www.ejthr.com

European Journal of Tourism, Hospitality and Recreation
Vol. 3, Issue 1, pp. 135-156, 2012

Printed in Portugal

WINE-BLOGS INFLUENCE AND BLOGS' COMMUNITY CONNECTIVITY: A social network analysis

Kostas Zafiropoulos
University of Macedonia, Greece

ABSTRACT: Wine-blogs are important for the information provision and promotion of wine industry and tourism activities related to wine. The paper records 1,305 wine-blogs and their inter-linkages through their blogrolls. It uses Social Network Analysis to study blogs' networking characteristics and connectivity. By proposing five graph-theoretic indexes, and by summarizing them to one overall index, the study explores which blogs are most active and connected. The property of skewnness is discussed. According to the overall index, the top-10 wine-blogs are described. Highly connected wine-blogs can be regarded as influential within the bloggers' community. Locating them may be significant since it allows for the study of information flow to readers. **Keywords:** wine-blogs, social network analysis, principal components analysis, connectivity, influence.

RESUMEN: Los blog sobre vinos son importantes como fuentes de información y como medios de promoción del sector vinícola y del enoturismo. Este artículo se ocupa de 1305 blogs sobre vinos y sus interrelaciones a través de los respetivos listados de blogs. Es utilizado el Análisis de Redes Social (*Social Network Analysis*) para investigar las características de rede y conectividad de los blogs. A través de la creación de cinco indexes gráfico-teóricos y de su aglutinación en un índex global, el estudio identifica los blogs más activos y con más conexiones. La propiedad de distorsión (*property of skewnness*) es discutida. Son descritos los diez blogs de vinos con mejor puntuación en el índex global. Los blogs de vinos fuertemente conectados pueden ser considerados como medios influyentes en la comunidad de bloguistas. Su localización puede ser importante, una vez que permite conocer los flujos de información que llegan a los lectores. **Palavras clave:** blogues de vinos, análisis de redes sociales, análisis en componentes principales, conectividad, influencia.

RESUMO: Os blogues sobre vinhos são importantes como fontes de informação e como meios de promoção do setor vinícula e do enoturismo. Este artigo abrange 1305 blogues sobre vinhos e as suas interligações através das respetivas listas de blogues. É utilizada a Análise de Redes Social (Social Network Analysis) para investigar as características de rede e conetividade dos blogues. Através da criação de cinco indexes gráfico-teóricos e da sua aglutinação num índex global, o estudo identifica os blogues mais ativos e com mais conexões. A propriedade de distorção (property of skewnness) é discutida. São descritos os dez blogues de vinhos com melhor pontuação no índex global. Os blogues de vinhos fortemente conetados podem ser considerados como meios influentes na comunidade de bloguistas. A sua localização pode ser importante, dado que permite conhecer os fluxos de informação que chegam aos leitores. **Palavras chave:** blogues de vinhos, análise de redes sociais, análise em componentes principais, conetividade, influência.

Kostas Zafiropoulos is an Assistant Professor at the Department of International and European Studies, University of Macedonia, Thessaloniki, Greece. He holds a PhD in Quantitative Methods and he teaches statistics and research methods. His research interests include the study of Web2.0 applications especially in Tourism and Politics. Author's contact: kz@uom.gr

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 also fits 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).

Globally, the wine industry attracts much attention although it is not a growth industry and to millions of investors and hundreds of millions of consumers, the wine industry provides a far more fascinating product than its shares of global expenditure or GDP might suggest (Anderson et al., 2001). Nowadays, the wine industry has adopted the Internet for promotion 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, 15,535 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 behaviour 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 blog belongs to, and the size of blog's ego-network. Next, the paper constructs an overall index which serves to categorize wine-blogs according to their magnitude of influence.

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). Wine blogs are established by a blogger or group of bloggers who write about wine and are quickly becoming a new source of information for wine lovers and other wine bloggers (Thach, 2010).

Generally, blogs are personal online journals. Bloggers personalize and actualize content and information online and present their views to a broad audience. Blogs help self-expression and have the power to engage people in collaborative activity, knowledge sharing, reflection and debate (Blood, 2002; Efimova et al., 2005; Hiler, 2002; Punie & Cabrera, 2005) and to share their consumption experience with others (Matsumura 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 sharing 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 explore them. An organizational or corporate blog is a means of communication between an organization and its public and enables the industry to interact meaningfully with their potential customers (Jüch and Stobbe. 2005, Kosonen et al. 2007, Lu and Hsiao, 2007). Wineries, marketers, retailers and wine businesses establish their

wine blogs to promote business, to promote and sell wine and wine related products (Thach, 2009). Some of the most popular wine blogs are commercialized (Santos, 2011). 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).

This recent explosion of blogs and consumer generated content have tremendous impact on the decision making behavior of Internet users (Sigala, 2007). Bloggers provide more authentic information, gained through personal experience (Sharda & Ponnada, 2007), trust one another, and take into consideration others' opinions when making a purchase decision (Kozinets, 2002). Salganik et al. (2006) proved that people might form completely different evaluations of the same song due to the perception of others' behavior and Chen (2008) also found that consumers could be influenced by the choices of others when making decisions about purchasing books online. 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).

