The dynamics of viral marketing

Lada Adamic
School of Information
U. Michigan, Ann Arbor

joint work with Jure Leskovec (CMU)
and Bernardo Huberman (HP Labs)

NetSci
May 22nd, 2006
The dynamics of Viral Marketing

Outline

- prior work on viral marketing & information diffusion
- incentivised viral marketing
- cascades and stars
- network effects
- product and social network characteristics
Information diffusion

- Studies of innovation adoption
  - hybrid corn (Ryan and Gross, 1943)
  - prescription drugs (Coleman et al. 1957)
  - poison pills and golden parachutes (Davis and Greeve, 1997)

- Models (very many)
  - Rogers, ‘Diffusion of Innovations’
  - Watts, information cascades, 2002
Using online networks for viral marketing

- Burger King’s subservient chicken
Motivation for viral marketing

- viral marketing successfully utilizes social networks for adoption of some services
  - hotmail gains 18 million users in 12 months, spending only $50,000 on traditional advertising
  - gmail rapidly gains users although referrals are the only way to sign up

- customers becoming less susceptible to mass marketing
- mass marketing impractical for unprecedented variety of products online
  - Google AdSense helps sellers reach buyers with targeted advertising
  - but how do buyers get good recommendations?
Marketing in the long tail

The web savvy consumer and personalized recommendations

- > 50% of people do research online before purchasing electronics

- personalized recommendations based on prior purchase patterns and ratings
  - Amazon, “people who bought x also bought y”
  - MovieLens, “based on ratings of users like you…”
  - Epinions, “based on the opinions of the raters you trust” (Richardson & Domingos, 2002)

- ratings have been shown to affect the likelihood of an item being bought
  - Resnick & Zeckhauser, 2001: eBay
  - Judith Chevalier and Dina Mayzlin, 2004: Amazon and BN
Is there still room for viral marketing next to personalized recommendations?

- We are more influenced by our friends than strangers

- 68% of consumers consult friends and family before purchasing home electronics
  (Burke 2003)
Incentivised viral marketing

- Senders and followers of recommendations receive discounts on products

  10% credit

  10% off

- Recommendations are made to any number of people at the time of purchase
- Only the recipient who buys first gets a discount
Product recommendation network

- purchase following a recommendation
- customer recommending a product
- customer not buying a recommended product
the data

- large online retailer (June 2001 to May 2003)

- 15,646,121 recommendations
- 3,943,084 distinct customers
- 548,523 products recommended
- 99% of them belonging 4 main product groups:
  - books
  - DVDs
  - music
  - VHS
data attributes

- recommendations
  - sender (shadowed)
  - recipient (shadowed)
  - recommendation time
  - buy bit
  - purchase time
  - product price

- additional product info
  - categories
  - reviews
  - ratings
### Summary Statistics by Product Group

<table>
<thead>
<tr>
<th></th>
<th>products</th>
<th>customers</th>
<th>recommendations</th>
<th>edges</th>
<th>buy bits</th>
<th>buy edges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>103,161</td>
<td>2,863,977</td>
<td>5,741,611</td>
<td>2,097,809</td>
<td>65,344</td>
<td>17,769</td>
</tr>
<tr>
<td>DVD</td>
<td>19,829</td>
<td>805,285</td>
<td>8,180,393</td>
<td>962,341</td>
<td>17,232</td>
<td>58,189</td>
</tr>
<tr>
<td>Music</td>
<td>393,598</td>
<td>794,148</td>
<td>1,443,847</td>
<td>585,738</td>
<td>7,837</td>
<td>2,739</td>
</tr>
<tr>
<td>Video</td>
<td>26,131</td>
<td>239,583</td>
<td>280,270</td>
<td>160,683</td>
<td>909</td>
<td>467</td>
</tr>
<tr>
<td>Full</td>
<td>542,719</td>
<td>3,943,084</td>
<td>15,646,121</td>
<td>3,153,676</td>
<td>91,322</td>
<td>79,164</td>
</tr>
</tbody>
</table>
observations on product groups

- There are relatively few DVD titles, but DVDs account for ~ 50% of recommendations.

