

Mark Tremayne

University of Texas at Austin

. Applying Network Theory to the Use of External Links on News Web Sites

Chapter 3 of the book

*Internet Newspapers: Making of a Mainstream Medium*

Lawrence Erlbaum Associates

August 2004

Abstract

This study tested emerging network theory against a sub-sample of the Web: stories on national news Web sites. It found that news Web stories contain links to external sites less frequently than just a few years ago. As each organization builds up its own archive of Web content, this material appears to be favored over content that is off-site.

Key words: hypertext, news flow, Web growth, network, network theory, random graph theory, power law, power law distribution, power curve, **Pareto**'s law, preferred attachment, link, hyperlink, external link, link proportion, hub, context, preferred attachment, blog, bloggers, news topic, New York Times, Washington Post, USA Today, US News and World Report, Time, ABC, CBS, CNN, FOX News, MSNBC.

### **Applying Network Theory to the Use of External Links on News Web Sites**

Compared to traditional media, the Web allows journalists greater flexibility in constructing sophisticated stories, stories that can be consumed in different ways by unique subsets of readers. Hypertext is allowing for stories that are less linear and more

segmented than was possible before. As a result, news stories on the Web are more heavily linked every year (Tremayne, 2004b). The unique features of the Web were identified a decade ago as potential difference-makers in stimulating public dialogue of important issues (Rheingold, 1994; Rosen, 1995). Web news editors influence the flow of user traffic on the broader Web by providing links to certain sites and avoiding others. And site managers encourage particular patterns of use within their sites. News Web sites such as The *Dallas Morning News* and the *Chicago Sun Times* and even, at one point, *National Public Radio*, have posted policies on their sites prohibiting so-called “deep linking” to news stories. They want Web traffic to pass through the homepage where the expensive advertising is displayed.

Beyond trying to influence how a reader gets to a site, the Web editor can guide the reader’s path through the site. Some sites offer links to related material, both internally and externally, and in this way the collective choices of news editors on the Web help to create the infrastructure of public discourse on the Web. This study examines the use of external links, those that lead off a news organizations site and out into the larger Web.

Now comprised of billions of individual pages, the Web is one of the largest networks ever. Emerging theory is beginning to explain how networks grow, how information flows along networks and why some nodes in a network become hubs. In a preliminary study examining the first few years of Web news, the author used network theory to identify trends in the use of external hyperlinks on national news Web sites (Tremayne, 2004a). This chapter will reexamine those findings and test the trends against two more years of Web news data. A database of more than 1,700 stories will be

used to track changes in the use of external links over six years. First, though, an explanation of network theory and how it will be applied in this chapter.

### *Network Theory and the Growth of the Web*

Swiss mathematician Leonhard Euler introduced the study of graphs, or networks, when studying a problem concerning bridges and land masses (Barabasi, 2002). In this case, the network consisted of 4 land masses (nodes) connected by 7 bridges (links). A network is any distribution of similar objects that are interconnected in some way. Networks occur in nature, such as neural networks, or can be man made, such as power distribution grids. An explanation for how commonly occurring networks structures form came more than two hundred years after Euler's discoveries in the work of mathematicians Paul Erdos and Alfred Renyi (Barabasi, 2002; D. Watts, 2003). They proposed random graph theory, the idea that networks are essentially the end result of links being formed between random pairs of nodes. Each link forms independently of all others and clustering of links around some nodes is explained by chance. This explanation went unchallenged as most researchers focused instead on information flow on networks rather than evolution of the network itself.

In the 1960s, Rogers studied the flow of ideas through social systems (Rogers, 1983). This work explored both the individual adoption process and the large group diffusion process. In the latter area, Granovetter made the case that ideas are spread most widely not by persons in the middle of a cluster of acquaintances, but by peripheral members who spend time in multiple social circles (Granovetter, 1973). Milgram (1967) demonstrated what has come to be known as the "six degrees of separation." Although he

never used that phrase, Milgram found that it took almost that many friendship links to connect any two people in the United States.

Granovetter's work suggested that the independence assumption of random graph theory was violated in the case of social networks. His conceptualization of the "strength" of a social tie included how much *time* individuals dedicated to its establishment and maintenance. Because an individual's time is limited, establishment of future ties is dependent, in part, on the existence of current ties. Certainly geography and other factors also limit possible social ties. For this type of network, and many others, theories such as random graph theory could not suffice. An alternative explanation is needed.

Barabasi and Albert used the rapid growth of the Web as a testing ground to construct a theory of network development (Barabasi & Albert, 1999). Each Web page has a number of ties (links) to other Web pages. If random graph theory applied to the World Wide Web, there would be a small number of pages with few or no links, a small number with a huge number of links, and the majority with an average number of links; in other words, a normal distribution. Barabasi and Albert found instead a scale-free power-law distribution, like the one in Figure 1.

