

Network Manipulation (with application to Political issues)

Panagiotis Takis Metaxas
Department of Computer Science
Wellesley College
Wellesley, MA 02481, USA
pmetaxas@wellesley.edu

I. INTRODUCTION

We live in an increasingly interconnected world, one in which a growing number of people turn to the web to make important medical, financial and political decisions [1]. As more people use the Web’s search engines daily as their primary source for locating information on many important issues, search engines are in the position to influence what is perceived as relevant information through their mechanism of ranking web pages. However, as studies have shown [2], interested groups and individuals can also make use of web spamming mechanisms to trick search engines in ranking their pages higher than those of their rivals.

The battle for controlling the messages in cyberspace is spreading over many ideological, cultural, and political issues where controversial positions vie for the public support. For example, consider issues such as abortion legality and morality, children vaccination risks, creationism vs. evolution, homosexuality, etc. [3]. Nowhere this battle is more obvious than when it comes to issues related to national elections. Obviously, the stakes here are high. If one is able to influence the public elect officials that will be favorable to his/her agenda, this will have far-reaching implications.

The term “Web Spam” or Adversarial Web Search [4] is broadly used to describe misinformation planted on the Web. Google-bombs are probably the best known examples of Web spamming, because of their broad coverage in the press [5]. Web spammers are creating misinformation using “text spam” and “link spam”, while using “cloaking” to cover their tracks. This generic categorization of practical actions does not explain why they are successful, or why Google Bombs may or may not be successful.

Recently, spamming techniques have been introduced in Social Media [6], making it more appropriate to talk about “Network Manipulation” rather than just Web spam. The basic techniques of this manipulation, however, can be traced in the long history of propagandistic techniques in society [2].

In this presentation proposal we provide an overview of the technical side of Network Manipulation, and discuss its connection to Social Propaganda. Further, we describe how it has been extended in the area of Social Media and discuss some of its successes and failures when it comes to political

| Ranking Techn. | Net. manipulation | Soc. Propaganda |
|------------------|-----------------------------|-------------------------|
| Doc Similarity | keyword stuffing | glittering generalities |
| Site popularity | link farms | bandwagon |
| Page reputation | mutual admiration societies | testimonials |
| Anchor text | Google bombs | card stacking |
| Real-time search | Twitter bombs | plain folks |

Table I
RANKING TECHNIQUES BY SEARCH ENGINES ARE LISTED ALONG WITH THE RESPONSE OF THE WEB GRAPH MANIPULATORS AND THEIR CORRESPONDING PROPAGANDISTIC TECHNIQUES.

issues, especially related to congressional elections. We end with a discussion of what the search engines have said to have done so far, and what is likely that they have done without admitting it.

II. WHAT IS NETWORK MANIPULATION

Network Manipulation is the attempt to modify the Web graph and/or a social network, and thus influence online network tools in ways beneficial to the manipulators. The modification of a network is in terms of altering its structure and/or its contents. The online network tools that manipulators try to influence are typically search engines and online social media.

One can explain most of the major technological developments in the area of search engine technologies as their attempt to stop the successes of Web Graph manipulators (See Table I). For example, Google’s attempt to combat link farms (groups of interlinked web sites controlled by the same entity) with the introduction of the famous Page Rank algorithm was countered by the introduction of “mutual admiration societies” (organized groups of manipulators who have achieved high reputation independently for unrelated themes) [2]. In terms of propagandistic techniques, this corresponds to “testimonials” often used by advertising companies: A famous actor playing the doctor on TV urges the audience to buy a particular pain killer, as if he is an expert in medicine.

Each of the network manipulation techniques is implemented with altering the structure or the contents of the network’s components. For example, in creating mutual admiration societies, the so-called “black hat” search engine

optimization companies organize themselves exchanging links [7]. To create Google bombs, they announce the terms to be targetted as anchor text and the links to support, sometimes even in the open [8], as we discuss next.

Ever since the appearance of the “miserable failure” hoax that produced search results that included President George W. Bush (and later, President Barack Obama [9]), Google bombs have attracted a lot of attention in the media, since they appear that with little effort, net manipulators can game the sophisticated algorithms of the search engines. This practice of “gaming” the search engines is implemented with mislabeled anchor text techniques (corresponding to the “card stacking” propagandistic technique; see Appendix??), in which web site masters and bloggers use the anchor text to associate an obscure, negative term with a public entity [10]. In particular, during the 2006 US midterm congressional election, a concerted effort to manipulate ranking results in order to bring to public attention negative stories about Republican incumbents running for Congress took openly place under the solicitation of the progressive blog, MyDD.com (My Direct Democracy) [8].