O'Neill et al. (2002) who studied wine tourism in Australia, found that visitors' recommendations boost wine sales when vacationing opinion leaders told their experiences. Later on, 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 (607) also in German (35), Dutch (4), Italian (41), French (39) Japanese (2), Chinese (5) Spanish (22), Portuguese (11) , Hungarian (2), Norwegian (2), Indonesian(1) and Slovak (1). In the list there are Winery Blogs (67) and Wine Podcasts (43) too.

Annunziata et al. (2008) classify wine-related 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 reviews, are topics highly discussed followed by wine news, wine making, wine events, wine tourism, geoponic, cheap wines, wine education, images, wine marketing, wine packaging, book reviews, culture and wine, wine commentaries, 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.

BLOGS AND SOCIAL NETWORK ANALYSIS

Social networks are defined as "a collectivity of individuals among whom exchanges take place that are supported only by shared norms of trustworthy behavior" (Liebeskind et al., 1996, p.429). Social networks are formed from the inherent tendency of people to come together in the structure of society (Backstrom et al., 2006). Nowadays, technology is increasingly incorporated into people's day-to-day social relationships (Farnham et al., 2004) and "social media" technologies have altered the underlying architectures of social interaction (Boyd, 2007). Thus, new online social networks have emerged linking people, organizations, and knowledge and new ties are developed among people sharing interests. Time and space become less important in the communication, but it is easier to communicate with large groups of community members and to bring unconnected community members into direct contact (Wellman, 2001). Blogs are intrinsically social (Nardi et al., 2004) and provide appropriate features for the bloggers to interact with their readers and other bloggers, to manage blog interactivity and hint at a blogger's connections (Du & Wagner, 2006; Sigala, 2008; Su et al., 2005). Therefore, blogs that share similar interests, views, or opinions are usually inter-connected, which form social networks among the bloggers (Warmbrodt et al., 2008).

Social Network Analysis (SNA) is a sociological methodology for analyzing patterns of relationships and interactions between social actors and seeks to explain social phenomena through a structural inter-

pretation of human interaction (Marlow, 2004; Wasserman & Faust, 1994). SNA focuses on the dialectical relationships between agents, nodes, and actors (Phillips & Ryan, 2007) and its goal is to identify the key actors (age, gender, socioeconomic status, and education) and the properties of their relationships (nature, intensity, and frequency of the relationships) (Chau & Xu, 2008; Krackhardt, 1996). SNA methods have been employed to study many areas like organizational behavior, inter-organizational relations, citation patterns, computer mediated communication (Chau & Xu, 2008). According to Wellman (2004), SNA is a powerful method for "explaining variances in resources, social behavior and socio-economic outcomes. When applied to knowledge management, it can identify patterns of interaction and knowledge--exchange flows within a network. It shows how knowledge-intensive work is done and can illustrate complex communication channels within a network (Phillips & Ryan, 2007).

SNA is rooted in the concepts of nodes and connections (Balancieri et al., 2007). Nodes are the social actors and can be persons, groups, organizations, nations, communities, neighborhoods, departments within organizations, offices, journal articles, web pages, blogs and so on and "connections" refer to channels of communication (Balancieri et al., 2007; Martino & Spoto, 2006; Quan-Haase and Wellman, 2006; Watts, 1999; White et al., 2004). When used to mine a network, SNA can help to reveal the structural patterns that may have important implications. For example, central nodes can be revealed which act as hubs, leaders, or gatekeepers, or bridging different communities (Albert et al., 2000).

A Social Network can be represented in three ways: the first one is by giving a simple list of all the elements taken from the set of social actors, and the list of the pairs of elements that are linked by a social relationship of some kind. The second has a form of a matrix. If two social actors I and J have a relation then 1 is placed at the cell (i,j), otherwise 0 is placed in this cell. Finally, a description of a Social Network may have a form of a graph where social actors can be represented by nodes, and the connections with each other can be represented by edges between these nodes (Marlow, 2004; Martino & Spoto, 2006).

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? Previous 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, 2004; 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 blog post is often eloquent. A long post often suggests some necessity of doing so. Therefore, they used the length of a post as a heuristic measure for checking if a post is influential or not.

Adar & Adamic (2005) and Ghosh & Lerman (2008) studied community identity in relation to influence. Both studies are based on the assumption that blogs within the same community tend to influence each other much as close affiliates would. Ghosh and Lerman (2008) defined influence as the number of paths, of any length, that exist between two nodes. They used the influence metric to partition a network into groups or communities by looking for regions of the network where nodes have more influence over each other than over nodes outside the community. Influence based ranking scores enabled the writers to identify the most influential nodes, but has also shown that as the weight of indirect links grows, the rank of the nodes that act as bridges between communities increase.