Figure 1 goes here

The general formula for a distribution of this type is:  $N(k) \sim k^{-g}$ , where the parameter  $g$  is the degree exponent. This value varies from network to network. For the

Web, they found a curve with a degree exponent of 2.1 for incoming links to a page and an exponent of 2.5 for outgoing links from Web pages.

Figure 1 shows a small number of pages with hundreds or thousands of links and a large number of pages with few links. The ones with hundreds of links or more are called hubs.

Distributions following a power law also conform to Pareto's law, the so-called 80/20 rule. Pareto, a nineteenth-century engineer, observed that approximately 20 percent of the population earns 80 percent of the income. This pattern can be observed in a number of other contexts and in each case a plot of individual cases will look like Figure 1.

Only in the late 1990s did researchers begin understanding why naturally occurring networks follow this pattern (Adamic & Huberman, 2001; Barabasi & Albert, 1999; D. J. Watts, 1999). Using computer simulations Barabasi and Albert discovered two principles that explain all distributions that follow a power law: *growth* and *preferential attachment* (Barabasi & Albert, 1999).

*Web Growth.* Another example from the Web helps illustrate the theory. If we plot Web pages by links pointing to them we will find that some (yahoo.com or ebay.com) are linked to by thousands or even millions of Web pages. But most Web pages have very few other pages that link to them. A plot of this would look very much like Figure 1. So why do the hubs have so many links pointing to them? One reason is that some sites have been around a long time (for the Web this means since the mid-1990s). As more pages and more links appear the older sites have a disproportionately greater chance of being linked to than those that are young. So the very *growth* of a

network favors the original nodes. But if this were the only way hubs grew, then instead of a power curve we would see a relatively straight line with old nodes having the most links, “middle age” nodes having a mean number of links and young nodes having few links (see Figure 2). The reason this is not the case has to do with the second principle of network growth: preferential attachment.

Figure 2 goes here

Recognition of the benefits of early entry into the Web drove the dot.com rush of the 1990s. But it is the second principle of network theory that causes the line in Figure 2 to curve, and allows the rich to get richer still.

*Preferred Attachment.* When a node in a network has more connections than other nodes it becomes even more useful to connect to comparatively. So eBay benefited by being on the Web early, but its edge grew because buyers and sellers would prefer to reach the largest possible market. In this way the straight line of Figure 2 bowed so that just a few sites obtained nearly all the auction traffic. Hundreds of copycat sites exist but each has a relatively small number of other sites that link to it.

Naturally occurring distributions of this type depart from scale-free power curves in one respect; they are not truly scale-free. The tails of true power curves stretch to infinity, whereas real networks are finite (D. Watts, 2003). One reason for this is the cost associated with adding links. Network theory assumes a low cost. Where the cost of establishing new ties is high, growth of hubs is stunted.

The growth and preferred attachment principles apply to networks of all types. **An earlier study** analyzing the use of hyperlinks in Web news stories found a similar power law distribution with 31 percent of stories in the sample accounting for 80 percent of all the links (Tremayne, 2004b). That study confirmed the growth principle; news stories on the Web have more links in each succeeding year. The study also suggested that preferred attachment in this case stemmed from editor decisions regarding the need for context in stories. Stories about international relations contained significantly more links than spot news stories about crimes and accidents. The contextual possibilities of the former, historical, geographical, social and political, were offered as an explanation for the increased hyperlinking on those stories.

### *Context and the Linking Debate*

When reporting on events journalists must decide what context is necessary to understand the story. The debate over how much context and interpretation are required is an old one. Walter Lippmann concluded that *news* should be about events and facts but that *truth* will “set them into relation with each other”(Lippmann, 1922). Although not his intention, the media’s pursuit of truth has led to a greater role for context and interpretation. Schudson documented a “decline of facts” and an increase in interpretation in reporting on the presidential State of the Union address from 1790 to 1980 (Schudson, 1982). Barnhurst and Mutz also found an increase in context and interpretation in newspaper reporting over the last century (Barnhurst & Mutz, 1997).

As news organizations migrated to the Web one facet of the medium created a new opportunity for putting current events into context: hypertext. On the Web,

additional information could be linked off the primary text, either to another Web site altogether, or to material within the domain. This development helped address two of the problems journalists routinely face. The first is the question of how much of the previous day's events need to be recapped in today's story. On the Web, the journalist could link to yesterday's news and dispense with a background paragraph. The second benefit concerns alternative points of view, those that are often excluded in favor of more mainstream views (typically two). A Web story might still focus on the two dominant opinions but provide, as well, links to other groups or organizations. Even for dominant viewpoints links to organization sites can provide greater depth for reader understanding.

