

Weblogs as indicators of urban social relationship

Abstract

Drawing on a sample of 2,598 weblogs that are identified by location in the United States, this project extracts the outward links of these weblogs from the NITLE census dataset. A total of 44 U.S. city/region units represented by the first three digits US zip codes are examined for how far, geographically, their bloggers link. An index of Spatial Social Capital & Openness (SSC&O) is created to measure the urbanism reflected in the physical distance and scale of social relationships. The results demonstrate that bloggers on the West Coast have higher SSC&O values, both domestically and internationally, than bloggers in other parts of the US. New suburban technology centers, link further than traditional major cities. Furthermore, the findings suggest a shift of urban characteristics from the Northeast to the West Coast, and from traditional metropolitan centers to the then suburbs or fringes of major cities.

Networked urbanity

Urban culture is a set of specific social mentalities decided by a type of open and loose interpersonal relationship, which is contrasted to primary ties found in countryside. Castells points out that urban culture is a product of “specific system of social relations by a given ecological context (the city)” (2002, p.35). In his discussion (2000) of the network society Castells suggests that the society is constructed around flows of such matters as capital, information, technology, organizational interaction, images, sounds and symbols. It is not the place per se, but the space of flows among nodes and hubs that illustrates power and strength of the place. A space does not exist by itself; instead, it is defined by the exchange of flows in the network. Nonetheless, it does not mean that a space of flow is placeless, since “this network links up specific places, with well-defined social, cultural, physical, and functional characteristics” (p. 443). While most people’s life and experience is rooted in places, the power and wealth is projected in a much broader space. Castells believes that “elites are cosmopolitan and people are local” (p.446). In the other word, elites tend to define their community in a spatial bound rather than a geographical one. Their personal micro-network “project them in functional macro-networks throughout the global set of interactions in the space of flows” (Castells, p.446).

Lefebvre suggests that the only way to study a society’s classes, ideas, values and political systems is to recognize the space on all planes of a world scale (Philips, 1996). A place is tied with “local” in a physical sense, while a space is a conceptual and abstract form as a result of training and education, politics and so on (1996b). A city is not only a place, but also a spatial congregation of the similar mindsets. Such an attribute endows urban cultures both physical and spacial connotations. To further clarify this issue, we will start with the discussion of the concepts of place and space. Splenger in his article “The soul of the city” elaborates the difference of the city people and country people by saying: “Today a Brandenburg peasant is closer to a Sicilian peasant than he is to a Berliner” (1922, p.66). Here is a very interesting notion of space and place. Brandenburg is no doubt closer to Berlin than to Sicilian, in the sense of geographical proximity, or say, place; however, the shared culture between

Brandenburg peasants and Sicilian peasants draws them much closer to each other. If they had been able to interact with each other at the beginning of the 20th century, the communication would have had formed a shared space. In general, the shared space among the city people from different places is made more possible by their tendency to travel around and their easy access to the public transportation. In the past century, shared spaces beyond geographical territory have become more and more common with the development of transportation and communication technology.

The term “cosmopolitan” has been used to describe a modern urbanization characterized by openness, sophistication and excitement; however, it originally referred to the social-psychological trait pertinent to the *world* opposite to the *local*. Loosing sense of place comes with the “civilization in its giant cities” through which the man becomes “intellectual nomad” (Splenger, 1922, p. 65). In *Gemeinschaft* culture, a community is defined by its territory, and place is always overlapping with space, while in *gesellschaft* culture, a community is based on the common culture, and place is not equal to space. Simmel also emphasizes that urban mentalities are transcending the city’s physical limit and are the “range of the person constituted by the sum of effects emanating from him temporally and spatially” (1950, p. 56).