Karpf (2008) illustrated four distinct areas of influence: network centrality, link density, site traffic, and community activity. He used the Blogosphere Authority Index that combines data from four measures

of online influence into a single ranking system. The four measures used are: 1. The Network Centrality Score. It is an applied sociometric variable that is culled from blogrolls, which are self-reports of recommended or approved blogs. 2. The Hyperlink Authority Score that was derived directly from Technorati.com authority tracking system. 3. The Site Traffic Score. He used Sitemeter.com that directly measures the number of unique visitors a website receives every day. 4. The Community Activity Score. Nearly all of the blogs include a mechanism for readers to write comments in response to the author's post. Total Comments/Week was used to serve as a site's Community Activity Score.

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 1,305 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 exclusively on wine or wine and food, but not on other topics or posting infrequently on wine, 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 1,305 by 1,305 non--symmetric binary adjacency matrix. In this matrix if blog I links to blog J then 1 is placed at the cell (i,j), otherwise 0 is placed in this cell. 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 the bloggers' community.

When 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 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. A 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 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, a 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 1-cliques used in the analysis contain at least three blogs, that is they do 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 skewnness is discussed. For a certain SNA index, skewnness refers to the property that a minority of blogs present high values of the index, while the majority have only small values. For example, regarding the number of incoming links, skewnness implies that only a few blogs have many incoming links and the majority have only a few incoming links. A simple procedure to explore skewnness is to use a scaterplot of the particular index versus the ranks of this index.

Next, the paper combines the five proposed indexes to form one final overall index of blog connectivity. The method followed is the Principal Components Analysis. Scores of the Principal Components can be sorted in order to distinguish which are the blogs that have the largest overall network connectivity. Technically, factor scores which have values larger than unity represent blogs with high overall community connectivity scores.

FINDINGS

For the 1,305 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 regard to 1-cliques formation, but on the other hand few wine-blogs (presented above) are extremely con-

nected and involved within the community. This conclusion can be supported by studying skewnness. An easy way to present skewnness is to draw the scaterplot of numbers of 1-cliques that a blog belongs to, versus the ranks of these numbers. From Figure 1d the property of skewnness becomes obvious.

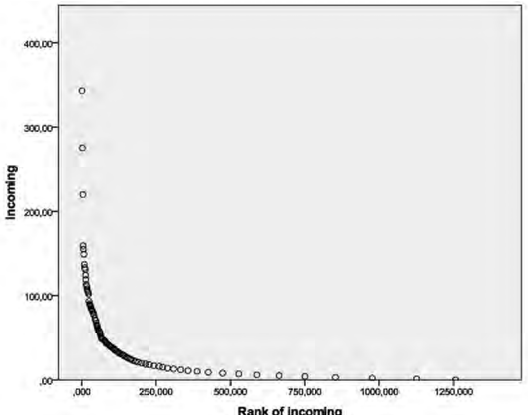
The Ego-network size ranges from zero to 449, with a mean equal to 22.6 and a standard deviation equal to 35.2. Five blogs have ego--network size greater that 200: fermentation.typepad.com, vinography.com, drvino.com, goodwineunder20.blogspot.com, goodgrape.com. Fifty percent of the blogs have ego-network size less than 12 and 75% of the blogs have size less or equal to 25. This means that 25% of the blogs have the highest ego-network sizes, which reach the maximum value 449. Skewnness is also obvious here when constructing the scaterplot of ego-network size versus ranks of ego-network size (Figure 1e).

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 there are some blogs that have no incoming or outgoing links, while some blogs have up to 334 outgoing and 343 incoming links. Figure 1a and 1b present the skewed scaterplots of incoming and outgoing links, respectively, versus their ranks.

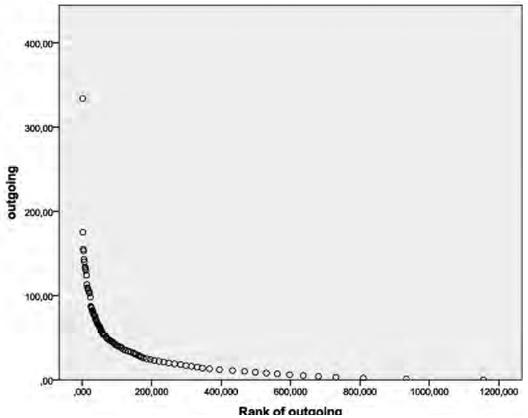
Normalized betweenness is a graph theoretic index associated with centrality of the network. The greater the normalized betweenness the more central a blog is. Table 1 shows that mean normalized betweenness equals 0.1. Because standard deviation equals 0.4, which is greater that the mean, one can conclude 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. Figure 1c presents the skewnness of the scaterplot of normalized betweenness versus its rank.