These lessons have not been completely ignored by commercial news providers in the U.S. Take, for example, **a campaign story posted to the NYTimes.com on March 3, 2004 at 3:33 p.m. EST. A regular feature of the online *New York Times*, the "Times on the Trail" is presented in "blog" style, updated frequently by the paper's political reporters. Unlike most of the stories on NYTimes.com, this one uses external links liberally. The most recent update concerned a controversy over campaign ads for President George W. Bush that included images of firefighters in New York in the aftermath of the 9/11 attacks. The updates contains links to the campaign Web sites of Bush and John Kerry, as well as a link to the International Association of Firefighters. Other updates contain links to a dozen other politician Web sites.**

**Along the side of the news updates are links to other campaign articles, not just those produced by the *Times* but to 30 other publications, including *USA Today*, the *Washington Post*, and the *Wall Street Journal*. In addition to these established**

media outlets, the *Times* also links to several blog sites such as Wonkette.com and Bopnews.com.

In total, the “Times on the Trail” story contained 62 external links. It provided interested readers with a wealth of information about the 2004 elections including easy access to stories on other sites. Additionally, it gave readers a gateway to candidate Web sites where, at least potentially, readers could actually get involved in the democratic process.

With so many external links, this Web news story was atypical. But the existence of a few heavily linked stories follows the pattern of the Web as a whole.

#### *Hypothesis and Research Questions*

Network theory rests on the premises that networks grow and evolve and that links are added not randomly but attach to nodes by the principle of preferential attachment. For the case of journalism, the decisions of Web editors regarding the proper hyperlinks for a given story has been offered as the mechanism through which preferred attachment operates (Tremayne, 2004b). And a study on the use of external links tested the idea that an editor preference for internal linking rather than external linking would lead to relative decline in external linking (Tremayne, 2004a). Here we test network theory on a much larger sample with the following hypothesis:

H1: The proportion of external links in Web news stories will decrease over time.

**In addition to this hypothesis, four research questions are explored. Because previous research found that overall use of hyperlinks varied by news topic (Tremayne, 2004a), this study investigated the following question:**

**RQ1: How does the use of hyperlinks vary by news topic?**

**An earlier study (Tremayne, 2004a) found U.S. government and military sites to be the most common type of external link, followed by commercial sites. External site type was measured again using the expanded data set and addressed to this question:**

**RQ2: What types of sites are linked to most frequently by national news stories on the Web?**

**The earlier study also made note of a decline in the use external links to Web sites outside the U.S. Here we will formally explore this trend and see if it continues:**

**RQ3: Does the downward trend in the use of links to sites outside the U.S. continue in 2003 and 2004?**

**Finally, the earlier study found a small but statistically insignificant difference in the use of external links by print-affiliated Web sites versus those affiliated with broadcast companies. To see if this difference has reached significance, this question is posed:**

**RQ4: Are there any differences in the use of external hyperlinks between Web sites of print and broadcast companies?**

## Method

Answering the research questions required a content analysis of news Web sites. An after-the-fact data collection scheme, such as one that might be conducted for a newspaper content analysis, would not be practical for the ever-changing Web. Instead, a sample of Web news content was created by periodically recording the home pages of 10 national news Web sites.

### *The Sample*

News organizations with the largest circulations or audiences that had a Web presence in 1999, and that were national in scope, were selected for analysis.

Five are the Web counterparts of newspapers or magazines. These are *The New York Times*, *The Washington Post*, *USA Today*, *US News and World Report*, and *Time*. The other five are the Web counterparts of broadcast companies. They are *ABC*, *CBS*, *CNN*, *Fox News*, and *MSNBC*.

### *Sampling*

A systematic sample was recorded using the following scheme. In March of each year, the 10 Web sites were visited every four days with a random start for a total of seven days each year. Main page stories appearing on the screen of a standard-sized monitor were examined. This resulted in 1,758 stories over the six years examined, or an average of 293 stories per year.

### *Key Variables*

Over the course of the project a number of variables were measured including topic, hyperlinks per story, node modality (audio, video, text, etc.),

**internal versus external links, etc. Five of these are of interest here: hyperlinks, external links, type of external site (commercial, governmental, etc.), location of site and topic.**

***Hyperlinks* were defined as any clickable text or graphic that leads the user to additional material directly related to the story. Also included were rollover graphics that present new material when the cursor is placed on them (without clicking). What constituted “related” was left to the coders but specifically excluded standard site navigation icons and advertising. Also excluded were next page links and those beyond the first level.**

***External hyperlinks* were those that led to a Web address off the site of the news story. Sites that were co-owned or in cooperation with each other sites (such as *MSNBC* and the *Washington Post*) were not considered external.**

***Type of External Site and Location of Site* were determined by the Web address. Sites were categorized by the URL suffix (com, gov, mil, org, etc.) and, if the country of the site was still unclear, by examining the site for a postal address.**

***Topic* was coded after data collection using the headline and subhead. There were twelve categories including miscellaneous. When two or more news categories applied, the coder was instructed to select the one considered primary.**

***Coding and Reliability.* Coding occurred at two points. In the first years of the study, two coders accessed, on computers in separate rooms, two stories on each of the seven coding days. The ten Web sites were each represented by one or two stories in the reliability sample. These 14 stories represented five percent of the total sample that year. Reliability on links per story was very high, .99 by Krippendorff’s alpha (this statistic**

adjusts for small reliability samples). By 2001 software made it easier to preserve digital copies of hypermedia stories; another test of reliability was conducted in 2003. For this portion of the data, a simple random sample of 100 stories was selected. Agreement was .94 on links per story, .87 on link type, and .95 on type of external site, .81 on location, and .86 for story topic (by Krippendorff's alpha).

