Using Hyperlink Network Analysis

An overview of
Intermedia Agenda Building of the Blogosphere:
Public Relations in The Network

Adam Saffer
Ph.D. Student
University of Oklahoma
Adam.Saffer@gmail.com

Presentation Overview

- ❑ Overview of hyperlink network analysis

- ❑ Steps for using hyperlink network analysis

- ❑ Steps for using content analysis

- ❑ Background on research

- ❑ Methods used in research

- ❑ Findings of research

Overview of HNA

- Hyperlink network analysis (HNA) is closely related to social network analysis.

- HNA studies the relationships between actors. Actors can be people, organizations, or websites/blogs.

- Hyperlinks create a network of connections that are sent to (outdegree) and received (indegree) by blogs.

- Blogs that receive a high number of links are centrally located and have content that other blogs direct their readers to visit.

- Centrality in network analysis is synonymous to influence.

Steps for HNA

- Identify the blogs of interest. This will be your "seed sites" list.
  - You could use a list of blogs that often mention your client.
  - Keep in mind the more blogs, the slower the web crawl.

- Use a web crawler to mine the hyperlink data on the blog websites.
  - Options for web crawlers include: SocSicBot or VOSON.
  - Download data into a network matrix form for analysis.

- Analyze network matrix data to identify most central blogs.
  - Options for network analysis software include: NodeXL and UCINET.

- See Ackland (2011) for further insights into hyperlink network analysis.

Steps for Content Analysis

- Identify time frame for studying blog posts.

- Gather blog posts from blogs in the network.

- Code for date, topic, and sources.
  - Practitioners could code for talking points from pitches.
  - Also, some might find it helpful to code for organizational spokespeople when appropriate.

- Correlate codes between blogs.

- Analyze the relationships between correlations and blogs' centrality.

Background on Research

- The research is part of the Ketchum Excellence in Public Relations Research Award.

- The idea came while working on a project for a client at Ketchum that wanted to know the most important blogs for pitching stories.

- The study focused on three topic-specific blog networks: environmental, financial, and information gadgets.*

- Topic-specific blogs create a network by sending hyperlinks between blogs.

*A fourth network—TV show reviews—was removed after the hyperlink analysis for lack of connections between blogs.

Background on Research (cont.)

- The blog networks can be studied using hyperlink network analysis.

- The study used intermedia agenda building theory as a framework (Denham, 2010; Ragas & Kiousis, 2010; Reese & Danielian, 1989), which stems from agenda setting theory.

- The theory suggests media outlets (blogs) set each others agenda. Media outlets look to see what their competitors are covering and will sometimes cover the same topics.

- Public relations practitioners play a role in setting the agenda by pitching stories to these outlets.

- The study was interested in studying how content moved through blog networks.

Background on Research (cont.)

- Research Questions:
  - RQ1: What are the structural characteristics of the blog networks?
    - The network structures were assessed with the measures *of density, geodesic distance,* and *overall centrality scores.*
    - The question sought to study how the blogs connect to one another, which is an indication for how stories might be shared.
  - RQ2: How does the structure of a blog network affect how content is shared?
    - Blog posts were analyzed to identify which blogs shared topics and sources.
    - The content analysis results were paired with the hyperlink network. analysis results.

Methods Used in Research

- Hyperlink Network Analysis
  - Top 50 blogs as identified by Technorati for each topic-specific blog network were used as "seed sites" for the web crawl.
  - Web crawler VOSON mined the hyperlink data from the blogs.
  - The web crawl added blogs not on the top 50 blog list because the top 50 blogs linked to unlisted blogs.
  - Data were transformed into a network matrix.
  - NodeXL was used to analyze the data and identify the most central actors.

Methods Used in Research

- Content Analysis of Blog Posts
  - Two weeks of blog posts from the 20 most central blogs in each of the three blog networks were content analyzed.
  - The unit of analysis was the blog posts.
  - Blog posts were coded for date, general topic based on the headline and lead paragraph, and sources used.
  - The topics and sources were correlated.

# Findings of Research

- Results of RQ1:
  - The three networks had low overall connections.
  - Financial blog network was the most well connected (dense) with the least distance across the network (geodesic distance) and greatest average betweenness centrality.

- Results of RQ2:
  - Blogs in the financial network shared the most topics in common with other blogs.
  - Across the three networks, the most central blogs shared more content and sources in common with other blogs.

