A SOCIAL NETWORKING ANALYSIS OF WINE BLOGS

Vasiliki G. Vrana
Technological Education Institute of Serres
Serres, Greece
e-mail: vrana@teiser.gr

and

Kostas V. Zafiropoulos
University of Macedonia
Thessaloniki, Greece
e-mail: kz@uom.gr

ABSTRACT

Wine-blogs are important for the information provision and promotion of wine industry and tourism activities related to wine. The paper records 1305 wine-blogs and their inter-linkages through their blogrolls. It uses Social Network Analysis to study blogs’ networking characteristics and connectivity. The study proposes five graph-theoretic indexes and explores whether wine-blogs are making use of the features that Web2.0 offers, or they are just a collection of users’ websites with limited “social” or “networking” characteristics. The property of skewness is discussed. Findings show that only few wine-blogs are highly connected. However, they can be regarded as influential within the bloggers’ community.

Key Words: wine-blogs, social network analysis, connectivity, influence

INTRODUCTION

The potential of the web as a commercial medium and market has been widely documented (Kiang et al., 2000). However, there are certain product categories on the web, which require more information for people to make a purchase (Leskovec, et al., 2007). Tourism products fit this category as they can hardly be evaluated prior to their consumption (Rabanser & Ricci, 2005) and depend on accurate and reliable information (Kaldis et al., 2003), thus elevating the importance of interpersonal influence (Lewis & Chambers, 2000). Wine, fits also this category due to its intrinsic characteristics, namely it is a product of “experience” and is highly differentiated and marketed internationally (Stricker et al., 2003). Nowadays, the wine industry has adopted internet and myriads of websites exist for wine related issues on the web, promoting either wineries, wine-shops, wine regions or wine portals (Nóvoa, 2010). Wine related enterprises have begun to use innovative tools and to develop managerial abilities to face new challenges coming from the web (Annunziata et al., 2008). Top wine country producers have adopted wine portals as a first level entry point to information (Nóvoa, 2010). Blogging in the wine industry is also widespread. Santos (2011) suggests that international wine bloggers are a niche community, relatively young and dynamic in the world of blogs. Technorati.com, the most popular real-time search engine dedicated to the blogosphere tracks on 12-09-2011, 15535 blogs involving food, beverage and food-related fields. This category of blogs is ranked 3rd among the top 18 topics on blogging (Technorati.com, 2008). However, little research effort has been devoted to investigating gastronomy and beverage blogs (Wang, 2011). Previous research focuses on revealing types of wine blogs, topics of discussion (Annunziata et al., 2008; Thack, 2009; Thack, 2010; Vrana et al., 2011), critical roles of blogs in predicting readers’ intention to taste local food and beverages (Wang, 2011) and examination of the profile, motivations and several technical and marketing features of international wine bloggers (Santos, 2011). What has not been investigated yet is wine blogs’ influence, networking characteristics and blogs’ connectivity.

Social influence, “describes the phenomenon by which the behavior of an individual can directly or indirectly affect the thoughts, feelings, and actions of others in a population” (Cialdini 2001; Song et al., 2007 p. 971). In this vein, social influence is found within the blogs social network community (Tan et al., 2011). It is interesting to explore which are the blogs that have an influence both to bloggers and the readers and users of blogs. The paper explores which blogs are most influential using Social Network Analysis. It uses five graph theoretic indexes: the number of incoming links, the number of outgoing links, normalized betweenness, the number of 1-cliques that a blogs belongs to, and the size of blog’s ego-network.

WINE BLOGS

Importance of Wine Blogs
Wine blogs are “interactive websites in which a blogger writes wine reviews and informational or opinion pieces about wine, and encourages readers to type their responses to the blog so that others may read and respond as well” (Thach, 2009 p. 144). Generally, blogs are personal online journals. Bloggers personalize and actualize content and information online and present their views to a broad audience. Blogs have the power to engage people in collaborative activity, knowledge sharing, reflection and debate (Efimova et al., 2005; Punie & Cabrera, 2005). They are used to share consumption experiences (Matsumura et al., 2008) and have tremendous impact on the decision making behavior of Internet users (Sigala, 2007) as people take into consideration others’ opinions when making a purchase decision (Kozinets, 2002). In the tourism sector many travelers tend to use blogs for searching for travel information, tips, selecting travel suppliers and destinations (Sigala 2009; Wang, 2011) and travel planning (Akehurst, 2009; Litvin et al., 2008). Blogs are also ideal for experts to broadcast their expertise to a large audience (Wagner & Bolloju, 2005). In this vein everyday consumers and wine writers establish their wine blogs to share the passion for wine, to share their knowledge, to provide wine news, to review wines and provide ratings and share their wine-travel experiences (Thach 2009, Zafiropoulos et al. 2010). According to Santos (2011) wine blogs can be used independently by people to say some provocative things about wine (ratings, restaurants, food, etc.) or can create online communities of people interested in share thoughts, activities and reactions about wine.