### Results

The sampling procedure yielded 1,758 stories over the six year period, an average of 293 stories per year. Most stories in the sample had few hyperlinks, including 300 with none and 148 with only one. One story had 137 hyperlinks. The mean was 8.50 hyperlinks per story. As predicted by network theory, a graph of stories by number of hyperlinks follows a power curve (see Figure 3).

Figure 3 goes here

The 1,758 stories contained 9,166 hyperlinks. Of these, 13 percent were links to external Web sites (1,934 links). The hypothesis and research questions concern variations in this percentage over time, by news topic, and by medium, as well as an examination of what types of external sites are linked to most often. Because the links variables follow a power curve rather than a normal distribution, nonparametric statistics were employed in the analysis that follows.

*The trend over time.* If the hypothesis is correct, the use of external hyperlinks should diminish over time relative to the use of internal links. Figure 4 illustrates the

trend. The percentage of links that go to sites external to a news organization's site is diminishing over time. **This trend was tested statistically by creating a "gap" variable.**

Figure 4 goes here

**The gap variable** was calculated for each case by subtracting the number of external links for a story from its internal links. For all 1,758 stories, the mean gap between external and internal links is 6.3 links. As hypothesized, the gap grows year to year. A Kruskal-Wallis test confirmed significant differences amongst years ( $H=134.5$ ,  $p<.05$  by Bonferroni procedure). Additionally, a test of linear trend using an ANOVA procedure also confirmed the trend ( $F=55.6$ ,  $p<.001$ ). ANOVA is not typically used for data that do not follow a normal distribution because of a loss in statistical power, but with larger sample sizes can yield accurate p values (Wilcox, 2001). The analysis confirmed the hypothesis. The use of external links in stories on national news Web sites is diminishing relative to internal links. Even in absolute terms the number of external links is at its lowest in 2004 (0.7 links per story), compared to a high in 2000 (1.6 links per story).

*External Links and News Topic.* The first research questions asks: "How does the use of hyperlinks vary by news topic?" Four of the 12 news topics had less than 50 cases in the sample and are excluded from analysis here, as is the miscellaneous category. The remaining stories had an average of 6.9 fewer external links than internal ones. Figure 5 shows the results by topic.

Figure 5 goes here

The proportion of external links was particularly low for two news categories: financial and international relations stories. International relations stories had almost nine fewer external links than internal. Possible reasons for this are discussed in the next section. Financial stories were also much more likely to be linked internally rather than externally. Many of the sites in this sample have built up a large business archive, including profiles of most publicly-traded companies. This allows for a significant amount of internal linking.

Two topics, health and spot news, had a relatively small gap, meaning a more equal distribution of internal and external links. The differences in the internal-external gaps between topic are statistically significant (Kruskal-Wallis  $H=122.3$ ,  $p<.001$ ).

*External Site Type* The second research question asks: “**What types of sites are linked to most frequently by national news stories on the Web?**” The external Web site links that appeared in national Web stories during the study period were categorized by domain suffix (.com, .edu, .gov, etc.). Dot-com sites were also categorized as media and non-media.

Military and government Web sites, both in and outside the U.S., are the most common external link category (40%), followed by commercial non-media sites (23%), non-profit organizations (21%), media sites (13%), and educational institutions (3%).

*Location of External Sites* **Research question three concerns the location of the sites for which external links appear.** Most of the external links, 73 percent, are to U.S.-based Web sites. This figure alone is unsurprising, but the trend over time is

revealing. In 1999, when online journalism was still in its infancy, links to Web sites outside the U.S. constituted 60 percent of external links. By 2004, that number was down to just 5 percent. Figure 6 shows the results by years.

Figure 6 goes here

The drop in the use of external links to Web sites outside the U.S. is due mainly to a decreased use of non-US government or military sites. In 1999, government Web sites outside the US accounted for 71% of all government links. That percentage subsequently fell every year to reach a low of just 4% in 2004. Increasingly, when US news Web sites include links to external sites, those sites are US-based.

*Print versus broadcast* Research question four asks about differences between U.S. print and broadcast companies in the use of external linking. In overall terms, the broadcast sites used more hyperlinks (10.9 per story) than the print sites (5.8) over the six years in question, a statistically significant difference (Mann-Whitney  $U=262979$ ,  $p<.001$ ). As for external links specifically, national print and broadcast Web sites were closer, but broadcast sites offered more (1.3) than print sites (0.9) at a statistically significant rate (Mann-Whitney  $U=328875$ ,  $p<.001$ ). As a percentage of all links offered, however, more links on print stories were external (15.9%) than for stories on the broadcast sites (11.7%).

