

Issues in Measuring Power and Influence in the Blogosphere

Norhidayah Azman
Learning Societies Lab
School of Electronics and
Computer Science
University of southampton
nba08r@ecs.soton.ac.uk

David E. Millard
Learning Societies Lab
School of Electronics and
Computer Science
University of southampton
dem@ecs.soton.ac.uk

Mark J. Weal
Learning Societies Lab
School of Electronics and
Computer Science
University of southampton
mjlw@ecs.soton.ac.uk

ABSTRACT

Power and influence in the blogosphere can be elusive in nature yet they still play a perceived role in determining future events. Cases like the resignation of Senator Trent Lott and the Dan Rather scandal have been quoted as evidence of the power of blogs, highlighting the potential role of blogs as a new medium for instigating change. However, supporting evidence of power in the blogosphere is often anecdotal. Over the past century, power definitions have been continuously debated amongst political scientists. Based on their theories, this paper defines power as the ability to produce effects among others when making decisions. The blogosphere's emergence echoes the phenomena of 17th century pamphleteering, where the invention of the printing press had facilitated publishing beyond institutional control such as churches and the monarchy. In Web Science, power in the blogosphere fits in as one of the issues that emerge due to the macro nature of the blogosphere. Previous work on identifying power and influence has resulted in papers positioning blogs within a hierarchy, based on metrics such as links, comments and phrases. These metrics are constrained by what can be observed. Moreover, quantifying power would involve finding a tractable, concrete link between blog activity and an action. External influences such as traditional media also make it difficult to correctly analyze blogs-to-action correlations. Nonetheless, the use of data propagation could be suitable to measure power. One potential methodology is to correlate blog trends to a tractable action like an e-petition.

Keywords

power, influence, blogosphere

1. INTRODUCTION

Power and influence in the blogosphere can be elusive in nature yet they still play a perceived role in determining future events. When Malaysia's ruling government lost many seats during the general elections in March 2008, the previous Prime Minister said that "We ... lost the Internet war ... We made the biggest mistake in thinking that it was not important ... We thought that the newspapers, the print media, the television was supposed to be important, but the young people were looking at SMS and blogs." This

highlights the perceived role of blogs as a new medium for instigating political change.

In Section 2, we outline the position of blogs in Web Science and the various examples of perceived power in the blogosphere. Section 3 presents the parallelisms between blogs and 17th century pamphlets, followed by Section 4 where we discuss the different definitions of power and what this paper defines as power. Section 5 details related work on identifying influence in the blogosphere, which leads to Section 6 that discusses the difficulties of reconciling theoretical definitions of power to the actual experiments trying to measure power. Finally, Section 7 proposes the use of data propagation as a crude measurement of power.

2. BLOGS IN WEB SCIENCE

Web Science aims to bring a multi-disciplinary approach towards studying the Web as an entity. Web applications' micro properties are initially tested in small groups, but once these applications are released for public use on the Web, macro properties begin to emerge, sometimes unexpectedly, thus requiring a better understanding of how this occurs [11]. One of the aims of Web Science is to identify emergent macro issues and deploy research methods from sociology to computer science in order to address them.

Power in the blogosphere fits in as one of the issues that emerge due to the macro nature of the blogosphere. Individual, unread blogs may seem powerless, but when given the millions of blogs currently online, we have begun to see instances of how blogs have empowered its authors.

For example, Senator Trent Lott was heavily criticized by the blogosphere for making racist comments during a speech, a fact that was ignored by the mass media at the time. Due to increased exposure in blogs, this issue was later taken up by the press. He then resigned as Senate Minority Leader in 2002 [9].

Another example is the Dan Rather scandal in 2004. In his television show 60 Minutes, he presented documents which supposedly showed irregularities in George W. Bush's National Guard service. However, certain bloggers began to claim that the documents were doctored and should not have been used by credible mainstream media as proof to support the story. As a result, Dan Rather made a public

Copyright is held by the authors.