Table 1. Descriptive statistics of the five proposed SNA indexes

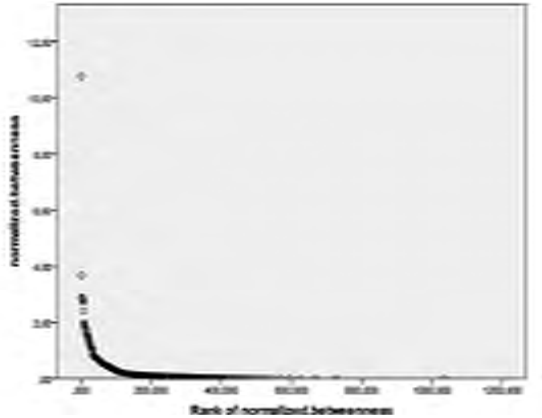
	N	Min	Max	Mean	S. D
Incoming links	1305	0	343	12.9	25.9
Outgoing links	1305	0	334	12.9	23.1
Normalized betweenness	1305	0	10.76	0.1	0.4
Number of 1-cliques that a blog belongs to	1305	0	11406	65.2	404.7
Size of Ego-network	1305	0	449	22.6	35.2



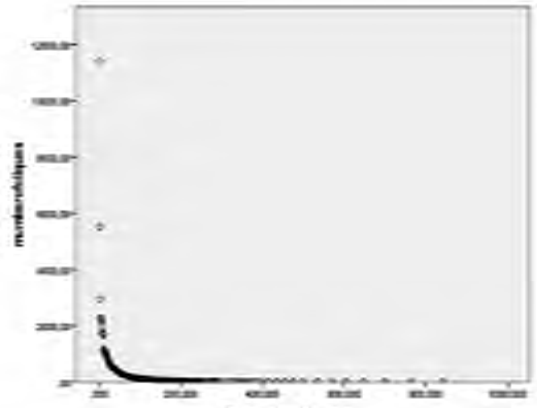
a) number of incoming links vs. its rank



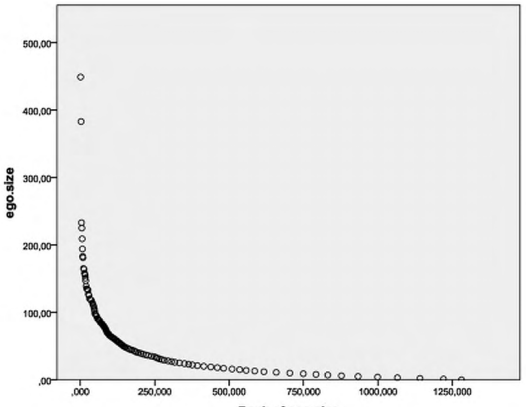
b) number of outgoing links vs. its rank



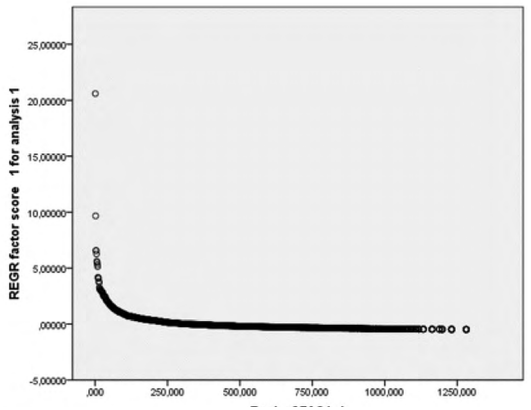
c) normalized betweenness vs. its rank



d) number of 1-cliques that a blog belongs to vs. its rank



e) size of Ego-network vs. its rank



f) factor scores vs. rank of factor scores

Figure 1. Scaterplots of the proposed indexes and the overall factor score versus their ranks

All of the above proposed indexes seem to behave similarly regarding the property of skewness (Figure 1). This is indeed true as it can be seen from Table 2, which presents the correlation coefficients of the 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 a strong evidence that all the indexes measure network centrality and blogs connectivity.

Table 2. Correlations matrix of the five SNA indexes

	Incoming links	Outgoing links	Normalized betweenness	Number of 1-cliques that a blog belongs to
Outgoing links	0.418*			
Normalized betweenness	0.615*	0.717*		
Number of 1-cliques that a blog belongs to	0.610*	0.586*	0.794*	
Size of Ego-network	0.888*	0.774*	0.731*	0.677*

(: $p<0.01$)*

Since all the proposed indexes are inter-correlated, it is a standard procedure to produce an overall index that will summarize them and will not allow for overrepresentation of some properties, through the indexes, and the underrepresentation of some other indexes. Principal Components Analysis (PCA) is a usual approach for data reduction. PCA produces one component which explains 75% of the total variance. Factor loadings are presented on Table 3. All factor loadings are positive and greater than 0.8. Through PCA, factor scores can be calculated. These factor scores can be used to summarize the indexes. By construction, factor scores have a mean equal to 0 and a standard deviation equal to 1. It is a standard procedure to separate cases (i.e. the wine-blogs) using their factor scores. Blogs with factor scores greater than 1 are usually considered to be the ones that have the highest values of the factor scores and consequently of the five proposed indexes, while those blogs with factor scores smaller than -1 are considered to have the lowest values of the factor scores and the five indexes, as well. Ninety blogs have factor scores over 1. That is, only a percentage of 7% of the wine blogs are really networked and active. They constitute a minority of the 1,305 recorded blogs.

Factor scores are also characterized by the property of skewnness. In Figure 1f, the scaterplot of factor scores versus their ranks reveal this property. Only few wine-blogs have the highest scores. The property of skewnness 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 4 presents the percentiles of these distributions. One can see from Table 4 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, their ego-network size is up to 21 and their overall factor score nearly equals zero (-0,08). Values grow rapidly at the largest percentiles, and this is due to the property of skewnness. The highest values are present after the 90% percentile for all the proposed indexes.