The new digital technology serves to transform this “Intellectual nomad” to a “spatial intellectual nomad”, since it enables intellectual exchange, social activity and business transaction on the Internet, while people physically stay in the fixed place. The new economy based on information technology is functioned more in a ‘space’ than a ‘place’, since the constraints imposed by physical distance are gradually removed and the interactions in the digital space become a day-to-day experience. As a result, the accumulation and wisdom and talents, as well as the intellectual encounter can be realized with people physically dispersed in separated places. On the other hand, personal relationship is made more complex in the sense of distance and geography. Although the face-to-face interaction is still preferred, there is a strong tendency of spatial relationship in nature (Halavais, 2000). One of the most remarkable impacts of computer-assisted communication on people is the deprivation of the sense of place. A computer conferencing system called WELL was born in 1986, and it facilitated its members to join more than a hundred ongoing computer conferences and news groups. Some of its users describe their experience as being “in something” or “I’m somewhere”, while others comment that it is “a place that’s no place” (Philips, 1996).

Emergence of the Internet gave rise to many of such shared space to people in different places. Curry’s (1996a) examination on Lefebvre’s literature on language, ideas, and space postulates that reader, author and text construct a set of spaces which can be seen as an author-reader community. When author can write and reader can read in a totally neutral manner in an ideal setting, they are transcending of time and space. However, since there is no neutral reader and no interest-free writer, writing and reading are carried in a position of time and space. The text is “written for a special time and place and for the reader able to fit within that time and place” (p.185).

A broad definition of a weblog is any “hierarchy of text, images, media objects and data, arranged chronologically, that can be viewed in an HTML browser” (Winer, 2003). A personal weblog is an online journal regularly updated in a reversed chronicle order and it can be categorized into three types: first, a personal diary

recording daily activity and thoughts; second, a collection of links to other websites worthy of recommendations, usually with a few words of comments; third, a forum devoted to specific topics. The majority of blogs are in the first category, or mix of first and the second types, and they are highly personalized. “Kittyblogs” or “bloggeral” (blog doggerel) are terms that refer to the extreme examples of over-sharing triviality (Mcfedries, 2003). Blogs mainly consisted of links to other websites serve references or reminders. Through writing and linking to other sites of interests, a blogger reconstructs and presents a self image, consciously or unconsciously, in cyberspace.

A single personal website, such as a blog, can easily be submerged in a mighty torrent of millions of other similar sites, and the seeming triviality of its content might well lead to it being neglected. However, the totality of this micro-content across sites vividly and objectively depicts the social landscape and ideology at certain points of time and space. The rich, unsolicited source material and easy accessibility make weblogs a valuable alternative to both macro and micro social science research. The increasingly penetration of computer technology, such as data mining or knowledge discovery, into social science research empowers the more sophisticated analysis on much larger scale of data. Bloggers, the individuals who write blogs, provide semantic data for the research in public agenda or sentiment. Furthermore, the data constructs a complete dimension for social structure analysis. Blogging technology breaks down the traditional centralized authorship and makes everyone (with internet access) a potential author that can publish his writings to public. The occurrence of decentralized publishing sources consequently re-defines the community space. One of the important missions in studying blogosphere is to see the difference between this redefined space and the real physical place. Curry suggests that the act of being in cyberspace involves acting in places such as home, office and neighborhood (1996b), therefore, cyberspace and physical place could be determinant to each other.

Important features of blogging technologies include pinging, subscribing, commenting, tracking, hyperlinking, through which they constitute a set of affiliation networks, where people with the same interests or close social ties interact with and refer to each other. For the purpose of regular reading, bloggers place links of other blogs in their index pages (permanent links) or subscribe these blogs through RSS feeds. According to Rebecca Blood (2000), blogs started as link-driven sites and each was a mixture of links to news, commentary and thoughts. In this sense, one cluster of hyperlinks could be viewed a virtual neighborhood or community. They can be grouped by political preference, culture or academic interest, health or spiritual support, or for family and friendship ties.