Blogs are not only employed in personal environments but also in organizations and enterprises (Kolbitsch and Maueur, 2006) and as they are growing in popularity, businesses and organizations are looking for ways to exploit them. Wineries, marketers and retailers and wine businesses, establish their wine blogs to promote business, to promote and sell wine and wine related products (Thach 2009). This use of blogs may have important implications. For example, Stormhoek Vineyards, a small winery in South Africa tripled its sales in 2 years by using a wiki and blogs to create groups for wine tasting parties (Lai & Turban, 2008).

McMillan (2007) commenting on the U.S. wine blogs wrote “The impact of these bloggers is growing because they are viewed as offering objective community-driven commentaries on wine at a time when the community of established traditional wine writers can’t possibly cover all the fine wine being made”. Recently, Thach (2010, p.2) mentioned “wine blogs are not monitored and there are no official guidelines or rules regarding what can be published – therefore, there are many diverse opinions about wines and wine brands. Because of this some bloggers may write either positive or negative reviews about wine that can help or hinder wine sales”.

Previous research on wine blogs

In the “complete list of wine blogs” Yarrow (2010) recorded 882 wine blogs which cover many different aspects. The majority of wine blogs are written in English. Annunziata et al. (2008) classify wine-related blogs in: blogs as a tool for enthusiastic admirers, which give the writers an opportunity for storytelling, giving opinions and sharing information; and company blogs where the producer and the community of users communicate. According to Thach (2010) nine major types of wine blogs exist: “review of wines”, “wine and food”, “wine education”, “winemaking & viticulture”, “specific region”, “wine & culture”, “winery blog”, “wine business” and “other”. Vrana et al. (2011) investigated the topics discussed in wine blogs. Wine reviews are the main common topic. Wine dining, wine publications, wine tasting, and restaurant review are topics highly discussed followed by wine news, wine making, wine event, wine tourism, geoponic, cheap wines, wine education, images, wine marketing, wine packaging, book reviews, culture and wine, wine commentary, recommendations, and wine serving. In a recent study, Santos (2011) shed light on blogging practices and motivations. The study examines the profile, motivations and technical and marketing features of international wine bloggers. The findings indicate that wine bloggers tended to blog for self-promotion rather than diversion, using Wordpress platform, Google analytics counter and Creative Commons license. Furthermore, the study indicates the minor role of advertisement and the practice of rating wines as a strategy to increase the profitability of the blog.

BLOG INFLUENCE

The blogosphere is conversational (Song et al., 2007) consisting of millions of individual blogs. Users interact through the unique technological capabilities and enhanced blogging tools for between-blog features interactivity (Woodly, 2008) and form communities, sub-communities and cliques (McGlohon et al., 2007). Within any community there may be some particularly prominent members who start major conversations and there may be others who are more active in gathering content from many conversations and only a few that attract a large readership (McGlohon et al., 2007; Wagner & Bolloju, 2005). The important role of these blogs has attracted growing attention, as they are more influential than others. A question that arises is how do we measure influence?
Previews studies have taken different approaches in order to identify influential bloggers. Blog features, in-links (incoming links) and out-links (outgoing links) (Adar and Adamic, 2005; Agarwal and Liu, 2008) have been used in order to detect influence. The overall distribution of in-links between blogs is highly unequal (Drezner & Farrell, 2008; McGlohon et al., 2007; Zafiropoulos and Vrana, 2010). However, the authors of the most well-known and read blogs manage to make themselves a “celebrity” among the community of bloggers. These blogs are regularly the most linked by others (Trammell & Keshelashvili, 2005; Ali-Hasan & Adamic, 2007). In this vein the median blogger has almost no influence while a few “elite” blogs can operate as both an information aggregator and as a “summary statistic” for the blogosphere. Agarwal (2008) and Agarwal et al. (2008) claimed that identifying the influential bloggers at a blog site requires the integrated use of the information specific to a blog site that is out-links, in-links and comments. Their preliminary model takes into consideration a set of four properties: 1. Recognition - An influential blog post is recognized by many. 2. Activity Generation - A blog post’s capability of generating activity can be indirectly measured by how many comments it receives the amount of discussion it initiates. 3. Novelty - Novel ideas exert more influence. Hence, the number of out-links is an indicator of a post’s novelty. 4. Eloquence - An influential is often a response to the author’s post. Total Comments/Week was used to serve as a site’s Community Activity Score. As at this point only data from the blogs’ network are available and no data for the users’ community, this particular study is limited to study bloggers’ popularity within the context of bloggers community. When many blogs link to a particular blog, this blog has a high degree of incoming links. It can be argued that highly linked blogs enjoy some appreciation or are recommended by the bloggers’ community. Thus, heavily linked blogs may be regarded to be influential in the sense that their comments are read and recommended by others. Originating from this idea, this paper expands the analysis in the following way: It studies blogs’ community characteristics and the degree of blogs’ connectivity. That is the paper studies how wine blogs are organized in small communities along with other blogs, expanding in this way their territory of influence. This property is not