Table 3. Factor loadings of the five proposed indexes.

Indexes	PC
Size of Ego-network	0.944
Normalized betweenness	0.894
Number of 1-cliques a blog belongs to	0.849
Incoming links	0.820
Outgoing links	0.808

Table 4. Percentiles of the proposed indexes

Percentile	Incoming links	Outgoing links	Normalized betweenness	Number of 1-cliques a blog belongs to	Ego-network size	Overall score
10%	1	0	0	0	2	-0,45
20%	1	0	0	0	3	-0,42
30%	2	1	0	0	5	-0,39
40%	4	2	0	3	8	-0,33
50%	5	5	0	5	11	-0,28
60%	7	8	0,01	8	16	-0,19
70%	10	12	0,02	14	21	-0,08
80%	16	19	0,07	29	31	0,12
90%	31	34	0,17	92	55	0,65
100%	343	334	10,76	11406	449	20,59

The procedure introduced in this paper can lead to the overall ranking of wine blogs according to the overall factor score. As an example, Table 5 presents the Top-10 wine-blogs according to the sorted factor scores.

Table 5. Top-10 wine-blogs according to overall factor score, with scores on the five proposed indexes.

Wine-blogs	Number of 1-cliques a blog belongs to	Ego--network size	Outgoing links	Incoming links	Normalized betweenness	Factor score (PC)
fermentation.typepad.com	11406	449	334	275	10.76	20.6
vinography.com	5522	383	45	343	1.75	9.7
goodwineunder20.blogspot.com	2239	225	153	137	2.86	6.6
goodgrape.com	2947	209	98	159	2.4	6.3
wannabewino.com	1679	181	143	106	2.9	5.6
lennthompson.typepad.com	1709	194	103	131	2.8	5.5
drvino.com	1975	233	29	220	1.6	5.3
jimsloire.blogspot.com	486	183	175	41	3.7	5.1
wildwallawallawinewoman.blogspot.com	2276	135	113	84	1.2	4.2
dirtysouthwine.com	1179	158	124	87	2.8	4.1

By navigating to the top-10 wine-blogs, one can see that half of the top-10 wine-blog authors are professionals, as in the case of fermantationtypepad.com and vinography.com, jimsloire.blogspot.com, wildwallawallawinewoman.blogspot.com, goodgrape.com. These professional are managers of marketing and CRM firms or public and media relations firms serving the wine industry. Vinography.com and lennthompson.typepad.com have contributors, while blogs are authored by individuals and not by the wine industry. This is important, since it could lead to the conclusion that influence of wine-blogs is not depended by the industry-versus-personal blog issue. Rather, it is personal and the single-authored blogs that are most active and connected.

Most of the authors reside in USA. Most of the blogs were created from 2004 to 2008, while drvino.com was created in 2002. The authors of fermentation.typepad.com, goodgrape.com, and wildwallawallawinewoman.blogspot.com have certificates in wines and spirits

or viniculture. Seven out of ten blogs have been awarded some kind of prizes or awards, particularly goodwineunder20.blogspot.com, drvino.com, goodgrape.com and vinography.com and fermentation.typad.com. Several blogs are maintained by one author, but vinography.com and lennthompson.typepad.com have many contributors. Two blogs have sponsors. Six of the blogs have advertisements or links to commercial products or sites related to wine. Six of the blogs are promoting trips related to wine. Seven of the blogs are providing links to producers or other sellers. Four of the blogs offer education or links to wine education resources. Almost all of the blogs of the blogs provide wine reviews. Six blogs offer food or restaurant reviews or recommendations, or links to food blogs.

CONCLUSIONS

The paper proposed a method for measuring wine-blogs connectivity and influence. Social Network Analysis provide 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 to find 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.

One of the discussion topics in wine-blogs is wine tourism. Bloggers frequently discuss their experiences of wine tasting while visiting specific geographic regions. Some blogs are more oriented in presenting and promoting wines of specific regions and tourism of these regions. As mentioned in the influence of wine-blogs section of this paper, blog readers share their interest with bloggers and trust discussions and recommendations when they wish to be informed about wines and wine tourism. Combining this with the fact that tourism is the largest selling industry on the Internet, it is interesting, therefore,

to find out which are the wine-blogs that are really influential. As the paper showed, only a minority of wine-blogs are connected but despite being few in number, they have the ability to address a much wider audience than the majority of the wine-blogs.

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 and properties. Studying the travel blogs of Travelpod.com, the authors found out that only few travel blogs are well connected and can serve as focal points, where information provision might be larger that in the rest of the blogs, since the likelihood that these blogs are visited is greater. In conclusion, skewnness 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

Adar, E., & Adamic, L.A. (2005). Tracking information epidemics in blogspace. Proceedings of the *2005 IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 207-214) Washington DC: IEEE Computer Society.

Agarwal, N. & Huan, L. (2008). Blogosphere: Research issues, tools, and applications. *KDD Explorations,10(1)*, 19–29.

