EVA: Extraction, Visualization and Analysis of the Telecommunications and Media Ownership Network

Kim Norlen, Gabriel Lucas, Michael Gebbie, John Chuang

School of Information Management & Systems
UC Berkeley
http://denali.berkeley.edu/eva/
Ownership Network

- Directed social network:
  - Nodes: firms
  - Edges: equity possession

- Flow of capital, information, and control

- Industrial organization
  - horizontal merger
  - vertical integration

Portion of telecom ownership network
Corporate Transparency

- Essential for public discourse concerning anti-trust, regulation, investor confidence
- SEC requires full disclosure of mergers, acquisitions and other relationships...
- ... but companies have little incentive to do so (until now)
- Research challenge:
  - automated construction of telecom/media/IT ownership network dataset from publicly accessible documents
EVA System

Information Extraction

- Data Sources:
  - U.S. SEC 10-K documents (Corporate Annual Reports)
  - Industry Standard Deal Tracker Database
  - Columbia Journalism Review “Who owns what”

- Dataset: 8,343 companies, 6,726 relationships
EVA System

- Information Extraction
- Information Visualization
EVA System

- Information Extraction
- Information Visualization
- Network Analysis
Extraction from SEC 10-Ks

3,374 10-K docs (1.5GB free text)

Company name dictionary & thesaurus (8,343)

Keyword-based retrieval
Noise Filters
Probabilistic Ranking
Human review

Precision

364,581 paragraphs (<5%)
14,759 paragraphs (20%)
3,249 paragraphs (55.4%)*
652 relationships (100%)

* State-of-art IE systems achieve 50-70% precision for entity event finding.
Tough Case #1: Ambiguity

SEC 10-K document filed by Aether Systems Inc, for year ending Dec 31 2000:

“In connection with the acquisitions of Cerulean, Sinope, RTS and Motient, the Company has accrued $29,800 as of December 31, 2000 for the remaining portion of the purchase price…”

Later in the same document:

“… On November 30, 2000, we acquired Motient's retail transportation business unit for $49.2 million in cash…”
Tough Case #2: Directionality

- SEC 10-K document filed by Nextel Communications, for year ending Dec 31 1998:

  "... we **acquired** all of Motorola’s 800 MHz SMR licenses in the continental United States in exchange for 41.7 million shares of class A common stock and 17.8 million shares of nonvoting class B common stock."

- Who acquired whom?
**Visualization**
Network Analysis: Key Findings

- 6,726 relationships between 7,253 companies
  - additional 1,090 companies with no relationships
- Node degree and component size both follow power law distribution:
  - Top ten companies are parents for 24% of relationships
  - Largest component: 4400+ firms (53.6% of network)
Network Analysis: Key Findings

- The largest bi-component contains 234 companies and includes many competitors
  - AT&T, MCI WorldCom
  - British Telecom, Deutsche Telecom
  - AOL-Time Warner, Comcast
  - Bertelsmann, Yahoo!
  - CBS, NBC, Disney (ABC)
  - Cisco, Intel, Microsoft, Sony

- Various network prominence metrics:
  - Degree: Clear Channel Communications, Liberty Publishing
  - “Freeman” Betweenness: Liberty Media, Time Warner
  - Depth/Radius: Comcast
  - Cutpoints: Clear Channel, Time Warner
  - Cliques: Liberty Media, AT&T
  - Ego: Liberty Media, Comcast
Conclusion

- EVA uses:
  - Information extraction and visualization techniques to gather and present corporate ownership relationships from heterogeneous data sources
  - Social network analysis techniques to identify prominent firms and reveal industry structure
- EVA helps:
  - Regulators, policy researchers, investors, and general public by bringing greater transparency to public disclosure of corporate inter-relationships