**METHODOLOGY**

This research used the Wine Blog Search Engine”, and initially recorded the Top 100 Wine Blogs (http://alawine.com/wine-blog-rankings.html). Wine blog ratings in Top100 list are standardized composite scores based on multiple relative link popularity rankings from three top search engines and Technorati.com, as well as Google page rank scores. Top100 blogs served as starting points and through their blogrolls new blogs were found. Next, this research used snowball sampling, starting with the top 100 wine blogs, blogrolls of these newly found blogs were used to locate new blogs etc. The procedure resulted to a record of 1305 wine blogs. Sampling took place during June - July 2010. Along with the blogs, the links among them were recorded. Defining which blog is a wine-blog can be a really complex matter, regarding the snowball sampling. For this research, blogs posting frequently exclusively on wine or wine and food, but not on other topics, were considered eligible to be included in the overall data set. Links were represented by a graph describing the social network of wine blogs. This network is associated with a 1305 by 1305 non-symmetric binary adjacency matrix. The network adjacency matrix has been used by Zafiropoulos et al (2010) and Vrana et al (2010) for the study of wine-blogs, and by Vrana & Zafiropoulos (2010) for the study of travelers’ blogs. Influential blogs have the potential to address many visitors and it can be argued that they can determine the flow of information to users and other blogs as well. Social Network Analysis (SNA) can be used to study blogs connectivity and popularity with regards to inter-linkage of blogs within the wine blogs community.

As at this point only data from the blogs’ network are available and no data for the users’ community, this particular study is limited to study bloggers’ popularity within the context of bloggers community. When many blogs link to a particular blog, this blog has a high degree of incoming links. It can be argued that highly linked blogs enjoy some appreciation or are recommended by the bloggers’ community. Thus, heavily linked blogs may be regarded to be influential in the sense that their comments are read and recommended by others.
necessarily directly linked to the property of centrality as measured by the incoming links. Rather it is about the individual blogger’s effort to establish a network of closed blogs or bloggers, those who probably share common characteristics or interests. The paper uses five Social Network Analysis indexes to measure the abovementioned properties:

1. Number of incoming links of a blog. It represents the number of blogs that link to a specific blog and it can be considered to be a kind of index of recommendation.
2. Number of outgoing links. It represents the number of blogs that a specific blog links to. It can be considered to be a blog connectivity index regarded as a means to reach the blogs’ community.
3. Normalized betweenness. The centrality of a node in a network is a measure of the structural importance of the node. A person’s centrality in a social network affects the opportunities and constraints that they face. Betweenness centrality is defined as the number of geodesic paths that pass through a node. It is the number of "times" that any node needs go through a given node to reach any other node by the shortest path. The node with high betweenness can serve as a liaison between disparate regions of the network. Betweenness is therefore a measure of the number of times a vertex occurs on a geodesic (Zafiropoulos & Vrana 2011). The normalized betweenness centrality is the betweenness divided by the maximum possible betweenness expressed as a percentage.
4. Number of 1-cliques that a blog belongs to. An 1-clique is a maximal subgraph which contains blogs which are linked directly to each other. That is, for a given blog the 1-clique groups contains all the blogs that this particular blog is connected to, and also the blogs of the 1-clique are connected with each other, regardless if the linkages refer to incoming or outgoing links. So, an 1-clique contains the closest blogs to a particular blog, which also are linked to each other. The number of 1-cliques that a particular blogs belongs to, is a measure of a blog’s network connectivity and belonging to certain families of blogs. It should be noted that the set of 1-cliques used in the analysis, contains only 1-cliques with at least three blogs, that is it does not contain pairs of connected blogs.
5. Ego-network size. For a specific blog, an ego-network contains all the blogs that are directly linked to this blog. The size of an ego-network is the number of blogs in the ego-network. Again as in the case of 1-cliques this number is a measure of a blog’s and community connectivity.

