Mapping communities in large virtual social networks

Using Twitter data to find the Indie Mac community

Michiel van Meeteren, Ate Poorthuis, Elenna Dugundji
University of Amsterdam
Outline

• Background on Indies and their (online) social networks
• Research goals and methods
• Research findings
• Implications for ethics and marketing
Indies

- “Independent” Macintosh (and iPhone) software companies
- Virtual cluster/community of practice
- Early Twitter adopters
- Culture/ethos on “quality,” “proper” behavior, competition and information sharing
- Informally bounded, no formal organization
Use of Social Software

- Identification and socialization
- Satisfaction of information needs: “Virtual watercooler”
- Marketing
Research Goals

• “Map” the virtual social network of the Indie community starting from an acknowledged “tastemaker” in the community

• Explore the possibilities of research on large and noisy Twitter datasets

• Analyze the importance of social position for the diffusion of products within the network (in progress)
John Gruber

- Professional independent tech blogger (www.daringfireball.net)
- Widely respected in the Indie community
- 150,000 RSS feed subscribers
- 1.3 million page-views a week
- On Twitter: Following 244, Followed by 32,718 (Sept 20, 2009)
Findings I:

Word clouds of Gruber network and subcommunities

Total Gruber network
N=51501

Community 1:
N= 13978

Community 2:
N=11522

Community 3:
N= 591

Community 4:
N= 1428

Community 5: Indie community
N= 12669

van Meeteren, Poorthuis, Dugundji

EDF 2009
Findings II:

Frequencies of occupations of 25 highest ranking nodes in the Indie subcommunity

- Developer
- Apple
- Journalist
- Other
Utility for marketing

- Defining subcommunities form finely grained target audiences
- Market research tool
- Niche (direct) advertising tool
Potential ethical issues

- Finds out characteristics of people which they do not necessarily know about themselves and which they did not publicly disclose.

- When refined these methods could potentially jeopardize (political) activism using Twitter as a medium (e.g. Iran elections).
Conclusions

• Analyzing large and noisy Twitter data results in clear, delineated subcommunities

• Network structure of subcommunities relevant for “tastemaking”

• Opens marketing possibilities and moral hazards
More information

• Ate Poorthuis  
  (atepoorthuis@gmail.com)

• Michiel van Meeteren  
  (michielvanmeeteren@gmail.com)

• Elenna Dugundji  
  (e.r.dugundji@gmail.com)