Can these efforts be stopped? Search engines have tried to counter this bad publicity by announcing initially a plethora of features for ranking, and recently more sophisticated algorithms than Page Rank [11]. These changes seemed to bear fruits in the 2008 congressional elections where very few spamming sites rose to the top of search results [12].

It appears, however, that these supposedly “new, sophisticated” (and secret) algorithms would not scale: They are likely pre-computed search results on a white-list of search query terms that would likely be bombed.

In particular, these new features and algorithms appeared to be highly effective in the announced launching of 98 Google bombs [13] during the 2010 congressional elections, as we predicted [14]. Both in 2008 and 2010, searches related to congressional candidates would bring up in the top six results a ranked list of the same sources: one or more of the campaign sites of the candidates, their official web pages, their wikipedia page, and google images (Figure 1). This was consistent across the candidates independently of their visibility or of the fact that they were under attack by thousands of political spammers. Moreover, the relative location of each result in a period of 29 weeks remained remarkably steady (Figure 2).

At the same time, they proved to be completely ineffective in at least two examples of network manipulation that were under their radar screen: The Decor-My-Eyes site [15] that successfully used bad publicity to rank high, and the JCPenney case [16] that used anchor text manipulation of link farms. After these cases became known, their relative position changed dramatically moving downwards by dozens of locations per day.

These counter examples provide strong evidence that we are not talking about new sophisticated algorithms in effect,

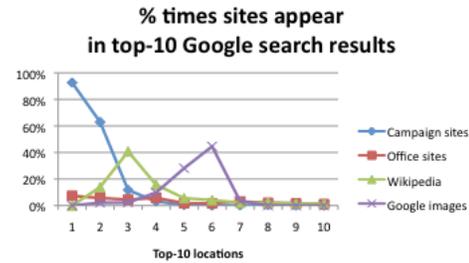


Figure 1. Percentage of times a site appeared in a particular position in the top-10 search results.

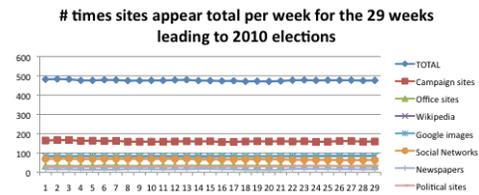


Figure 2. Relative change in position of collections of sites during the 29 week period preceding the 2010 congressional elections.

but for old-fashion, hand-crafted list of blacklisted sites and white-listed terms.

ACKNOWLEDGMENT

The author would like to thank Eni Mustafaraj, Era Vuksani, Ljubica Ristovska and Dana Bullister for collecting some of the data presented in this paper. This work was partially supported by a Brachman-Hoffman Fellowship.

REFERENCES

- [1] Pew Foundation, “Pew internet and american life project.” <http://www.pewinternet.org>, 2008.
- [2] P. Metaxas, “Web spam, social propaganda and the evolution of search engine rankings,” *Lecture Notes in Bus. Info. Proc. (LNBIP)*, accepted; to appear 2010.
- [3] M. Hindman, K. Tsioutsoulis, and J. Johnson, “Googlearchy: How a few heavily-linked sites dominate politics on the web,” in *Annual Meeting of the Midwest Political Science Association*, April 3-6 2003.
- [4] C. Castillo and B. D. Davison, “Adversarial web search,” *Foundations and Trends in Information Retrieval*, vol. 4, pp. 377–486, June 2010.
- [5] T. McNichol, “Your message here.” *New York Times*, Jan. 22 2004.
- [6] C. Grier, K. Thomas, V. Paxson, and M. Zhang, “@spam: the underground on 140 characters or less,” in *Proceedings of the 17th ACM conference on Computer and communications security, CCS '10*, (New York, NY, USA), pp. 27–37, ACM, 2010.