## Discussion and Conclusions

This study examined a six year sample of national news stories on the Web and found the use of external links is continuing to decline in ways predicted by network theory. News Web sites are linking to external sites less frequently than just a few years ago. **At the same time, use of internal links has moved sharply higher (Tremayne, 2004b). The conclusion is clear:** As each organization builds up its own archive of Web content, this material is being favored over content that is off-site. This is just one example of *preferred attachment* which is the driving principle of network theory. As predicted by network theory, **a plot of external links per story** has a distribution that follows a power curve.

Figure 7 goes here

Figure 7 shows that most stories in the sample (1,272 out of 1,758) have no external links while a few have dozens. **Although we can only infer the reasons behind this example of preferred attachment, one obvious possibility is that Web managers are trying to keep users from leaving their pages. If this is indeed the case, it's another example of market-driven journalism (McManus, 1995). To more fully explore this possibility, a survey of Web editors or a series of in-depth interviews could be conducted. Among other possible explanations for the trend are fear of association with external sites and/or legal concerns associated with possible libelous statements on those sites.**

Although not **tested here as a hypothesis**, a similar pattern is occurring with links to foreign Web sites. In 1999, 60 percent of external links went to sites outside the U.S.

By 2002 that number was just below 20 percent and this study found it falling even further, to just 5 percent in 2004. The foreign Web sites seeing the largest decrease are government-run sites. Two possible causes are suggested here. One is that Web editors track user traffic and have found those links to be little used. Another possibility is a post-9/11 effect; a rallying around the flag by Web editors in the U.S.

Some topics are more likely to receive external linking than others. Stories about spot news and health had a more evenly divided source of linked material than other topics. In the case of spot news, this has more to do with the relative lack of links overall than with special attention being paid to external linking. Spot news is the least “linked” news category of the twelve. Because of the nature of these stories, Web editors have relatively few options for including linked material. In some cases, links to other Web sites may be the only choice.

Health stories, on the other hand, present a more interesting case. It may be that Web editors recognize the public benefit of directing interested users to important medical information. And because of the delicate nature of medical advice, they may see this as a safer choice.

Financial news and international relations stories had the widest gaps between internal and external linkages. The former is not surprising. Many news Web sites have built extensive business news sections, partly because early Web adopters tended to be more affluent than the population as a whole. This is a classic case of the growth principle combining with preferential attachment (by editors) to generate a fairly huge amount of internal content that editors can choose from on a daily basis. And the more

links there are on these stories the longer this highly sought after readership will stay on the site.

The gap for stories about international relations is unfortunate. These types of stories provide for many opportunities to link to external sites. Rather than doing this, Web editors are teaming up with other content providers, such as encyclopedia publishers, to keep users on their sites. The benefit of this is that some contextual material not available through traditional media channels is available online. Of course, the downside is that U.S. users are increasingly isolated from a world of ideas and points-of-view.

This study found relatively little difference in the using of external links between news sites run by print companies and those of broadcast outlets. While the broadcast sites used slightly more external links per story, the print sites still have a greater overall percentage of external links.

Are the documented patterns inevitable? Because the linking patterns evident on the news Web sites examined here are predicted by network theory it may seem that way. But there is another possibility. Independent content providers, called bloggers, thrive by linking users to a virtual public sphere where current events are debated. The blogosphere is reciprocal in nature, those who link to other sites are linked to in turn. Bloggers often link to traditional media sites, but many, in response to the anti-deeplinking practices of media sites, are no longer providing those links. The impact of this for most media sites is small, at present. If the amount of traffic on non-media sites grows, Web editors may want to reconnect with that network of users. Such a move is

advocated by many independent journalists (Dube, 2003). If that occurs, a new preferred attachment principle may create a different pattern in the use of external links.

### *The “Hubs” of External Linking*

**The existence of hubs in networks was long ago observed. The value of network theory is that it provides a systematic explanation for the distribution common to so many networks and helps us understand how hubs form. Like the use of links overall, this study found that the use of external links in news stories on the web also follows a power distribution. A fair question, then, is what types of stories are the “hubs” of external linking? The sample in this study contained one story from 2003 that contained more than 40% of the external links for all stories from that year: 84 external links in all. The story was posted to USNews.com on March 3, 2003 and, like the *New York Times* story mentioned earlier, is presented in blog style with the most recent update at the top. This ongoing story is the “War on Terror” and the latest update concerned the arrest of an al Qaeda suspect in Pakistan. The page provides a series of stories related to the war on terror and links to 34 USNews.com stories. Additionally, it provides external links to raw source material, mostly documents on U.S. government Web sites, 49 in all. There are also links to 14 other media sites that have related stories and 12 links to non-profit sites such as the ACLU.**

