenue portion of the product that is generated by outgoing links to other products in the network is:

\[
Outgoing \ Value(u) = \sum_{v \in \text{Out}(u)} \alpha_{u \rightarrow v} \cdot \text{impressions}(u) \cdot P(v)
\]
Then, the **network value** of the product (the value it provides the system)

\[
\text{NetworkValue}(v) = \text{IntrinsicValue}(v) + \text{OutgoingValue}(v)
\]

\[
= \left[ Q(v) - \sum_{u \in \text{In}(v)} \alpha_{u \to v} \cdot \text{impressions}(u) \right] \cdot P(v)
\]

\[
+ \sum_{u \in \text{Out}(v)} \alpha_{v \to u} \cdot \text{impressions}(v) \cdot P(u)
\]
When applied over a given graph once, our model pushes extrinsic revenues back to the originating items located one link away from the recommended products.

Some of the revenue from sales of item C that is attributed to item B should in fact be attributed backwards to item A, which generated part of item B's traffic to begin with.

In fact, B’s actual contribution to C's revenue should be decreased by the proportion of A’s contribution to B

\[
OutgoingValue_n(v) = \text{OutgoingValue}_{n-1}(v) - \beta^{n-1} \cdot \frac{\text{IncomingValue}(v)}{\text{Revenue}(v)}
\]

\[
\cdot \text{OutgoingValue}_{n-1}(v) + \beta^{n-1}
\]

\[
\cdot \sum_{u \in \text{Out}(v)} M(v, u) \cdot \frac{\text{IncomingValue}(u)}{\text{Revenue}(u)} \cdot \text{OutgoingValue}_{n-1}(u)
\]
An empirical application: The Amazon books product network

- Over 900,000 books sold on Amazon.com on a particular day in 2010.

<table>
<thead>
<tr>
<th>Number of Nodes</th>
<th>916,944</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Edges</td>
<td>4,101,567</td>
</tr>
<tr>
<td>Average in-degree (std)</td>
<td>4.47 (9.2)</td>
</tr>
<tr>
<td>Average out-degree (std)</td>
<td>4.47 (1.1)</td>
</tr>
<tr>
<td>Density</td>
<td>$4.87 \times 10^{-6}$</td>
</tr>
<tr>
<td>Fraction of reciprocal links</td>
<td>0.32</td>
</tr>
<tr>
<td>Clustering Coefficient</td>
<td>0.35</td>
</tr>
</tbody>
</table>
Some assumptions we make in the empirical analysis

- We take the network “as is”
- Demand computed based on SalesRank
- We use the UPD (“Ultimate Purchase Data”) network as a proxy for the RCR ($\alpha$)
The "UPD" network in Amazon

What Other Items Do Customers Buy After Viewing This Item?

- **Linked: How Everything Is Connected to Everything Else and What It Means**
  - by Albert-Laszlo Barabasi
  - Paperback
  - ★★★★★☆ (116)
  - $10.88

- **Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives**
  - by James H. Fowler
  - Audio CD
  - ★★★★★☆ (41)
  - $15.60

- **Everything Is Obvious: *Once You Know the Answer***
  - by Duncan J. Watts
  - Hardcover
  - ★★★★★☆ (51)
  - $17.16

- **Sync: How Order Emerges From Chaos In the Universe, Nature, and Daily Life**
  - by Steven H. Strogatz
  - Paperback
  - ★★★★★☆ (66)
  - $11.43
Some basic value numbers for the Amazon network

<table>
<thead>
<tr>
<th>Average Revenue  [$]</th>
<th>Average Network Value [$]</th>
<th>Average Incoming Value [$]</th>
<th>Average Intrinsic value [$]</th>
<th>Average Outgoing Value [$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>35.92</td>
<td>35.92</td>
<td>7.73</td>
<td>28.2</td>
<td>7.73</td>
</tr>
</tbody>
</table>
We want to examine the relationship between network value and revenue tier

There is considerable literature about the extent, antecedents and importance of the “long tail”


However - the inter-product effect on value has not been examined
## Value type results for different revenue tiers