- [7] srainwater, "Nigritude ultramarine faq." <http://www.nigritudeultramarines.com/>, 2004.
- [8] T. Zeller Jr., "Gaming the search engine, in a political season." *New York Times*, Nov. 6 2006.
- [9] D. Sullivan, "Obama Is "Failure" At Google & "Miserable Failure" At Yahoo." <http://searchengineland.com/yahoo-obama-is-a-miserable-failure-16286>, January 22., 2009.
- [10] T. McNichol, "Engineering google results to make a point." *New York Times*, Jan. 22 2004.
- [11] S. Hansell, "Google keeps tweaking its search engine." *New York Times*, Jun. 3 2007.
- [12] P. Metaxas and E. Mustafaraj, "The battle for the 2008 us congressional elections on the web," in *Proceedings of the Web Science 2009 Conference*, (Athens, Greece), March 2009.
- [13] C. Bowers, "Call to action: Google-bombing the election." *Daily Kos*, <http://www.dailykos.com/story/2006/10/22/133437/99>, Last retrieved on Nov. 23, 2010.
- [14] S. L. Stirland, "Google is latest weapon vs. gop." *Politico*, <http://www.politico.com/news/stories/1010/43767.html>, Last retrieved on Nov. 23, 2010.
- [15] D. Segal, "A bully finds a pulpit on the web." *New York Times*, Nov. 20 2010.
- [16] D. Segal, "The dirty little secrets of search." *New York Times*, Feb. 12 2011.
- [17] D. Welch, "Power of persuasion - propaganda," *History Today*, vol. 49, no. 8, pp. 24–26, 1999.
- [18] A. M. Lee and E. B. Lee(eds.), *The Fine Art of Propaganda*. The Institute for Propaganda Analysis. Harcourt, Brace and Co., 1939.

III. APPENDIX: ON PROPAGANDA THEORY

We offer here a brief introduction to the theory of propaganda detection. For more information, see [2].

There are many definitions of propaganda, reflecting its multiple uses over time. One working definition we will use here is

Propaganda is the attempt to modify human behavior, and thus influence people's actions in ways beneficial to propagandists.

Propaganda has a long history in modern society and is often associated with negative connotation. This was not always the case, however. The term was first used in 1622, in the establishment by the Catholic Church of a permanent Sacred Congregation *de Propaganda Fide* (for the propagation of faith), a department which was trying to spread Catholicism in non-Catholic Countries [17]. Its current meaning comes from the successful Enemy Propaganda Department in the British Ministry of Information during WWI. However, it was not until 1938, in the beginning of WWII, that a theory was developed to detect propagandistic techniques. For the purposes of this paper we are interested in ways of detecting propaganda, especially by automatic means.

First developed by the Institute for Propaganda Analysis [18], classic Propaganda Theory identifies several techniques that propagandists often employ in order to manipulate perception.

- **Name Calling** is the practice of giving an idea a bad label. It is used to make people reject and condemn the idea without examining the evidence. For example, using the term "miserable failure" to refer to political leaders such as US President George Bush can be thought of as an application of name calling.
- **Glittering Generalities** is the mirror image¹ of name calling: Associating an idea with a "virtue word", in an effort to make us accept and approve the idea without examining the evidence. For example, using the term "patriotic" to refer to illegal actions is a common application of this technique.
- **Transfer** is the technique by which the propagandist carries over the authority, sanction, and prestige of something respected and revered to something he would have us accept. For example, delivering a political speech in a mosque or a church, or ending a political gathering with a prayer have the effect of transfer.
- **Testimonial** is the technique of having some respected person comment on the quality of an issue on which they have no qualifications to comment. For example, a famous actor who plays a medical doctor on a popular TV show tells the viewers that she only uses a particular pain relief medicine. The implicit message is that if a famous personality trusts the medicine, we should too.
- **Plain Folks** is a technique by which speakers attempt to convince their audience that they, and their ideas, are "of the people," the "plain folks". For example, politicians sometimes are seen flipping burgers at a neighborhood diner.
- **Card Stacking** involves the selection of facts (or falsehoods), illustrations (or distractions), and logical (or illogical) statements in order to give an incorrect impression. For example, some activists refer to the Evolution Theory as a theory teaching that humans came from apes (and not that both apes and humans have evolved from a common ancestor who was neither human nor ape).
- **Bandwagon** is the technique with which the propagandist attempts to convince us that all members of a group we belong to accept his ideas and so we should "jump on the band wagon". Often, fear is used to reinforce the message. For example, commercials might show shoppers running to line up in front of a store before it is open.

The reader should not have much trouble identifying additional examples of such techniques used in politics or advertising.

¹Name calling and glittering generalities are sometimes referred to as "word games."