Agarwal, N., Huan, L., Tang, L., & Yu, P. (2008). Identifying the Influential Bloggers in a Community. *WSDM'08*, February 11–12, Palo Alto, California, USA .

Akehurst, G. (2009). User generated content: the use of blogs for tourism organizations and tourism consumers. *Service Business, 3 (1)*,51-61.

Albert, R., Jeong, H., & Barabási, A.-L. (2000). Error and Attack Tolerance of Complex Networks. *Nature, 406*, 378-382

Ali-Hasan, N., & Adamic, L. (2007). Expressing Social Relationships on the Blog through Links and Comments. *International conference Webblogs and Social media*. Colorado, Usa, March 27-28. Retrieved from: http://www.icwsm.org/papers/2--Ali-Hasan--Adamic.pdf

Anderson, K., Norman, D., & Wittwer, G. (2001). Globalization and the world's wine markets: overview. Centre for International Economic Studies and School of Economics University of Adelaide. Discussion Paper 014.

Annunziata, A., Misso, R., & Vecchio, R. (2008). Information and Communication Technologies in the Campania Region Wine Business. AWICT-SAE 2008 May 22-23, Alexandroupolis, Greece. Retrieved from: http://ageconsearch.umn.edu/ bitstream/ 37674/2/ICT_in_the_Campania_Region_wine_business.pdf

Backstrom, L., Huttenlocher, D., & Kleinberg, J. (2006). Group Formation in Large Social Networks: Membership, Growth, and Evolution. *KDD'06*, August 20–23, Philadelphia, Pennsylvania, USA.

Balancieri, R., Cuel, R., & dos Santos Pacheco, R.C. (2007). Social Network Analysis for Innovation and Coordination. Proceedings of *I-KNOW '07* Graz, Austria, September 5-7, 2007. Retrieved from, http://triple-i.tugraz.at/blog/wp-content/uploads /2008/ 11/9_social-network-analysis-for--innovation-and-coordiniation.pdf

Blood, R. (2002). *The Weblog Handbook: Practical Advice on Creating and Maintaining Your Blog*. Cambridge MA: Perseus Publishing

Boyd, D. (2007). Social Network Sites: Public, Private, or What? *Knowledge Tree* 13, May. Retrieved from http://kt.flexiblelearning.net.au/tkt2007/?page_id=28

Chau, M., & Xu, J. (2008). Using web minining and social network analysis to study the emergence of cyber communities in blogs. *Terrorism Informatics. Integrated Series in Information Systems, 18 (II)*, 473-494.

Chen, Y.-F. (2008). Herd behavior in purchasing books online. *Computers in Human Behavior*, *24(5)*, 1977-1992.

Cialdini, R.B. (2001). *Influence: Science and Practice*, 4th ed., Needham Heights, MA: Allyn & Bacon.

Drezner, D., & Farrell, H. (2004). The power and politics of blogs. Paper presented at the 2004 Annual Meeting of the American Political Science Association, Washington, DC, August. Retrieved from: University of Toronto Scarborough. http://www.utsc.utoronto.ca /~farrell/ blogpaperfinal.pdf

Drezner, D., & Farrell, H. (2008). The power and politics of blogs. *Public Choice, 134 (1-2)*, 15–30.

Du, H., & Wagner, C. (2006). Weblog success: Exploring the role of technology. *International Journal of Human-Computer Studies, 64 (9)*, 789-798.

Efimova, L., Hendrick, S., & Anjewierden, A. (2005). Finding "the life between buildings": an approach for defining a weblog community. AOIR Internet Research 6.0: Internet Generations, Chicago. Retrieved from: https://doc.novay.nl/dsweb/Get/Document-55092/AOIR_blog_communities.pdf.

Farnham, S., Uberoi Kelly, S., Portnoy, W., & Schwartz, J. W. (2004). Designing Social Software for Co-Located Social Networks, Proceedings of the *37th Annual Hawaii International Conference on System Sciences* (HICSS'04)

Ghosh, R., & Lerman, K. (2008). Community detection using a measure of global influence. Proceedings of the *2nd SNA-KDD Workshop on Social Network Mining and Analysis* (pp. 20-35) Las Vegas, NV: Springer, Berlin.

Hiler, J. (2002). *Blogs as disruptive tech: How weblogs are flying under the radar of the content management giants*. Retrieved from: http://www.webcrimson.com/ourstories/ blogsdisruptivetech.htm.

Jóch, C., & Stobbe, A. (2005). Blogs: The new magic formula for corporate communications? Economics. Digital economy and structural change. *Deutsche Bank Research, 53 (August),* 22.

Kaldis, K., Boccorh, R., & Buhalis. D.(2003). Technology Enabled Distribution of Hotels. An Investigation of the Hotel Sector in Athens, Greece. In *Information and Communication Technologies in Tourism in 2003*, (pp. 280—287) Wien: Springer-Verlag.

Karpf, D. (2008). *Measuring Influence in the Political Blogosphere: Who is Winning and How Can We Tell?,* George Washington University's Institute for Politics, Democracy and the Internet. Retrieved from: http://www.the4dgroup.com/BAI/articles /PoliTechArticle.pdf

Kiang, M., Raghu T.S., & Huei-Min Shan, K. (2000). Marketing on the Internet — who can benefit from an online marketing approach? *Decision Support Systems, 27 (4)*, 383–393.