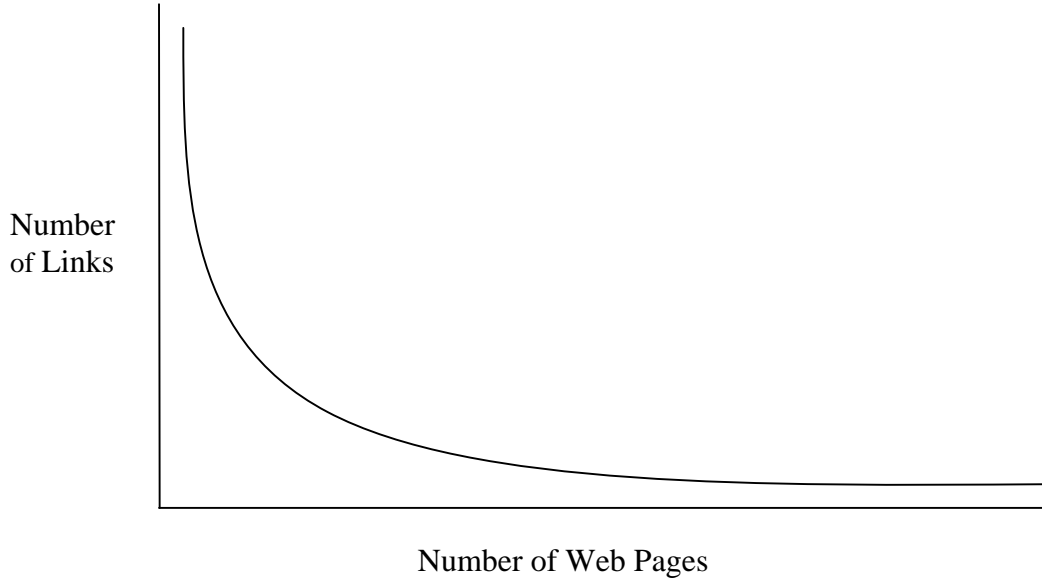
**News pages such as this one are exceedingly rare, but they do exist. Rather than adding external links to daily news stories, editors appear to favor these “repository” pages that can be added to and reposted to the site as events warrant.**

**The result is a small number of pages with dozens of external links and a large number of pages with few or no external links.**

**The USNews.com story on terrorism and the NYTimes.com story on the campaign trail share some features that might have resulted in them becoming “hubs” for external linking. Each concerned a topic that had been in the news a lot over the preceding two years. Each story was ubiquitous, covered by all news organizations, with little advantage in coverage from one outlet to the next. Each story involved government and related to, potentially, hundreds of government or political Web sites. As Web editors appear to favor U.S. government Web sites to all other types, this provided them with ample material for external linking.**

**If the patterns documented in this study continue, we can expect a growing percentage of stories that have few or no external links and a simultaneous rise in the number of external links on a very small number of news stories. These stories are likely to be on topics that have been in the news for many months or years and ones involving government.**

**Figure 1**  
*Power Law Distribution*



**Figure 2**  
Distribution Based on Growth Alone

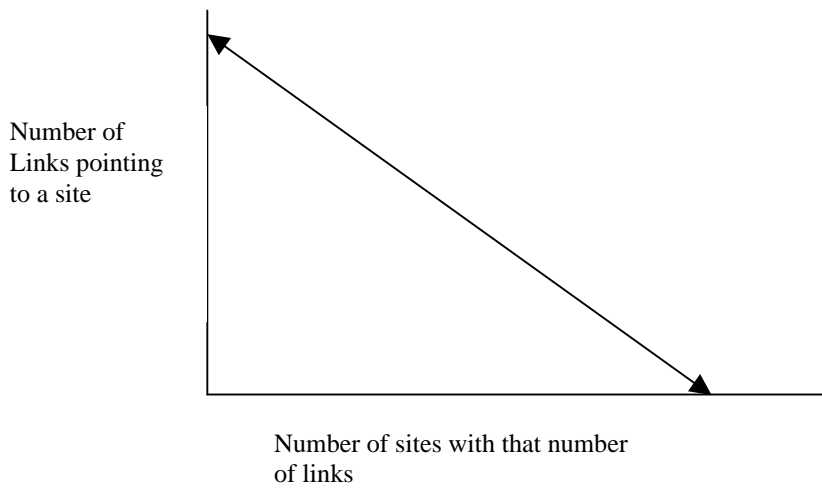


Figure 3  
Hyperlink per story by Number of Cases (count)