- recommendations per person
  - DVD: 10
  - books and music: 2
  - VHS: 1

- recommendations per purchase
  - books: 69
  - DVDs: 108
  - music: 136
  - VHS: 203

- Music recommendations reached about the same number of people as DVDs but used only 1/5 as many recommendations

- Very small number of unique edges for books, videos, music.
  - smaller than the number of nodes
  - the networks are highly disconnected
adoption of viral marketing program & social network connectivity

size of giant component

number of nodes

# nodes

1.7*10^5 m

by month

quadratic fit
viral marketing program not spreading virally

- 94% of users make first recommendation without having received one previously
- linear growth: ~ 165,000 new users added each month
- size of giant connected component increases from 1% to 2.5% of the network (100,420 users) – small!
- some subcommunities are better connected
  - 24% out of 18,000 users for westerns on DVD
  - 26% of 25,000 for classics on DVD
  - 19% of 47,000 for anime (Japanese animated film) on DVD
- others are just as disconnected
  - 3% of 180,000 home and gardening
  - 2-7% for children’s and fitness DVDs
Recommendation network between customers

1 million senders of book recommendations did not receive any
1.8 million recipients of book recommendations did not send any
medical study guide recommendation network
identifying viral marketing cascades

\[ t_1 < t_2 < \ldots < t_n \]

**Legend**
- Red circle: bought but didn’t receive a discount
- Blue circle: bought and received a discount
- Orange circle: received a recommendation but didn’t buy
measuring cascade sizes

- delete late recommendations
- count how many people are in a single cascade
- exclude nodes that did not buy

\[ n = 1.8 \times 10^6 \times x^{-4.98} \]

steep drop-off

very few large cascades
DVD cascades can grow large
possibly as a result of websites where people sign up to exchange recommendations

- shallow drop off – fat tail

~ $x^{-1.56}$

a number of large cascades
simple model of propagating recommendations
(ignoring for the moment the specific mechanics of the recommendation program of the retailer)

- Each individual will have $p_t$ successful recommendations. We model $p_t$ as a random variable.

- At time $t+1$, the total number of people in the cascade, $N_{t+1} = N_t \times (1+p_t)$

- Subtracting from both sides, and dividing by $N_t$, we have

$$\frac{N_{t+1} - N_t}{N_t} = p_t$$
simple model of propagating recommendations
(continued)

- Summing over long time periods

\[ \frac{dN}{N} = \sum p_t \]

- The right hand side is a sum of random variables and hence normally distributed.

- Integrating both sides, we find that \( N \) is lognormally distributed

\[
P(N) = \frac{1}{N \sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln(N) - \mu)^2}{2\sigma^2}\right)
\]
For high variance in # of recommendations sent

\[ \ln(P(N)) = - \ln(N) - \ln(\sqrt{2\pi\sigma^2}) - \frac{(\ln(N) - \mu)^2}{2\sigma^2} \]

- The lognormal behaves like a power-law with exponent 1
- We observe fat tails in cascade sizes

small if \( \sigma \) large
participation level by individual

Number of recommendations

Count

\[ \text{Count} = 3.4 \times 10^6 \times x^{-2.3} \]

\[ R^2 = 0.96 \]

very high variance
The most active person made 83,729 recommendations and purchased 4,416 different items!
Network effects
does receiving more recommendations increase the likelihood of buying?

![Graph showing the probability of buying for books and DVDs with increasing incoming recommendations.](image)
does sending more recommendations influence more purchases?

BOOKS

DVDs

Outgoing Recommendations

Number of Purchases

Outgoing Recommendations

Number of Purchases
the probability that the sender gets a credit with increasing numbers of recommendations

- consider whether sender has at least one successful recommendation
- controls for sender getting credit for purchase that resulted from others recommending the same product to the same person

![Graph showing the probability of credit levels off for DVDs](image-url)
Multiple recommendations between two individuals weaken the impact of the bond on purchases.

BOOKS

DVDs

Probability of buying

Exchanged recommendations

Probability of buying

Exchanged recommendations
product and social network characteristics influencing recommendation effectiveness
telling the world
vs. telling your friends

- consider
  - # reviewers per book
  - # recommenders per book

- what we observe (ratio of recommendations to reviews given in parentheses)
  - tell the world but not your friends
    - literature & fiction (0.57)
    - mystery & thrillers (0.36)
    - horror (0.44)
  - tell the world and your friends
    - biographies (0.90)
    - children’s books (1.12)
    - religion (1.73)
    - history (1.27)
    - nonfiction (1.89)
  - tell just your friends about personal pursuits
    - health, mind & body (2.39)
    - home & garden (3.48)
    - arts & photography (3.85)
    - cooking, food & wine (3.49)
  - tell your colleagues about professional interests
    - professional & technical (3.22)
    - computers & internet (3.10)
    - medicine (4.19)
    - engineering (3.85)
    - law (4.25)
recommendation success by book category