The hyperlinks at the individual level are a good indicator of social relationships. A cosmopolitan social relation stretches far beyond the merely local connection and extends to a world-wide network. Putnam uses the term “social capital” to describe the social connections or social networks that enhance social cohesion and improve the civic individual’s life (2000). Putnam believes that people with active connections with others, such as friends, family or neighbors, develop healthy characters that are good for the rest of society, since they are more tolerant and less cynical. The network of social capital can serve as the conduit of information flows and sharing, which enables smooth everyday business and social transactions. The link between the concept of social capital and urban study was implied by Jane

Jacobs' discussion of urban life and neighborhood (1961). Simmel (1921), together with many other urban thinkers indicate that one of the most significant characteristic of the metropolis is the extension of people's social relations beyond its physical boundaries. This is also reflected in Castell's notion that "elites are cosmopolitan and people are local." (2001)

Methodology

This study postulates that the range and density of social relation networks reflect the strength of social relationships and degree of openness of a city ecosystem. Permanent links at side bars of personal weblogs represent the social affiliation networks based on common interests or social ties. Since these links often represent a spatial social relationship, we invoke a *spatial social capital & openness index* (SSC&O) to measure social interaction patterns in urban cultures. O_i , the SSC&O of a city C_i is computed in two ways based on domestic links and international links. Consider a set of N cities in America represented by $C = \{C_1, C_2, \dots, C_N\}$, let L_{ij} be the number of Outbound links from a city C_i to C_j , $i \in [1, N]$ and $j \in [1, N]$, and d_{ij} be the Euclidian distance between C_i to C_j , then

$$O_i = \sum_{j=1}^N d_{ij} \cdot L_{ij}$$

And for links to foreign countries, $L_{i(N+1)}$ represents the total number of links from the city C_i to all foreign countries:

$$O_i = L_{i(N+1)}$$

The values of two O_i s represent the city's overall openness and range of social relationships.

Blog link data

About 10,000 U.S weblogs are considered from NITLE census data collected during June and July in 2003. The blog data from NITLE at 2003 consists of two parts: a url list with about 500,000 blog sites and a database with index pages from about 120,000 weblogs. Unfortunately, the latter dataset does not indicate the URLs of crawled pages. Since the two datasets were crawled at about the same time, the first dataset includes the majority of blog sites in the second dataset. In order to use the hyperlink data extracted from archived index pages, the first step is to match the index page with its URLs provided in the first dataset.

Since *html* codes on index page usually have the title area between the tags '<title>' and '</title>', and titles, especially the keywords in titles, are generally consistent over time, the titles retrieved from the database and the titles of the url list that are retrieved later from the Internet are compared to find the best match. Both methods of full match and partial match of title words are applied to extract the URL corresponding to the index page.

URLs of permanent links—commonly referred to as “blogrolls”—were extracted from the source code of index pages. In most cases, blogrolls appear on index pages, but for the majority of blogs hosted by livejournal and diaryland, links instead appear on member information pages. To exclude hyperlinks in text entries, which represent an only temporary connection, only URLs on sidebars of a blog, and grouped together and parsed only by line breaks, are considered. The links to such popular web services as blog hosts, counter statistics, commenting systems, and other forms of support are excluded from the collection.

Only those weblogs with the matching title and keyword information in both datasets, as well as hyperlinks information on index pages, are included in the analysis. Clearly, more recent archives would be desirable, but were not available. Moreover, the rise of web feeds and aggregators has diminished the use of blogrolls, complicating such an analysis. For this approach to be applied to the blogosphere as it is today, a significant period of linking relationships would need to be captured..

The geographical locations of blogs are obtained from various sources including user profiles, geocodes, local weather on index pages and domain registrations. All blogs retrieved from this dataset are plotted to their corresponding 3-digit zip code. The categorization of blogs into 3-digit zip codes represents a middle approach that standardizes different levels of geographical labels (city names, latitude-longitude, street addresses) while keeping some geographical detail (Lin & Halavais, 2004; Lin & Halavais, 2005). The distribution of blogs based on 3-digit zip codes has been found to be consistent with population size and socio-economic profiles.

Findings

The limitation of the dataset employed in this study yields only 44 areas represented by 3-digit zip codes with substantial blog numbers (more than 25 blogs at each city unit). Based on the procedures above, we render the SS&O Index at both domestic and international levels.

For each zip code, SSC& O values appear in table 1. Among the 10 zip codes that have the highest SS&O values, all but 1 are located on the West Coast, especially in California. On the contrary, the areas with low SS& O are located in the eastern and central US, with only 1 in California. Contrary to common understanding, big cities in the Northeast do not show a strong tendency of outreach.