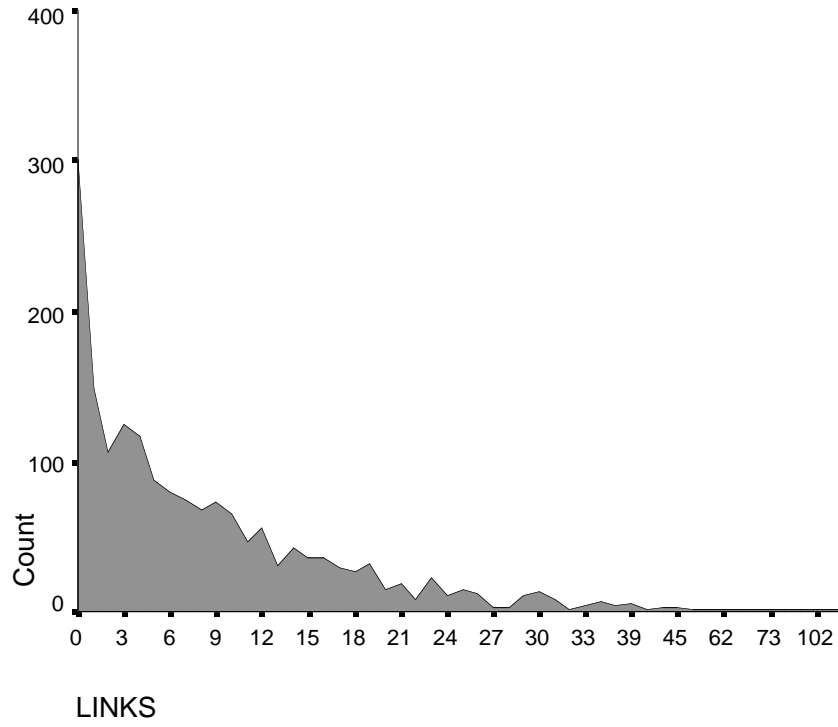


Figure 4  
External Hyperlinks by Year

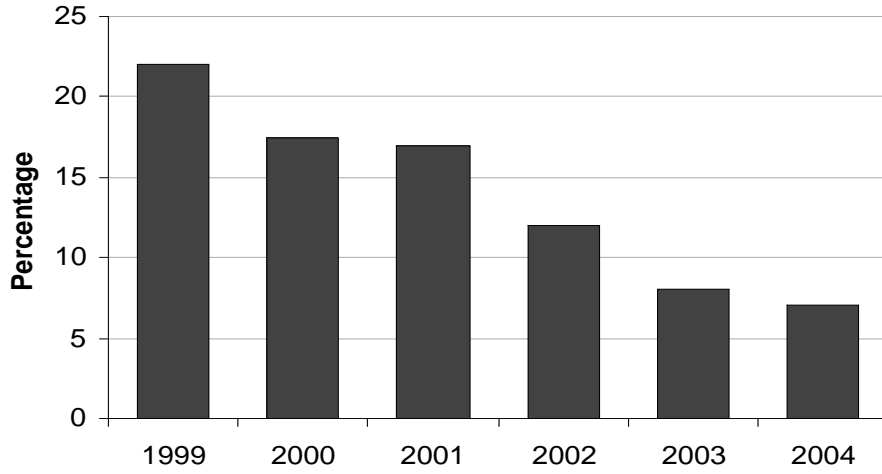


Figure 5  
External-Internal Link Gap by Topic

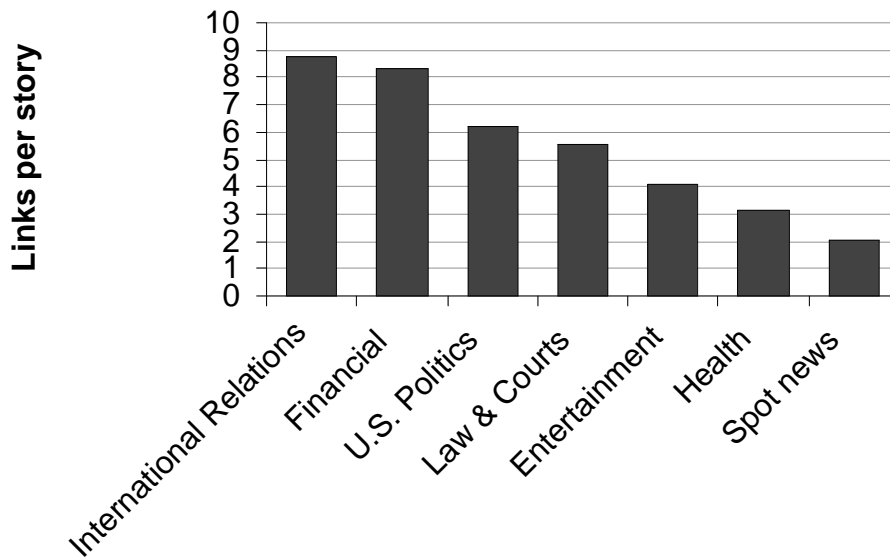


Figure 6  
Percentage of External Links to non-US Web sites

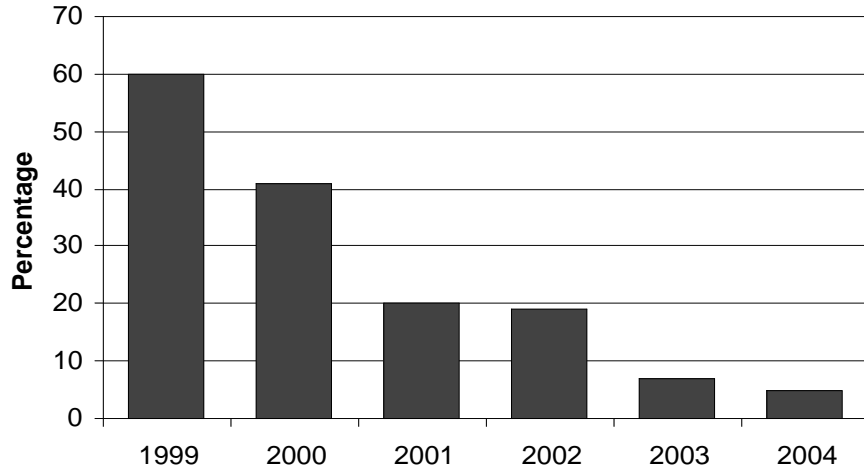
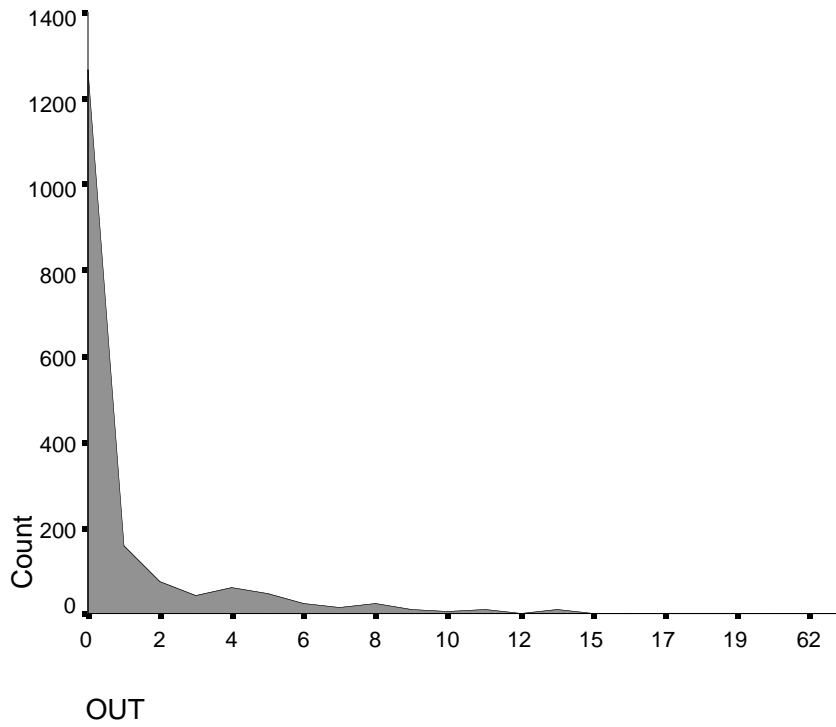


Figure 7  
External Hyperlinks per Story by Number of Cases



References

- Adamic, L. A., & Huberman, B. A. (2001). The Web's Hidden Order. *Communications of the ACM*, 44(9), 55 (56 pages).
- Barabasi, A.-L. (2002). *Linked: The new science of networks*. Cambridge, Mass.: Persues Publishing.
- Barabasi, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286, 509-512.
- Barnhurst, K. G., & Mutz, D. (1997). American journalism and the decline in event-centered reporting. *Journal of Communication*, 47(4), 27-53.
- Dube, J. (2003). *News sites loosen linking policies*. Retrieved October 28, 2003, from <http://www.cyberjournalist.net/news000720.php>
- Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78, 1360-1380.
- Lippmann, W. (1922). *Public Opinion*. New York: Penguin Books.
- McManus, J. H. (1995). *Market driven journalism: Let the citizen beware*. Thousand Oaks, CA: Sage.