- consider successful recommendations in terms of
  - av. # senders of recommendations per book category
  - av. # of recommendations accepted
- books overall have a 3% success rate
  - (2% with discount, 1% without)
- lower than average success rate (statistically significant at p=0.01 level)
  - fiction
    - romance (1.78), horror (1.81)
    - teen (1.94), children’s books (2.06)
    - comics (2.30), sci-fi (2.34), mystery and thrillers (2.40)
  - nonfiction
    - sports (2.26)
    - home & garden (2.26)
    - travel (2.39)
- higher than average success rate (statistically significant)
  - professional & technical
    - medicine (5.68)
    - professional & technical (4.54)
    - engineering (4.10), science (3.90), computers & internet (3.61)
    - law (3.66)
    - business & investing (3.62)
  - nonfiction – general (3.28)
- in the middle
  - literature & fiction (2.82)
  - religion & spirituality (3.13)
  - outdoors and nature (2.38)
professional and organized contexts

- In general, professional & technical book recommendations are more often accepted (probably in part due to book cost)
- Some organized contexts other than professional also have higher success rate, e.g. religion
  - overall success rate 3.13%
  - Christian themed books
    - Christian living and theology (4.7%)
    - Bibles (4.8%)
  - not-as-organized religion
    - new age (2.5%)
    - occult spirituality (2.2%)
- Well organized hobbies
  - books on orchids recommended successfully twice as often as books on tomato growing
Anime

- 47,000 customers responsible for the 2.5 out of 16 million recommendations in the system

- 29% success rate per recommender of an anime DVD

- Giant component covers 19% of the nodes

- Overall, recommendations for DVDs are more likely to result in a purchase (7%), but the anime community stands out
### Regressing on Product Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Transformation</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td></td>
<td>-0.940 ***</td>
</tr>
<tr>
<td># recommendations</td>
<td>ln(r)</td>
<td>0.426 ***</td>
</tr>
<tr>
<td># senders</td>
<td>ln(n_s)</td>
<td>-0.782 ***</td>
</tr>
<tr>
<td># recipients</td>
<td>ln(n_r)</td>
<td>-1.307 ***</td>
</tr>
<tr>
<td>product price</td>
<td>ln(p)</td>
<td>0.128 ***</td>
</tr>
<tr>
<td># reviews</td>
<td>ln(v)</td>
<td>-0.011 ***</td>
</tr>
<tr>
<td>avg. rating</td>
<td>ln(t)</td>
<td>-0.027 *</td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td>0.74</td>
</tr>
</tbody>
</table>

Significance at the 0.01 (***) , 0.05 (**) and 0.1 (*) levels
products most suited to viral marketing

- small community
  - few reviews, senders, and recipients
  - but sending more recommendations helps

- pricey products

- rating doesn’t play as much of a role
Conclusions

Overall
- incentivized viral marketing contributes marginally to total sales
- occasionally large cascades occur, most often for DVDs

Observations for future diffusion models
- purchase decision more complex than threshold or simple infection
- influence saturates as the number of contacts expands
- links lose effectiveness if overused

Conditions for successful recommendations
- professional and organizational contexts
- discounts on expensive items
- small, tightly knit communities
For more information

- short version of the paper:

- my publications:
  - http://www-personal.umich.edu/~ladamic

- Jure’s publications:
  - http://www.cs.cmu.edu/~jure/pubs/

- Bernardo Huberman’s Information Dynamics Lab at HP:
  - http://www.hpl.hp.com/research/idl
Extras
**pay it forward**

<table>
<thead>
<tr>
<th>product category</th>
<th>number of buy bits</th>
<th>forward recommendations</th>
<th>percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>65,391</td>
<td>15,769</td>
<td>24.2</td>
</tr>
<tr>
<td>DVD</td>
<td>16,459</td>
<td>7,336</td>
<td>44.6</td>
</tr>
<tr>
<td>Music</td>
<td>7,843</td>
<td>1,824</td>
<td>23.3</td>
</tr>
<tr>
<td>Video</td>
<td>909</td>
<td>250</td>
<td>27.6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>90,602</strong></td>
<td><strong>25,179</strong></td>
<td><strong>27.8</strong></td>
</tr>
</tbody>
</table>
when recommendations are sent
when purchases are made
when discounts are to be had
lag between time of recommendation and time of purchase

40% of those who buy buy within a day
but > 15% wait more than a week
observations

- purchases and recommendations follow a daily cycle
- customers are most likely to purchase within a day of receiving a recommendation
- acting on a recommendation at atypical times increases the likelihood of receiving a discount