In the analysis that follows all the above indexes are calculated. Normalized Betweenness, ego-network size, and number of 1-cliques, are calculated using UCINET for Windows, while number of incoming links, number of outgoing links are calculated as the sum of columns and rows, respectively, from the blog network adjacency matrix. In order to calculate ego-network sizes and 1-cliques the adjacency matrix is symmetrized, that is 1 is placed in cell i,j either if i links to j blog, or j links to i blog. Then the property of skewness is discussed. For a certain SNA index, skewness refers to the property that a minority of blogs present high values of the index, while the majority has only small values. For example, regarding number of incoming links, skewness implies that only few blogs have many incoming links and the majority has only few incoming links.

**FINDINGS**

For the 1305 wine blogs the five indexes are calculated. Regarding 1-cliques formation, 32% of the blogs do not belong to any 1-cliques. On the other hand the maximum number of 1-cliques that a blog belongs to is 11,406. Seventeen blogs belong to over 1,000 1-cliques and thirty blogs belong to over 500 1-cliques. Fermentation.typepad.com is a member of 11,406 1-cliques, while vinography.com, goodgrape.com, wildwallawallawinewoman.blogspot.com, goodwineunder20.blogspot.com and passionatefoodie.blogspot.com belong to more than 2,000 1-cliques each. Table 1 presents some descriptive statistics for the proposed indexes. One can conclude that there is a majority of wine blogs that are not really networked (with regards to 1-cliques formation), but on the other hand few wine-blogs (presented above), are extremely connected and involved within the community.

The Ego-network size ranges from zero to 449, with a mean equal to 22.6 and a standard deviation equal to 35.2. Five blogs have ego-network size greater than 200: fermentation.typepad.com, vinography.com, drvino.com, goodwineunder20.blogspot.com, goodgrape.com. Fifty percent of the blogs have ego-network size less than 12 and 75% of the blogs have size less or equal to 25. 25% of the blogs have the highest ego-network sizes, which reach the maximum value 449.

Descriptive statistics for the incoming and outgoing links are also presented in Table 1. The averages of incoming and outgoing links both equal 13. Standard deviations are almost 25 for both, while some blogs have no incoming or outgoing links, and some blogs have up to 334 outgoing and 343 incoming links.

Normalized betweenness is a graph theoretic index associated with centrality of the network. The greater the normalized betweenness the more central a blog is. Table 1 shows that mean normalized
betweenness equals 0.1 and the standard deviation is 0.4. Because standard deviation equals 0.4, which is greater than the mean, one can conclude that there is a diversity among blogs regarding centrality. There exist blogs with normalized betweenness equal to zero, while the index increases up to 10.76.

### Table 1
Descriptive statistics of the five proposed SNA indexes

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incoming links</td>
<td>0</td>
<td>343</td>
<td>12.9</td>
<td>25.9</td>
</tr>
<tr>
<td>Outgoing links</td>
<td>0</td>
<td>334</td>
<td>12.9</td>
<td>23.1</td>
</tr>
<tr>
<td>Normalized betweenness</td>
<td>0</td>
<td>10.76</td>
<td>0.1</td>
<td>0.4</td>
</tr>
<tr>
<td>Number of 1-cliques that a blog belongs to</td>
<td>0</td>
<td>11406</td>
<td>65.2</td>
<td>404.7</td>
</tr>
<tr>
<td>Size of Ego-network</td>
<td>0</td>
<td>449</td>
<td>22.6</td>
<td>35.2</td>
</tr>
</tbody>
</table>

All of the above proposed indexes seem to behave similarly regarding the property of skewness. Only few wine-blogs have the highest values of the indexes. Table 2 presents the correlation coefficients of the five proposed indexes. From Table 2 it becomes clear that all the correlations among the indexes are high and statistically significant. This property of indexes’ inter-correlations is of the essence for the graph-theoretic study of a network. Conclusions are further supported in such a case of strong index inter-correlations, and there is also strong evidence that all the indexes measure network centrality and blogs connectivity.