Web Science Conf. 2010, April 26-27, 2010, Raleigh, NC, USA.

apology and subsequently retired [22].

These two cases have been perceived as supporting evidence of the power of blogs. However, power evidence like these are often anecdotal. Different approaches are needed to observe and quantify power instead of depending solely on qualitative analysis.

3. BLOGS AND PAMPHLETEERING

The blogosphere's emergence echoes the phenomena of 17th century pamphleteering. The printing press had allowed anybody who could afford them to publish beyond institutional controls such as churches and the monarchy [8].

Some people became notable figures because of their involvement in pamphleteering. Jonathan Swift wrote the *Drapier's Letters* between 1724 and 1725 to voice his disagreement over the Irish government's privately minted copper coins. The contract awarded was deemed corrupt and the coins were claimed to be poor in quality. Despite the government's opposition towards *Drapier's Letters*, he managed to instigate a public uprising, resulting in a nationwide boycott and the eventual withdrawal of those coins [3].

Fast forward to the 21st century, similar issues are seen happening around the blogosphere. This century's blogosphere has provided the same potential for accumulating power, just like the printing press did after its invention in the 15th century. Several people have commented on the parallels between bloggers and pamphleteers. Dan Bricklin mentioned similarities between the two as early as 2001 [4]. A blogger is "today's highly democratic and accessible version of the pamphleteer," breaking existing information monopoly, distributing ideas and shaping public opinion "in the tradition of pamphlets" [14]. Rick Klau quoted a paragraph out of Pulitzer winner Bernard Bailyn's book 'The Ideological Origins of the American Revolution' and swapped the word "pamphlets" into both "weblogs" [12] and "Twitter", arguing that "the dynamics that make them so compelling are the very same ones that made pamphlets so powerful 225 years ago" [13]. From these arguments, there seems to be a consensus that pamphlets were considered powerful during its era, and the blogosphere appears to be emulating this, as shown by the anecdotal evidence of power mentioned in the previous section.

One aspect that facilitated the adoption of pamphlets was the ability to publish anonymously. Anonymity encourages expression by authors — both pamphleteers and bloggers — without fear of retaliation, social ostracism or privacy infringements. When Jonathan Swift wrote the *Drapier's Letters*, he wrote them under the pseudonym of M. B. Drapier to avoid retaliation.

The ability to publish anonymously has been a point of contention, between individuals looking to express their opinions against those who seek to remove this anonymity in order to hold these individuals accountable for their actions. In America, several landmark Supreme Court cases have been ruled in favour of protecting its citizens' freedom of speech, based on the First Amendment to the United States Constitution which allows citizens to associate and

express opinions freely without infringement. In the *McIntyre v. Ohio Elections Commission* case, the Supreme Court overturned an earlier statute that banned anonymous pamphleteering because this was unconstitutional according to the First Amendment [6]. In the *Dow v. Cahill* defamation trial, the defence sought to reveal the identity of an anonymous blogger. The case was lost when the Supreme Court judged that the perceived harm of such anonymous blogs do not commensurate the breach to an individual's freedom of speech [20]. Cardillo argues that readers should be left to evaluate the validity of anonymous literature themselves, without any source bias nor censorship from the authority [6].

While the above court cases suggest that bloggers do have some measure of immunity against attempted retaliation, this may not be the case for bloggers outside of America. In Iran, the government had traced and convicted activists who had used Twitter during the civil protests following the disputed Iranian elections in 2009 [5]. Similar issues have happened in the workforce, where two companies, Immunomedics and Raytheon Co., succeeded in forcing the identity disclosure of anonymous message board users by proving in court the extent of harm caused by their anonymous posts [15]. These court rulings on both blogs and pamphlets seem to point that parties such as the state and commercial bodies prefer to resort to legal outlets in order to minimize if not eliminate the impact of such publications which go against their interests.

This perceived impact of pamphlets and blogs, plus their ability to make a public call of action, leads into the next section, where we attempt to define power.