For SSC & O at international level, we also normalized number of outlinks to foreign countries by the total links. Once again cities from West Coast appear most often on the top of the index. The two indexes have significant correlation (.787).

Discussion

At this stage, we have only considered a relatively small number of US cities with significant long-standing populations of bloggers. Nonetheless, the result clearly demonstrates the patterns of linking behaviors in different American cities. Bloggers from the West Coast tend to link further than those from other part of the nation. An early study of the blog network in America finds that there are heavy linkage between

the West Coast and the Northeast (graph 1). This study further reveals that the direction of these links is in general pointed from the West to the East. We also compare this result with another study that used the same set of data to detect the topics of blogs in different area. We found a significant correlation between the high SSC& O values and the discussions of art, lifestyle and politics (table 3). These topics are typically interests of cultural elites, who tend to be urbanites residing in big cities. Interestingly, discussion of technology is not significant correlated with the SSC&O. Since the early adopters of blogging back in 2003 are mostly technology professionals, the geographical difference is not very significant in this category.

The cities on top of our SSC&O index include bay area (East Bay, Oakland, Santa Cruz), West Los Angeles (Los Angeles and Santa Monica), as well as Seattle and its vicinity, all of which are along the west coast, and they are not necessary the major city centers. The areas on bottom of the index are mostly located deep in land (Las Vegas, Detroit area, Chicago area, Dallas). Traditional centers, such as New York, Boston, Chicago or DC are not linking domestically or internationally as far as the cities in the West Coast. New York City and Boston are undeniably among those most famous “global cities”, with the most diversified population and active cultural scenes, but clearly remain the target of international attention, rather than actively turning their gaze outward. Another reason for these cities to be found low on the list may be that these cities have large populations as well as large numbers of bloggers. When we normalize the SSC&O by the blog numbers, it might take the edge off of the blogs with high SSC&O values.

However, our observation is consistent with the recent discussion on new geography shaped by new digital economy. The transformation of the manufacturing economy to an information economy since the 1980s has greatly changed the city landscape in America. According to Kotkin (2001), the majority of middle class families lived and worked in what he terms *Midopolis*, or the old suburbs. *Nerdistans* have grown up around the fringes of *Midopolis* and provide an escape from crowding and urbanism of *Midopolis*; since the early 1990s, there have been skilled professionals migrating from urban areas to the hinterlands. Central cities have either declined into recession or undergone a renaissance to transform into new centers of culture and entertainment.

The decline of the traditional industrial city is concomitant with the rise of the creative city, characterized by concentration of high-tech industries, knowledge-based production, comfortable lifestyles, and a creative class, which consists of the upper class of today’s society. These areas are not necessarily surrounded by governmental or corporation centers, as traditional industrial cities were, and they are much more fragmented and oriented to customization and creativity. Information-based business is able to be free from reliance on natural resources and cheap labor, which are geographically defined; therefore, businesses can start anywhere desirable (Kotkin, 2000). Florida (2000) contributes the rise of the creative class in the U.S. to the emergence of a “creative economy,” consisting of a workforce of creativity and empowered by information and knowledge: R&D personnel, writers, actors and artists lawyers, professors, CEOs, etc. Florida replaces the elite class with the term “creative class”, and keeps working class, such as those in manufacturing and low-skill services, out of his category. Florida found through a series of surveys that the creative class consists of people who are motivated by challenge and responsibility as well as a flexible working environment. The creative class does not particularly favor strong