- Milgram, S. (1967). The small world problem. *Psychology Today*, 2.
- Rheingold, H. (1994). *The electronic landscape: A writer's perspective*. Paper presented at the Writers Retreat On Interactive Technology and Equipment Conference, University of Vancouver, Vancouver, British Columbia.

- Rogers, E. M. (1983). *Diffusion of innovations*. New York: Free Press.
- Rosen, J. (1995, August 7). Cheap speech. *The New Yorker*.
- Schudson, M. (1982). The politics of narrative form: The emergence of news conventions in print and television. *Daedalus*, 3(4).
- Tremayne, M. (2004a). News websites as gated cybercommunities. *Convergence: The Journal of Research into New Media Technologies*, 10(4), In press.
- Tremayne, M. (2004b). The Web of Context: Applying Network Theory to the Use of Hyperlinks in Journalism Stories on the Web. *Journalism & Mass Communication Quarterly*, 81(2), 237-253.
- Watts, D. (2003). *Six degrees: The science of a connected age*. New York: W. W. Norton and Co.
- Watts, D. J. (1999). Networks, Dynamics, and the Small-World Phenomenon. *The American journal of sociology*, 105(2), 493 (435 pages).
- Wilcox, R. R. (2001). *Fundamentals of modern statistical methods: Substantially improving power and accuracy*. New York: Springer Verlag.