

Do you know what bloggers are saying about your business?

It's no secret that private bloggers tend to speak freely in their conversations about firms and their products. But how can firms tap into blogs and build their Business Intelligence?

Blogs are becoming more integrated into the digital marketing activities of firms

In 2013 Technorati found blogs to be the third most influential digital resource, after retail and brand websites, for consumers making purchase decisions. In 2014, the Social Media Marketing Industry Report identified blogging to be the 5th most commonly used social media platform. It was also the 4th most important social platform for marketers, after Facebook, LinkedIn, and Twitter. Indeed, 68% of these marketers indicated they planned to increase their blogging activities as part of their social media marketing mix.

Like other social media tools, blogs can bring firms closer to their consumers. They provide channels for feedback and comments about their products, user experiences, and promotional campaigns. As part of the overall social media mix, McKinsey's identify that they also assist firms in creating high levels of positive marketing buzz when launching their products and services – without significant expense!

Personal blogs are important to businesses

WordPress estimates that over 409 million people view more than 19.1 billion blog pages each month. And personal blogs far outnumber corporate blogs. Corporate blogs do provide firms with the ability to deliver publicity, promotions and communications activities in a more personal way than via a website. However, personal blogs are unique in the sense that people tend to freely express their emotions and their opinions in consumer-based discussions. Personal bloggers can therefore become involved in deeper discussions about their consumer experiences relating to a firm's products or services, and also the touch points relating to purchase interactions. Bloggers' comments have also been found to provide insights into the effectiveness of a firm's competitive positioning, its media relations, and even how it might better build the brand.

With these factors in mind, Dr. Michael Chau, a business analytics expert from HKU's Faculty of Business and Economics explains another important point. "As blogger interactions develop, bloggers tend to see each other as potential collaborators or as friends, forming a type of 'collective intelligence', arriving at deeper knowledge about a topic". Indeed, that Nielsen finds 92% of consumers place recommendations from friends and family ahead of all forms of advertising, reputable personal bloggers have the potential to become powerful opinion leaders within the blogosphere. The challenge for firms, however, relates to how to acquire relevant information from personal blogs.

Navigating the blogosphere to acquire business intelligence (BI)

The unique characteristics of blogs can present major challenges for collecting blog data, evaluating blog content, and analysing the underlying social networks. As Dr. Chau highlights, "blogs are very dynamic in that content can be frequently updated or removed. Bloggers also apply their own styles when linking or interacting with each other. The sheer volume of blog data also presents a challenge".

Bloggers interact through subscribing to other blogs, posting comments, and citing other blogs and bloggers within their own blog posts. These social interactions take place in

blogger communities. Bloggings form through explicit interactions where bloggers with a similar interest or background join each other to form groups. On the other hand, implicit or informal groups tend to form organically through ongoing interactions between bloggers. As interactions amongst bloggers develop, relationships tend to form and manifest themselves into virtual networks of bloggers. As Dr. Chau explains, “relationships in blogger communities tend to be characterised by memberships, a sense of belonging, shared values, and self-regulation”. In particular, the hidden or implicit structures that manifest from the blogger interactions need to be identified and studied to locate key players in the networks. These bloggers can serve as important opinion leaders and/or communication channels or bridges between different blogger communities. Navigating through the complexity of blogger interactions requires the application of automated BI techniques.

How to gather BI in the blogosphere

Dr. Chau and his colleague Dr. Jennifer Xu from Bentley University set out to tackle some of the issues related to gathering BI from blogs. Specifically, they designed a framework for generating BI from blog data. The framework was then applied to gather and analyse data from two sets of consumer-based blog discussions; one relating to iPods, and the other to consumer experiences at Starbucks. The following discussion focuses on Blog BI conducted for iPods; information about Blog BI on consumer experiences at Starbucks can be found in the paper by Dr. Chau and Dr. Xu: “Business Intelligence in Blogs”¹.

❖ Identify the explicit communities of interest within the blogosphere

Blog communities talking about iPods were identified from Xanga, which was the second most popular blog hosting site after Google’s Blogger. Specifically, web crawler software was applied to retrieve information about blog memberships. Bloggings that contained the word iPod in their titles or descriptions were identified. After removing blogs with single members, 204 valid groups were identified and groups were classified as positive, neutral, or negative towards iPods. The data set included 3,493 bloggers, and 75,445 blog entries. Blog communities were categorised based on the attitude of the group the blogger belonged to. The data identified 2,377 bloggers with a positive attitude toward the iPod, 225 with a negative attitude, and 891 were neutral.

❖ Collect information about the characteristics of bloggers in the explicit communities

The member lists from the explicit communities provided a massive amount of data. Blog crawler software was therefore designed to extract specific information from bloggers’ profiles, blog entries and comments pages. After re-examining the blogs’ contents for their relevancy to the iPod, the frequency with which the word iPod was mentioned in each blog was identified. ‘Splogs’ or spam blogs were filtered out and of the remaining legitimate blogs the one with the highest frequency mentioned iPods 115 times. Interestingly, it was found that only a small percentage (less than 2%) of bloggers used the name iPod frequently. 1,920 of the bloggers never mentioned the word iPod in their blogs! As Dr. Chau expressed, “the important point here is that these bloggers cannot be identified through a standard keyword search; they are only identifiable through their group memberships or bloggings”.