community ties, but individualism and weak ties instead. Brook depicts today's social elite as a hybrid of "Bourgeois" and "Bohemian", which he designates with the term "Bobo" (2001). Unlike the traditional elite class born wealthy and prestigiously in the East Coast or Saint Louis, Brook suggests, people today can achieve their high social status and success through education accessible to the public, talents, and most importantly, creativity. The majority of social scholars tend to agree on the fact that the creative class is selective in deciding their work and residence location. Joel Kotkin (2000) proposes that today wealth is accumulated where intelligence clusters, and he reiterates the mounting importance of geography in the digital era by declaring "if people, companies, or industries can truly live anywhere... where to locate becomes increasingly contingent on the peculiar attributes of any given location"(p.6). Florida's (2002) examination of census data finds that the creative class has been moving away from traditional corporation and industry centers and pursuing creative centers which "provide eco-systems of habitat where all forms of artistic and cultural, technological and economic can take root and flourish" (p.218). Fussal notes that "X" people, the young creative people who are seeking to break through the social class system, prefer places with access to multicultural scenes, and tend to move out when the place is "drifting too speedily middle- or prolewards" (1992, p.181). Cities of high creativity are usually associated with a low proportion of the working class and service class. Contrary to general perception, Florida suggests that businesses follow creative people, instead of the other way around. Regarding geographical preference for the creative class, Florida observes that creative capital is drawn to cities with 3 "T"s: Technology, Talent and Tolerance. A diversified cultural environment is more important than traditional infrastructure advantages such as urban malls, freeways, and skyscrapers, since the creative class values authenticity and quality of life. As for social interaction, the creative class prefers weak ties instead of community defined by geographical proximity, similar to what was identified in previous discussion of urban culture.

The cities on top of our SS&O index shared the similar characteristics with what Kotkin (2000) identified as new city centers. They have small working class populations and they are among the highest concentrations of technology-oriented companies in the country. Cities like Los Angeles, San Francisco, and Houston started to thrive or revive from 1990s, followed by the booming of luxury services, restaurant, the arts and other entertainments, which attracted the technological elite of the new rich. The rise of these new city centers does not diminish the importance of the old great American cities such as New York, Boston and Chicago. Because of the population saturation in these cities, new young urbanites tend to move elsewhere for affordable lifestyles. Such new technology elites happen to be the main early blogging adopters. Whether or not the western urbanites will be more urbanized than their predecessors in the Northeast or Chicago is hard to tell based on this research. In terms of far-reaching social relationship and openness to the outside world, the bloggers on the West coast demonstrate a clear advantage.

In all, the trend observed in this study is consistent with the new geography shaped by the rising of high-tech centers on the West Coast. Contrasted with traditional cultural elites from the East Coast and concentrated more on such careers as law, finance and fine art, the new urbanites from such technology centers as the San Francisco Bay area, Seattle and Bellevue, and Southern California have heavy technology backgrounds and skills. Without further inquiry of blogger's demographic

information, we can only speculate the reasons of why they tend to link further. One reason is that a lot of young technology professionals might migrate from other part of the country or from other part of the world, so they retain personal ties with their former homes. The other explanation is that there are more opinion leaders on the East Coast which attract more inbound links from the West Coast. Nonetheless, a person traveling around in his lifetime, with social ties everywhere, and interested in what is said in other places, represents the new cosmopolitan consciousness.

Reference:

- Blood, R. (2000). Weblogs: a history and perspective. Available:
http://www.rebeccablood.net/essays/weblog_history.html
- Castells, E. (2001). Rise of the network society. Blackwell Publishers
- Curry, M.(1996a) The work in the world: geographical practice and the written world.
MN: University of Minnesota Press.
- Curry, M. (1996b) Cyberspace and cyberplaces: Rethinking the identity of individual
and place. The 20th IAMCR conference, Sydney
- Florida, R. (2002). The rise of the creative class. NY, Basic books
- Halavais, A. (2000) National borders on the world wide web. New media & society,
1,3, 7-28
- Jacobs, J. (1961) The death and life of great American cities. Random House
- Kotkin, J. (2001). The new geography: how the digital revolution is reshaping the
American landscape. NY: Random House.
- Lin, J. & Halavais, A.(2005) Geographical distribution of weblogs in America,
presented at International Communication Association 2005 General Conference,
New York, New York, May .
- Lin, J. & Halavais, A.(2005) Mapping Bloggosphere in America, presented at
Thirteenth International World Wide Web Conference, New York, New York,
May.
- MacFriedes, P. (2003). Blah, blah, Blog. IEEE Spectrum.
- Philips, B. (1996). City lights: Urban-suburban life in the global society. NY, Oxford:
Oxford University Press.
- Putnam, R.(2000). Bowling Alone: The Collapse and Revival of American
Community. New York: [Simon & Schuster](#)
- Simmel, G. (1921) The metropolis and mental life. In Richard Sennett (Eds., 1950)
Classic essays on the culture of cities. NY: Meredith Corporation
- Splengler, O. (1928) The soul of the city (from The decline of the West, Vol. II). In
Sennett, R. (Eds.) Classic essays on the culture of cities.
- Winer, D. (2003) What makes a weblog a weblog? Available at
<http://blogs.law.harvard.edu/whatMakesAWeblogAWeblog>