<table>
<thead>
<tr>
<th>Revenue Percentile</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Revenue [$]</td>
<td>Incoming Value [$]</td>
<td>Intrinsic value [$]</td>
<td>Outgoing Value [$]</td>
<td>Network Value (3+4) [$]</td>
<td>Net Influence (4-2) [$]</td>
<td>Relative Net Influence (4-2)/1 [$]</td>
</tr>
<tr>
<td>0-20% Low sellers</td>
<td>6.21</td>
<td>0.83</td>
<td>5.38</td>
<td>3.03</td>
<td>8.40</td>
<td>2.2</td>
<td>35.38%</td>
</tr>
<tr>
<td>20-40%</td>
<td>11.75</td>
<td>2.11</td>
<td>9.65</td>
<td>4.26</td>
<td>13.91</td>
<td>2.15</td>
<td>18.33%</td>
</tr>
<tr>
<td>40-60%</td>
<td>18.15</td>
<td>3.63</td>
<td>14.51</td>
<td>5.49</td>
<td>20.00</td>
<td>1.86</td>
<td>10.21%</td>
</tr>
<tr>
<td>60-80%</td>
<td>30.43</td>
<td>6.34</td>
<td>24.09</td>
<td>8.14</td>
<td>32.23</td>
<td>1.8</td>
<td>5.90%</td>
</tr>
<tr>
<td>80-100% High sellers</td>
<td>113.08</td>
<td>25.72</td>
<td>87.36</td>
<td>17.72</td>
<td>105.08</td>
<td>-8</td>
<td>-7.07%</td>
</tr>
</tbody>
</table>

- Most of the value is intrinsic for all tiers
### Value type results for different revenue tiers

<table>
<thead>
<tr>
<th>Revenue Percentile</th>
<th>1 (Revenue [$])</th>
<th>2 (Incoming Value [$])</th>
<th>3 (Intrinsic value [$])</th>
<th>4 (Outgoing Value [$])</th>
<th>5 (Network Value (3+4) [$])</th>
<th>6 (Net Influence (4-2) [$])</th>
<th>7 (Relative Net Influence (4-2)/1 [$])</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low sellers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-20%</td>
<td>6.21</td>
<td>0.83</td>
<td>5.38</td>
<td>3.03</td>
<td>8.40</td>
<td>2.2</td>
<td>35.38%</td>
</tr>
<tr>
<td>20-40%</td>
<td>11.75</td>
<td>2.11</td>
<td>9.65</td>
<td>4.26</td>
<td>13.91</td>
<td>2.15</td>
<td>18.33%</td>
</tr>
<tr>
<td>40-60%</td>
<td>18.15</td>
<td>3.63</td>
<td>14.51</td>
<td>5.49</td>
<td>20.00</td>
<td>1.86</td>
<td>10.21%</td>
</tr>
<tr>
<td>60-80%</td>
<td>30.43</td>
<td>6.34</td>
<td>24.09</td>
<td>8.14</td>
<td>32.23</td>
<td>1.8</td>
<td>5.90%</td>
</tr>
<tr>
<td><strong>High sellers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80-100%</td>
<td>113.08</td>
<td>25.72</td>
<td>87.36</td>
<td>17.72</td>
<td>105.08</td>
<td>-8</td>
<td>-7.07%</td>
</tr>
</tbody>
</table>

- Most of the value is intrinsic for all tiers
- There are different effects at different book revenue tiers
  - The **Outgoing Value** is relatively higher for bestsellers
  - So is the **Incoming Value**
  - For low sellers the proportion of outgoing value is higher (12% vs. 8% for best sellers)
### Value type results for different revenue tiers

<table>
<thead>
<tr>
<th>Revenue Percentile</th>
<th>Revenue [$]</th>
<th>Incoming Value [$]</th>
<th>Intrinsic value [$]</th>
<th>Outgoing Value [$]</th>
<th>Network Value (3+4) [$]</th>
<th>Net Influence (4-2) [$]</th>
<th>Relative Net Influence (4-2)/1 [$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-20% Low sellers</td>
<td>6.21</td>
<td>0.83</td>
<td>5.38</td>
<td>3.03</td>
<td>8.40</td>
<td>2.2</td>
<td>35.38%</td>
</tr>
<tr>
<td>20-40%</td>
<td>11.75</td>
<td>2.11</td>
<td>9.65</td>
<td>4.26</td>
<td>13.91</td>
<td>2.15</td>
<td>18.33%</td>
</tr>
<tr>
<td>40-60%</td>
<td>18.15</td>
<td>3.63</td>
<td>14.51</td>
<td>5.49</td>
<td>20.00</td>
<td>1.86</td>
<td>10.21%</td>
</tr>
<tr>
<td>60-80%</td>
<td>30.43</td>
<td>6.34</td>
<td>24.09</td>
<td>8.14</td>
<td>32.23</td>
<td>1.8</td>
<td>5.90%</td>
</tr>
<tr>
<td>80-100% High sellers</td>
<td>113.08</td>
<td>25.72</td>
<td>87.36</td>
<td>17.72</td>
<td>105.08</td>
<td>-8</td>
<td>-7.07%</td>
</tr>
</tbody>
</table>

- Most of the value is intrinsic for all tiers
- There are different effects at different book revenue tiers
  - The relative proportion of network value out of the overall value for low sellers is much higher than for bestsellers
What is the story?
In-degree (recommending books) statistics for different revenue tiers at