4. DEFINITIONS OF POWER

The meaning of power has been continuously debated amongst political scientists over the past century. Various power theories have been proposed, underlining the subjectivity of power.

For example, Russell simply defined power as "the production of intended effects" [23]. More recently, Morriss looked at the different concepts of power [21]. First, the practical concept asked the questions "What can I do?" and "What can you do to/for me?". Then, the moral concept asks what could have been done. Finally, the evaluative concept considers the value of power distributions. Meanwhile, Lukes stated that there are three dimensions to power [19]. The first dimension is the manipulation of observable behaviour, in decision-making situations, involving observable conflict. The second dimension is the prevention of decisions, still involving observable conflict. The third dimension is the absence of observable conflict, where certain issues are downplayed, resulting in latent conflicts.

From these discussions, we can see that it is difficult to define power without considering other aspects: the context of the situation, the effect that shows power, the parties exerting power and those impacted by power.

However, one common thread amongst these arguments is that power involves the ability of A to cause an action by

B. This paper defines power as the ability to produce effects among others. This effect happens within a decision-making process, narrowing the scope of this definition to facilitate useful experimentation within the domain of blogs.

5. RELATED WORK

Which metrics should be used to measure power in the blogosphere? Previous work on identifying power and influence has resulted in papers positioning blogs within a hierarchy, based on various tractable metrics.

Adar et. al. ranked blogs according to both explicit and implicit link structure, where the blog that begins a trend first gains a higher ranking [1]. Agarwal et. al. ranked blogs according to the total comments generated by individual blog posts [2]. Song et. al. ranked blogs according to how efficient information flowed through them [24]. Leskovec et. al. created MemeTracker which tracks the volume of phrases mentioned in the blogosphere over time [16].

These papers have used existing objects such as links and phrases which are observable and easier to experiment with. We can derive that the authors were assuming that metrics such as links, comments and phrases were forms of action, carried out by their peers or readers. The author with the most amount of these metrics are considered powerful. The problem with metrics like these is that they are constrained by limits to what can be observed. Going back to Section 4, this paper's definition of power is the ability to produce effects among others within a decision-making process. This poses a problem if the effects that show power are not necessarily concrete and observable. The following section outlines the problems in attempting to quantify and subsequently study power empirically.

6. PROBLEMS IN MEASURING POWER

In Section 4, we have stated that power involves context, the effect of power and the parties involved in the power exchange.

If we were to assume power as the ability to attract the most number of blog readers, then the obvious metric that we could use is the number of hits or unique visitors that a blog has. However, not every blog has a hit counter, nor would these counts be easily accessible, if available, by researchers or the general public. Compare this to the availability of links and comments. Hence, when attempting to quantify the effects of power, we are limited to effects that can be observed, thus narrowing the scope of power that we can investigate experimentally.

Moreover, quantifying power in the blogosphere would involve finding a tractable, concrete link between blog activity and an online or offline action. Studies have shown correlations between blogs and financial stock activity [7], but again constraints impose a limit to which offline actions that can be studied. For instance, it is easier to measure stock activity as opposed to measuring corporate good will. Apart from that, the influence of external variables such as traditional media makes it difficult to correctly identify blogs-to-offline

correlations when studying offline actions.

Going back to Section 4, we discussed Lukes's three dimensions of power [19]. The use of tractable, observable metrics such as links and comments may satisfy Lukes's first dimension of power. However, it is quite difficult to find equivalent metrics that satisfy his second and third dimensions. Since the second dimension concerns the power to compel people not to make a decision, what is the most suitable parallel that we could find in the online world? This gets more complicated when we attempt to satisfy the third dimension, where latent conflicts mean that there researchers would find it difficult to see whether issues are being covered up or not, unless researchers themselves already know about those hidden issues.

7. DATA PROPAGATION

Given the arguments above, we are restricted to experimenting with metrics that are observable and tractable. Despite this limitation, there is still room to quantify power. One example approach is via data propagation.