❖ Analyse the posted content, identify the networks, and the central bloggers

Written text from the blog pages were classified and clustered. Sentences containing key words were extracted, and topics and opinions were analysed. For example, neutral comments about how users interact with their iPod were identified, such as “I downloaded

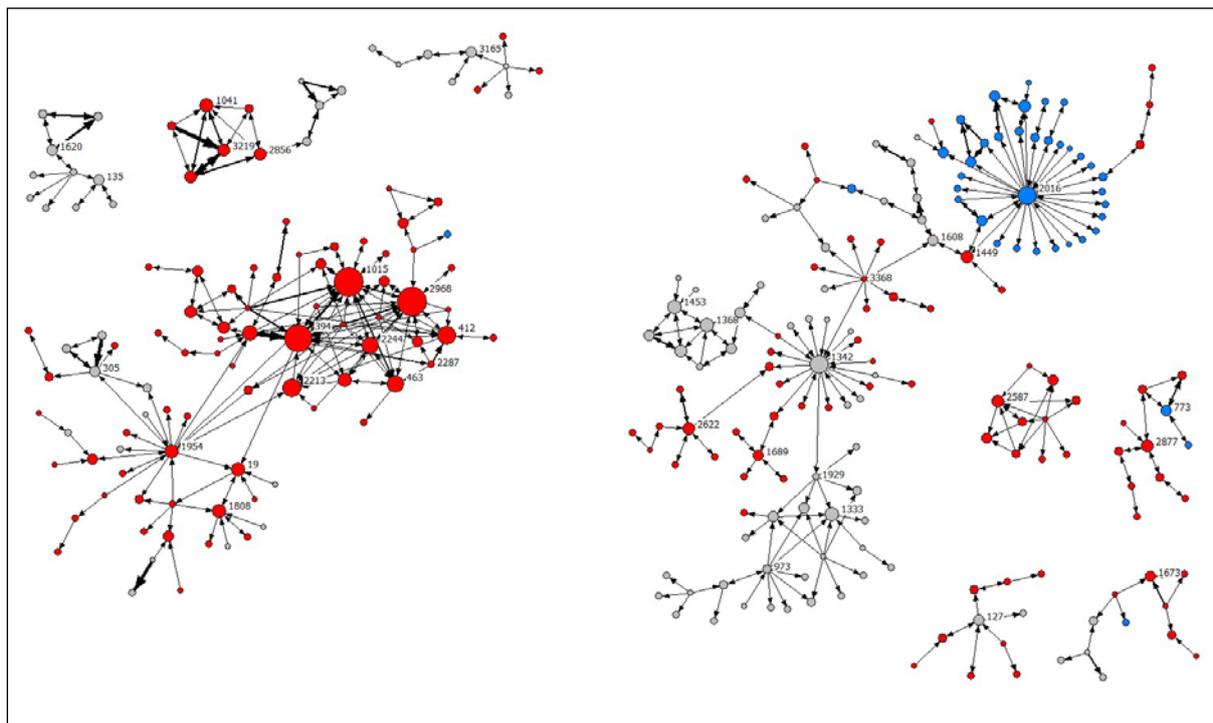
some new songs on my iPod”. Positive and negative comments tended to relate to blogger’s opinions of iPods, such as, “iPods are awesome”, and “I hate iPods”.

Three networks were identified; a subscription network, where bloggers subscribed to another’s blog; a comment network, where bloggers leave comments on another’s blog; and a combined network, where the links from both the subscription network and the comment network are included.

Central bloggers, or small numbers of bloggers involved in larger numbers of interactions were also identified. Their direct interactions, including incoming links and outgoing links within the networks were examined. A high degree of incoming subscription links indicates the blogger is popular or authoritative. Their comments are interesting and people like keeping track of their updates, so they tend to be liked or endorsed by other bloggers. Bloggers who are high out-degree subscribers tend to act as ‘hubs’ directing visitors to other interesting blogs they are connected to. Bloggers with a high in-degree of comments from many bloggers and high out-degree of posting comments on many blogs are the most powerful communicators within the blogosphere. To test the capability of the central bloggers to disseminate information, a simulated marketing message was ‘seeded’ 100 times through the networks. The tests confirmed that the top ten central bloggers identified, on average, achieved higher effectiveness in disseminating the message within their blogger networks.

❖ Identify the implicit blogger communities

Implicit blogger communities within the networks were identified through extracting and matching traces of links from bloggers’ subscription and comment interactions. Clusters of positive (red), negative (blue), and neutral (grey) bloggers were identified. For example, the analysis identified the existence of 292 clusters in the comment network; the largest cluster contained 124 members – 11.2 % of the total members in the network. Within the clusters of negative bloggers (indicated by the blue circles), the lead blogger was identified. This blogger had left comments on many of the other negative bloggers’ sites.



Major Clusters in the Comment Network in the iPod Dataset

Dr. Chau's final comment 'appears to say it all': "I would be surprised if Apple would not want to know more about why these consumers feel the way they do about the iPod!"

The paper by Dr. Chau and Dr. Xu is available at: http://www.business.hku.hk/~mchau/papers/BusinessIntelligenceInBlogs_MISO.pdf