Findings of Research

Financial Blog Network—5 Most Central Blogs
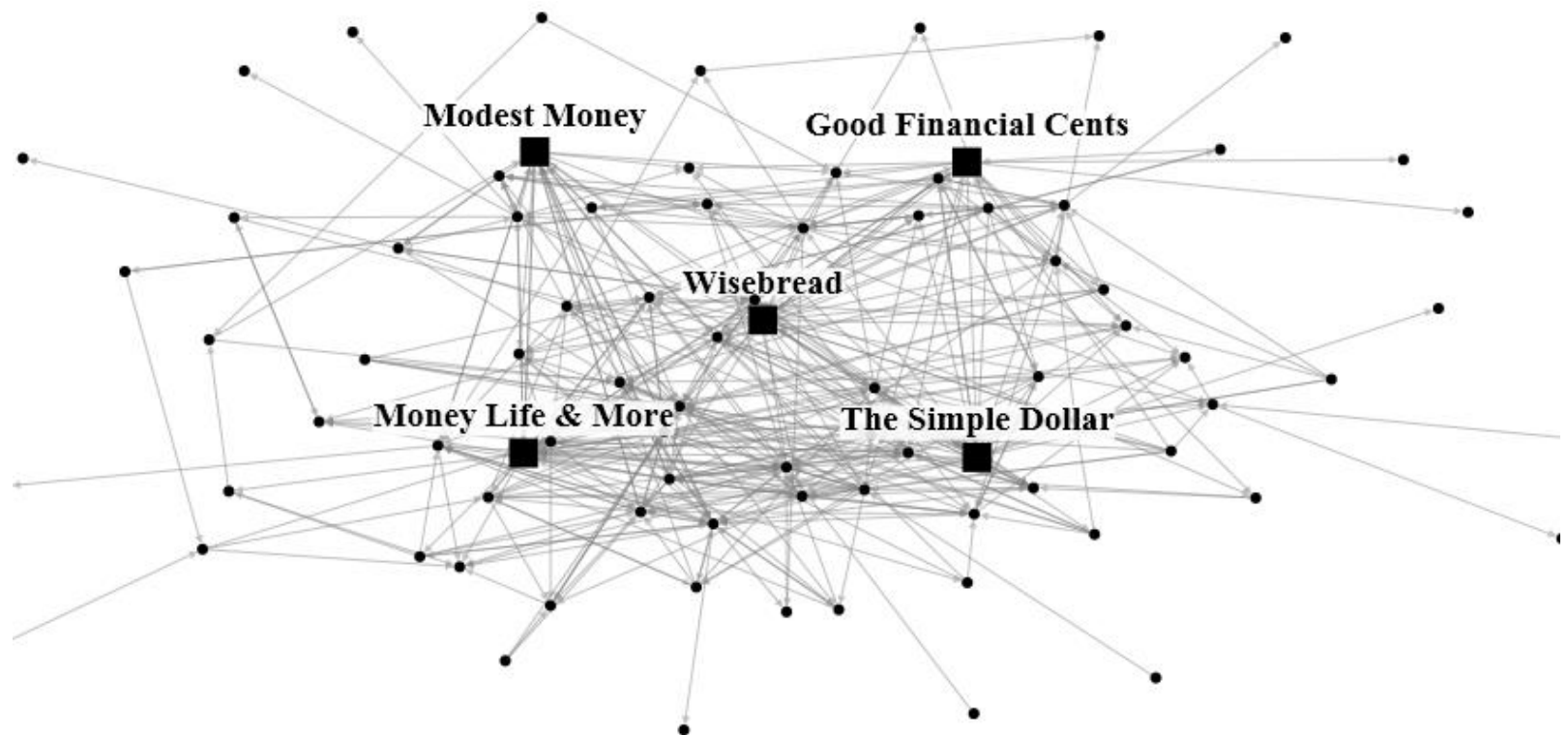
Implications for PR Practice

- Implication #1: (Media Relations Practitioners)
  - The most central blogs in a network are the most ideal blogs to target pitch stories.
  - If a practitioner wants a story to spread in a blog network, the study suggests that the most central blogs are most effective.
  - Practitioners should keep in mind that blog hyperlink networks are dynamic and can change over time with new blogs gaining popularity from other blogs sending links to them.
  - The structure of the network should be check regularly.

Implications for PR Practice

- Implication #2: (Media Relations Practitioners)
  - Hyperlink network analysis is a powerful tool for understanding blog networks.
  - Practitioners should use hyperlink network analysis to familiarize themselves with blogs by seeing which blogs point to other blogs.
  - If a blog is receiving a high number of links from other blogs, practitioners might consider also pitching to the other blogs.

Implications for PR Practice

- Implication #3: (Research Practitioners)
    - Network measures must be integrated into the measurement of influence. The number of followers, likes, and comments or retweets are invalid.
    - Network analysis offers a more accurate approach for determining a blog's influence because the method examines who links to who to determine who is the most influential.
    - The network measures are based on graph theory mathematics and encompass the relationships an actor has and the relationships of other actors in the network.
    - HNA offers a more holistic view of who is influential online.

References

- Ackland, R. (2011). WWW Hyperlink Networks. In D. Hansen, B. Shneiderman and M. Smith (Eds.), *Analyzing social media networks with NodeXL: Insights from a connected world* (pp. 181–200). Burlington, MA: Morgan-Kaufmann.

- Denham, B. E. (2010). Toward conceptual consistency in studies of agenda-building processes: A scholarly review. *The Review of Communication*, *10*(4), 306–323.

- Ragas, M. W., & Kiousis, S. (2010). Intermedia agenda-setting and political activism: MoveOn. org and the 2008 presidential election. *Mass Communication and Society*, *13*(5), 560–583.

- Reese, S. D., & Danielian, L. H. (1989). Intermedia influence and the drug issue: Converging on cocaine. In P. J. Shoemaker (Ed.), *Communication campaigns about drugs: Government, media, and the public* (pp. 29–46). Hillsdale, NJ: Lawrence Erlbaum Associates.