Based on our definition of power as the ability to produce effects among others within a decision-making process, then we can derive that the total number of people whose decisions are affected can be interpreted as a crude measure of that power. The metrics used by the past papers discussed in Section 5 highlights the importance of volume: more involvement, both in physical and virtual terms, arguably means more influence. Since data propagation allows us to track the growth of the number of people 'influenced' by an action or an idea, then we can assume that by investigating data propagation, we can track the growth, or decay, of power.

Several studies have been tracking data propagation on the Web. Different objects have been tracked, for example data flows along hyperlinks [1, 17] and text phrases [10, 16]. Different domains have been looked at, for example the flow of topics from mainstream news to the blogosphere [16] and the way e-mail chain letters are spread [18].

One potential approach we could use to measure power is to track the propagation of objects such as tweets onto tractable actions such as e-petitions. Although Choudhury et. al. [7] had already presented how blog activity can correlate to fluctuations in the stock market, there seems to be a lack of experimentation on correlations which involve persuading people to support a cause.

Building on the work that we have discussed in Section 5, we are constructing a toolkit that allows users to trace a variety of objects and correlate them to tractable actions. This toolkit uses the BackType API¹ to search for tweets related to phrases or links to a particular URL. The advantage of this API over the built-in search facility in the Twitter API² is that it also resolves shortened URLs. Once the related tweets are found, then the Twitter API is used to access tweet-specific data such as followers lists. Apart

¹<http://www.backtype.com/developers>

²<http://apiwiki.twitter.com/>

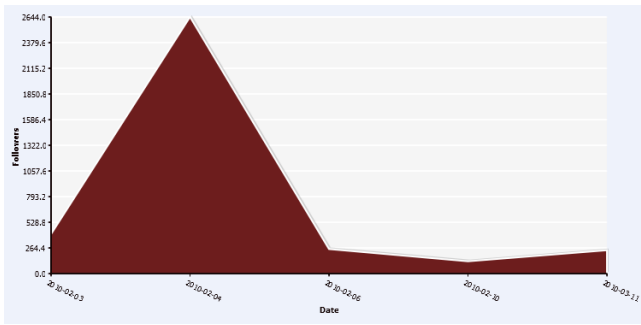


Figure 1: Number of followers that have seen a tweet linking to a particular e-petition (per day)

from tracing the propagation of phrases and links to particular URLs, one could also see the number of followers exposed to a particular tweet promoting an e-petition. Figure 1 demonstrates the initial output of this toolkit.

The second phase of this toolkit’s development is to correlate these propagation traces with tractable actions. For example, we could investigate any correlations between the number of followers who have seen certain e-petition tweets and the total signatures that have been collected within a particular timeline. This phase is still under development.

8. CONCLUSION

This paper has identified how power in the blogosphere is of interest in Web Science, citing cases like Senator Trent Lott’s resignation and Dan Rather’s scandal. These cases echo pamphlets from the 17th century, where they were used to bypass institutional controls and cause collective action amongst the public.

Although the definitions of power itself has been subject to debate, this paper defines it as the ability to produce effects among others within a decision-making process. Prior work in this area has relied on metrics such as links and phrases to rank blogs. The difficulty of observing more subtle effects of power means that any research work on power is constrained to studying concrete, tractable observations.

Nonetheless, the use of data propagation could be a suitable approach. One potential method to measure power is to correlate blog trends to a tractable action like an e-petition. Despite the limitations that we have discussed, it may still be possible to find meaningful empirical analysis of power in the blogosphere.