<table>
<thead>
<tr>
<th>Revenue Percentile</th>
<th>1</th>
<th>2</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Indegree</td>
<td>Average units sold by the incoming link's books</td>
<td>Fraction of incoming links from books of the same revenue tier</td>
</tr>
<tr>
<td>0-20%</td>
<td>2.48</td>
<td>0.874</td>
<td>0.58</td>
</tr>
<tr>
<td>20-40%</td>
<td>3.24</td>
<td>1.03</td>
<td>0.33</td>
</tr>
<tr>
<td>40-60%</td>
<td>4.18</td>
<td>1.16</td>
<td>0.28</td>
</tr>
<tr>
<td>60-80%</td>
<td>5.38</td>
<td>1.33</td>
<td>0.29</td>
</tr>
<tr>
<td>80-100%</td>
<td>10.12</td>
<td>2.16</td>
<td>0.44</td>
</tr>
</tbody>
</table>
Connectivity in product networks

First Issue

- In many product networks out-degree is constant limited to a few items
  - ~5 at the time in Amazon
  - But in-degree can be very high!

- We see that the in-degree is considerably higher for best sellers
  - Which leads to a higher *Incoming Value*
### In-degree (recommending books) statistics for different revenue tiers

<table>
<thead>
<tr>
<th>Revenue Percentile</th>
<th>1</th>
<th>2</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Indegree</td>
<td>Average units sold by the incoming link's books</td>
<td>Fraction of incoming links from books of the same revenue tier</td>
</tr>
<tr>
<td>0-20%</td>
<td>2.48</td>
<td>0.874</td>
<td>0.58</td>
</tr>
<tr>
<td>20-40%</td>
<td>3.24</td>
<td>1.03</td>
<td>0.33</td>
</tr>
<tr>
<td>40-60%</td>
<td>4.18</td>
<td>1.16</td>
<td>0.28</td>
</tr>
<tr>
<td>60-80%</td>
<td>5.38</td>
<td>1.33</td>
<td>0.29</td>
</tr>
<tr>
<td>80-100%</td>
<td>10.12</td>
<td>2.16</td>
<td>0.44</td>
</tr>
</tbody>
</table>
Second issue

- High selling books are connected more to other high selling books
  - Consistent with *assortative mixing* in social networks (Newman 2003)

- This will also lead to a higher *Incoming Value*

Overall

- While the *Outgoing Value* best sellers provide to others is higher than that of low selling items, the *Incoming Value* they get is even higher compared to low selling books.

- This will create a “deficit” in the relative network value.
## Average RCR among revenue tiers

<table>
<thead>
<tr>
<th>From Tier</th>
<th>0-20%</th>
<th>20-40%</th>
<th>40-60%</th>
<th>60-80%</th>
<th>80-100%</th>
<th>Average outgoing RCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low sellers</td>
<td>2.43%</td>
<td>3.52%</td>
<td>3.85%</td>
<td>3.87%</td>
<td>3.37%</td>
<td>3.41%</td>
</tr>
<tr>
<td>20-40%</td>
<td>2.16%</td>
<td>2.99%</td>
<td>3.44%</td>
<td>3.60%</td>
<td>3.44%</td>
<td>3.13%</td>
</tr>
<tr>
<td>40-60%</td>
<td>1.97%</td>
<td>2.49%</td>
<td>2.91%</td>
<td>3.27%</td>
<td>3.51%</td>
<td>2.83%</td>
</tr>
<tr>
<td>60-80%</td>
<td>1.83%</td>
<td>2.17%</td>
<td>2.51%</td>
<td>2.94%</td>
<td>3.49%</td>
<td>2.59%</td>
</tr>
<tr>
<td>80-100% Bestsellers</td>
<td>1.48%</td>
<td>1.76%</td>
<td>1.96%</td>
<td>2.23%</td>
<td>3.01%</td>
<td>2.09%</td>
</tr>
<tr>
<td>Average incoming RCR</td>
<td>1.97%</td>
<td>1.76%</td>
<td>1.96%</td>
<td>2.23%</td>
<td>3.01%</td>
<td></td>
</tr>
</tbody>
</table>
Third issue

- The conversion rate of the incoming links for high selling books is higher
- While the conversion rate of the outgoing links for high selling books is lower

Bestsellers absorb demand!
Third issue

- The conversion rate of the incoming links for high selling books is higher
- While the conversion rate of the outgoing links for high selling books is lower

**Bestsellers absorb demand!**

By the way…. we repeated our analysis with a **simulated RCR**. We experimented with:

- Normally distributed RCR with randomly chosen averages
- Normally distributed RCR with higher averages for high selling products
- Normally distributed RCR with lower averages for high selling products
Some final thoughts

- Products may have an indirect value
- It can be assessed
- Best sellers may “enjoy” the network more
- Additional support for the value of the long tail
- A lot can be still done….
Questions & Suggestions?

Gal Oestreicher-Singer
galos@post.tau.ac.il