Kolbitsch, J., & Mauer, H. (2007). The Growing Importance of e-Communities. *Web Lecture Notes in Business Information Processing, 1 (1)*,19-37.

Kozinets, R.V. (2002). The field behind the screen: using netnography for marketing research in online communities. *Journal of Marketing Research, 39*, 61–72.

Kosonen, M., Henttonen, K., & Ellonen, H-K. (2007). Weblogs and internal communication in a corporate environment: a case from the ICT industry. *International Journal of Knowledge and Learning*, *3 (4/5)*, 437-449.

Krackhardt, D. (1996). *Social Networks and the Liability of Newness for Managers.. In* Cooper, C.L. and Rousseau D. M. (eds.), *Trends in Organizational Behavior* (pp. 159-173). New York, NY: John Wiley and Sons, Ltd

Lai, L. S. L., & Turban, E. (2008). Groups Formation and Operations in the Web 2.0 Environment and Social Networks. *Group Decis Negot*, *17*, 387–402.

Leskovec, J., Adamic, L.A., & Huberman, B. (2007). The dynamics of viral marketing. *ACM Transactions on the Web. 1 (1)*, 1-46.

Lewis, R.C., & Chambers, R.E. (2000). *Marketing Leadership in Hospitality, Foundations and Practices*, 3rd ed. New York: Wiley.

Liebeskind, J.P., Oliver, A.L., Zucker, L., & Brewer, M.(1996). Social networks, learning and flexibility: Sourcing scientific knowledge in new biotechnology firms, *Organization Science*, *7 (4)*, 428-443

Litvin, S., Goldsmith, R., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism management*, *29 (3)*,458-468

Lu, H-P., & Hsiao, K-L. (2007). Understanding intention to continuously share information on weblogs. *Internet Research*, *17 (4)*, 345-361.

Marlow, C. (2004). Audience, structure and authority in the weblog community. In *The 54th Annual Conference of the International Communication Association*. Retrieved May, 1, 2009, from, http://www.researchmethods.org/ICA2004.pdf

Martino, F., & Spoto, A. (2006). Social Network Analysis: A brief theoretical review and further perspectives in the study of Information Technology. *PsychNology Journal*, *4 (1)*, 53 – 86.

Matsumura, N., Yamamoto, H., & Tomozawa, D. (2008). Finding influencers and consumer insights in the blogosphere, Proceedings of the *International Conference on Weblogs and Social Media*, (pp. 76-83). Menlo Park, CA: AAAI Press.

McGlohon, M., Leskovec, J., Faloutsos, C., Hurst, M., & Glance, N. (2007). *Finding Patterns in Blog Shapes and Blog Evolution*. School of Computer Science CMU-ML-07-100. Carnegie Mellon University.

McMillan, R. (2007). *State of the wine industry: forecast and strategic recommendations, 2007 – 2008* Silicon Valley Bank. Retrieved from: http://www.svb.com/Publications/Sector-Trends/Wine-Industry/State-of-the-Wine--Industry,-Forecast-and-Strategic-Recommendations-2007-2008/

Nardi, B.A., Schiano, D.J., Gumbrecht, M., & Swartz, L. (2004). Why we blog. *Communications of the ACM*, *47 (12)*, 41 - 46

Nóvoa, H. (2010). The Portuguese wine industry and Web 2.0: The Infovini project. *5th International Academy of Wine Business Research Conference*, 8-10 February. Auckland (NZ).

O'Neill, M., Palmer, A., & Charters, S. (2002). Wine production as a service experience – the effects of service quality on wine sales. *The Journal of Services Marketing, 16 (4)*, 342- 362.

Phillips, P. & Ryan, C. (2007). Building Research Clusters: Exploring Public Policy Options for Supporting Regional Innovation. *In* A Krattiger, RT Mahoney, L Nelsen, et al. (eds.) *Intellectual Property Management in Health and Agricultural Innovation: A Handbook of Best Practices*. (pp. 317-333). Oxford, U.K:MIHR and Davis, U.S.A.: PIPRA.

Punie, Y., & Cabrera M. (2005). The Future of ICT and Learning in the Knowledge Society. Report on a *Joint DG JRC-DG EAC Workshop* Seville, 20-21 October.

Quan-Haase, A., & Wellman, B. (2006). Hyperconnected net work. In Charles Heckscher and Paul Adler (eds.) *The Firm as a Collaborative Community* (pp. 281-333). New York: Oxford University Press.

Rabanser, U., & Ricci, F. (2005). Recommender Systems: Do They Have a Viable Business Model in E-Tourism? In *Information and Communication Technologies in Tourism in 2005* (pp. 160-171). Wien: Springer-Verlag.

Salganik, M.J., Dodds, P.S. & Watts, D.J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science, 311*, 854-856.

Santos, F. (2011). *Motivations and Characteristics of International Wine Bloggers*. American Association of wine economists. AAWE Working paper, No 92 Business.