9. REFERENCES

- [1] E. Adar, L. Zhang, L. A. Adamic, and R. M. Lukose. Implicit structure and the dynamics of blogspace. In *Workshop on the Weblogging Ecosystem, WWW2004*, New York, NY, 2004.
- [2] N. Agarwal, H. Liu, L. Tang, and P. S. Yu. Identifying the influential bloggers in a community. In *Proceedings of the international conference on Web search and web data mining*, pages 207–218, Palo Alto, California, USA, 2008. ACM.
- [3] M. Bragg. *In Our Time – A Modest Proposal*. Radio podcast, *BBC*, <http://www.bbc.co.uk/programmes/b00h3650>, 2009.
- [4] D. Bricklin. Pamphleteers and web sites. Weblog post, <http://www.bricklin.com/pamphleteers.htm>, 2001.
- [5] A. Burns and B. Eltham. Twitter free iran: an evaluation of Twitter’s role in public diplomacy and information operations in Iran’s 2009 election crisis. In *Communications Policy & Research Forum 2009*, University of Technology, Sydney, 2009.
- [6] R. M. Cardillo. I am Publius, and I approve this message: The baffling and conflicted state of anonymous pamphleteering post McConnell. *Notre Dame Law Review*, 80(5):1929–1960, 2004.
- [7] M. D. Choudhury, H. Sundaram, A. John, and D. D. Seligmann. Can blog communication dynamics be correlated with stock market activity? In *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, pages 55–60, Pittsburgh, PA, USA, 2008. ACM.
- [8] E. L. Eisenstein. *The Printing Press as an Agent of Change*. Cambridge University Press, 1980.
- [9] H. Farrell and D. Drezner. The power and politics of blogs. *Public Choice*, 134(1):15–30, 2008.
- [10] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins. Information diffusion through blogspace. In *Proceedings of the 13th international conference on World Wide Web*, pages 491–501, New York, NY, USA, 2004. ACM.
- [11] J. Hender, N. Shadbolt, W. Hall, T. Berners-Lee, and D. Weitzner. Web science: an interdisciplinary approach to understanding the web. *Comms. ACM*, 51(7):60–69, 2008.
- [12] R. Klau. Rick Klau’s weblog: Topical, polemical, and short. Weblog post, <http://tins.rklau.com/2003/04/topical-polemical-and-short.html>, 2003.
- [13] R. Klau. Rick Klau’s weblog: Twitter: Topical, polemical and short. Weblog post, <http://tins.rklau.com/2009/03/twitter-topical-polemical-and-short.html>, 2009.
- [14] D. J. Kochan. The blogosphere and the new pamphleteers. *Nexus Law Journal*, 11:99–109, 2006.
- [15] K. S. Lee. Hiding from the boss online: The Anti-Employer blogger’s legal quest for anonymity. *Santa Clara Computer & High Technology Law Journal*, 23:135, 2006.
- [16] J. Leskovec, L. Backstrom, and J. Kleinberg. Meme-tracking and the dynamics of the news cycle. In *Proceedings of 15th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, 2009*, Paris, France, 2009. ACM.
- [17] J. Leskovec, M. Mcglohon, C. Faloutsos, N. Glance, and M. Hurst. Cascading behavior in large blog graphs. In *SIAM International Conference on Data Mining*, 2007.
- [18] D. Liben-Nowell and J. Kleinberg. Tracing information flow on a global scale using internet chain-letter data. *Proceedings of the National Academy*

- of Sciences*, 105(12):4633–4638, 2008.
- [19] S. Lukes. *Power: A Radical View*. Palgrave Macmillan, 2nd edition, 2005.
- [20] S. E. Malloy. Anonymous bloggers and defamation: Balancing interests on the Internet. *Washington University Law Review*, 84:1187, 2006.
- [21] P. Morriss. *Power: a philosophical analysis*. Manchester University Press, 2002.
- [22] M. Munger. Blogging and political information: truth or truthiness? *Public Choice*, 134(1):125–138, 2008.
- [23] B. Russell. *Power: A New Social Analysis*. Routledge, new ed edition, 1975.
- [24] X. Song, Y. Chi, K. Hino, and B. L. Tseng. Information flow modeling based on diffusion rate for prediction and ranking. In *Proceedings of the 16th international conference on World Wide Web*, pages 191–200, Banff, Alberta, Canada, 2007. ACM.