The property of skewness is important and a main finding of this research. In order to make the property of skewness clear, it would be useful to present the distributions of the proposed indexes and the overall index. Table 3 presents the percentiles of these distributions. One can see from Table 3 that 70% of the wine-blogs of the study have up to 10 incoming links, 12 outgoing links, very small normalized betweenness (up to 0.02), they belong up to 14 1-cliques, and their ego-network size is up to 21. Values grow rapidly at the largest percentiles, and this is due to the property of skewness. The highest values are present after the 90% percentile for all the proposed indexes.

### Table 2
Correlations matrix of the five SNA indexes

<table>
<thead>
<tr>
<th></th>
<th>Incoming links</th>
<th>Outgoing links</th>
<th>Normalized betweenness</th>
<th>Number of 1-cliques that a blog belongs to</th>
<th>Size of Ego-network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outgoing links</td>
<td>0.418*</td>
<td>0.615*</td>
<td>0.610*</td>
<td>0.888*</td>
<td>0.677*</td>
</tr>
<tr>
<td>Normalized betweenness</td>
<td>0.717*</td>
<td>0.586*</td>
<td>0.794*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of 1-cliques that a blog belongs to</td>
<td>0.731*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of Ego-network</td>
<td>0.677*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(*: p<0.01)

### Table 3
Percentiles of the proposed indexes

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Incoming links</th>
<th>Outgoing links</th>
<th>Normalized betweenness</th>
<th>Number of 1-cliques a blog belongs to</th>
<th>Ego-network size</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>60%</td>
<td>7</td>
<td>8</td>
<td>0.01</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>70%</td>
<td>10</td>
<td>12</td>
<td>0.02</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>80%</td>
<td>16</td>
<td>19</td>
<td>0.07</td>
<td>29</td>
<td>31</td>
</tr>
<tr>
<td>90%</td>
<td>31</td>
<td>34</td>
<td>0.17</td>
<td>92</td>
<td>55</td>
</tr>
<tr>
<td>100%</td>
<td>343</td>
<td>334</td>
<td>10.76</td>
<td>11406</td>
<td>449</td>
</tr>
</tbody>
</table>

### CONCLUSIONS

The paper proposed a method for measuring wine-blogs connectivity and influence. Social Network Analysis provides some tools and indexes for this purpose. Blogs’ network connectivity is not limited only to the study of incoming and outgoing links but to the ability of blogs to form spheres of influence through the formation of interlinked and connected groups of blogs. Ego-networks and n-cliques are suitable mathematical entities for the study of such issues. The more holistic approach employed in this paper allows finding the most networked blogs. Such blogs have the power to address more peer bloggers and may be considered as influential. Studying these blogs and their content may be important and useful for both researchers and practitioners, since by locating them, one can be informed and analyze discussions of the most influential blogs.
This might be very useful particularly for small unknown wineries and regions. Less known wineries’ blogs and marketers in general, may contact active wine-blogs, to publicize their wines and regions interested in wine tourism may contact these blogs in order to attract visitors. Researchers, on the other hand, can follow the information flow of these blogs, and understand what are the most popular issues discussed in the blogosphere and what are the marketing strategies used to promote certain products related to the wine industry.

The findings of this paper are in accordance with relative findings regarding travel blogs. Vrana & Zafiropoulos (2010) and Zafiropoulos et al. (2010) have shown that travel blogs present similar patterns. Studying the travel blogs of Travelpod.com, the authors found that only few travel blogs are well connected and can serve as focal points, where information provision might be larger than in the rest of the blogs, since the likelihood that these blogs are visited is greater. In conclusion, skewness of connectivity characteristics seems to be a common property of the overall blog landscape of tourism. The present study followed the Social Network Analysis approach and a specific sampling procedure to locate and analyze wine-blogs. The analysis could be expanded by using complementary social network analysis indexes or other approaches such as the study of the content, the authority of the wine-blogs within the blogosphere and the community activity (for example by studying the comments of the wine-blogs posts). These techniques of further research may provide validation, reveal the limitations of the present findings or even offer new ways of assessing influential wine-blogs.

REFERENCES