Sharda, N., & Ponnada, M. (2007). Tourism Blog Visualizer for better tour planning. *Journal of Vacation Marketing, 14 (2)*, 157-167.

Sigala, M. (2007). WEB 2.0 in the tourism industry: A new tourism generation and new e-business models, *Ecoclub, 90*, 5-8.

Sigala M. (2008). Developing and implementing an eCRM 2.0 strategy: usage and readiness of Greek tourism firms. In O'Connor,P. Hφpken, W. and Gretzel, U. (Eds) *Information and Communication Technologies in Tourism 2008*, (pp. 463-474). Wien:Springer-Verlag.

Sigala, M. (2009). WEB 2.0, social marketing strategies and distribution channels for city destinations: enhancing the participatory role of travelers and exploiting their collective intelligence. In Gascσ-Hernαndez, M and Torres--Coronas, T. (Eds.) *Information Communication Technologies and City Marketing: Digital Opportunities for Cities around the World*, (pp. 220 – 244) IDEA Publishing

Song, X., Chi, Y., Hino, K., & Tseng, B. (2007). Identifying Opinion Leaders in the Blogosphere. *CIKM'07*, November 6-8, 2007, Lisboa, Portugal.

Stricker, S., Sumner, D., & Mueller, R. (2003). Wine on the web in a global market: A comparison of E-commerce readiness and use in Australia, California and Germany, *IV EFITA (European Federation for Information Technology in Agriculture) Conference*. July, Hungary.

Su, M. N., Wang, Y., & Mark, G. (2005). Politics as Usual in the Blogosphere. Proceedings of the *4th International Workshop on Social Intelligence Design* (SID 2005).

Tan, L.K.W., Na, J.C. & Theng, Y.L. (2011). Influence detection between blog posts through blog features, content analysis, and community identity. *Online Information Review*, *35 (3)*, 425-442.

Technorati.com (2008). *The State of the Blogosphere 2008*. Retrieved from: http://technorati.com/blogging/feature/state-of-the-blogosphere-2008/

Thach, L. (2009): Wine 2.0 - The Next Phase of Wine Marketing? Exploring US Winery Adoption of Wine 2.0 Components. *Journal of Wine Research*, *20 (2)*, 143-157

Thach, L. (2010). Wine blogs: Expressing diverse wine opinions in a new realm of online wine relationship marketing. *5th International Academy of Wine Business Research Conference*, 8-10 February 2010 Auckland (NZ).

Trammell, K. & Keshelashvili, A. (2005). Examining the new influencers: A self-presentation study of A-list Blogs. *Journalism and Mass-Communication Quarterly, 82 (4)*, 968-982.

Vrana, V., & Zafiropoulos, K. (2010). Locating Central Travelers' Groups in Travel Blogs' Social Networks, *Journal of Enterprise Information Management*, *23(5)*, 595 – 609.

Vrana, V., Zafiropoulos, K. & Vagianos V. (2011). An Exploration of Wine Blog Communication Patterns. In M. Sigala, U. Gretzel, and E. Christou (Eds.), *Web 2.0 in travel, tourism and hospitality: theory, practice and cases*. Farnham, UK: Ashgate Publishers.

Wagner, C., & Bolloju, N. (2005). Supporting knowledge management in organizations with conversational technologies: discussion forums, weblogs, and wikis. *Journal of Database Management, 16 (2),* i–viii

Wang, H.W. (2010). Exploring the factors of gastronomy blogs influencing readers' intention to taste. *International Journal of Hospitality Management, 30*, 503–514.

Warmbrodt, J., Sheng, H., & Hall, R. (2008). Social Network Analysis of Video Bloggers? Community. Proceedings of the *41st Hawaii International Conference on System Sciences.*

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Watts, D. (1999). *Small Worlds.* Princeton, NJ: Princeton University Press.

Wellman, B. (2001). Computer networks as social networks. *Science Magazine*, 293, 2031-2034.

Wellman, B. (2004). *Measuring Social Capital through Analysis of Social Networks*. University of Toronto. Session I: Expert Workshop on the Measurement of Social Capital Policy. Research Initiative/Social Sciences and Humanities Research Council/Statistics Canada (PRI/SSHRC/StatsCan)..

White, H., Wellman, B., & Nazer, N. (2004). Does citation reflect social structure? *Journal of the American Society for Information Science and Technology*, *55(2)*, 111- 126.

Woodly, D. (2008). New competencies in democratic communication? Blogs, agenda setting and political participation. *Public Choice, 134*, 109–123.

Yarrow, A. (2010). *The Complete List of Wine Blog's*. Retrieved from: http://www.vinography.com

Zafiropoulos, K., Vrana, V. & Vagianos, D. (2010). Conversation authorities among popular wine blogs. *EuroCHRIE, Passion for Hospitality Excellence*, Amsterdam 25th – 28th October.

Zafiropoulos, K., & Vrana, V. (2011). *Hyperlink analysis of political blogs communication patterns*. USA: Nova Science Publishers, Inc.

Submitted: 15th September 2011 *Accepted: 20th January 2012*
Final version: 12th January 2012 *Refereed anonymously*