Appendix:

Table 1: SSC&O at domestic level (weighted outlinks to other parts of the country)

3d- zipcode	Area	SSC&O (Domestic)	blog no.
946	Oakland	29097.36	34
902	Santa Monica, CA	26300.83	52
201	North DC	24974.34	82
945	East SF Bay, CA	22334.13	87
940	West Bay, CA	20407.59	50
900	Los Angeles	19781.4	95
980	Bellevue, WA	19002.38	48
951	San Jose	16319.07	38
981	Seattle	16089.58	127
941	San Francisco	15884.26	125
950	Santa Cruz	14210.58	66
112	Brooklyn	13989.27	63
554	Minneapolis	13780.86	41
972	Portland	13141.57	63
303	Atlanta	11626.71	55
802	Denver	11359.33	25
947	Berkeley	10936.35	39
100	New York City	10253.16	203
787	Austin	10240.5	70
770	Houston, TX	10070.82	57
21	Boston	9458.628	140
483	Northwestern suburban Detroit	9139.858	26
921	San Diego	9088.032	43
750	Suburban of Dallas	8930.301	59
606	Chicago	8897.09	120
852	East Phoenix	8573.606	49
70	Suburb of Jersey city	8042.959	48
31	Manchester, NH	7482.135	32
24	Suburb Boston	7366.7	56
926	Orange County	7305.905	29
200	DC	7000.188	61
701	New Orleans	6767.555	65
300	Suburb Atlanta	6522.57	28
152	Pittsburgh	6150.145	32
191	Philadelphia	5867.997	37
532	Milwaukee	5602.022	28
600	North Chicago	5364.972	26
752	Dallas	5353.37	39
891	Las Vegas	4319.206	25
913	San Fernando	4186.78	41
480	Northern suburban Detroit	4075.739	28
	Ann Arbor & Southern vicinity of		
481	Detroit	3648.951	36
146	Rochester	2655.759	26
190	Suburban Philadelphia Northwest	1198.538	25

Table 2: SSC&O at international level (percentage of outlinks to other parts of the world)

3d- zipcode	Area	SSC & O (international)	blog no.
945	East SF Bay, CA	0.195719	87
31	Manchester, NH	0.190751	32
770	Houston, TX	0.177177	57
902	Santa Monica, CA	0.176147	52
980	Bellevue, WA	0.175953	48
946	Oakland, CA	0.174263	34
950	Santa Cruz, CA	0.16622	66
981	Seattle, WA	0.165432	127
201	North DC	0.164823	82
900	Los Angeles	0.164794	95
	Northwest of North Suburban		
483	Detroit	0.157609	26
750	Suburban of Dallas	0.151429	59
70	Suburb of Jersey city	0.148997	48
554	Minneapolis	0.144578	41
951	San Jose	0.142361	38
300	Suburb Atlanta	0.1375	28
701	New Orleans, LA	0.132075	65
606	Chicago	0.127329	120
100	New York City	0.125727	203
303	Atlanta	0.1232	55
940	West Bay	0.122711	50
191	Philadelphia	0.119342	37
24	Suburb of Boston	0.118012	56
21	Boston	0.116784	140
532	Milwaukee	0.116129	28
787	Austin	0.114701	70
200	DC	0.11336	61
921	San Diego	0.112	43
112	Brooklyn	0.111423	63
891	Las Vegas	0.111111	25
852	East Pheonix	0.110368	49
941	San Francisco	0.109336	125
972	Portland	0.103053	63
926	Orange County	0.102041	29
802	Denver	0.100402	25
752	Dallas	0.096774	39
947	Berkeley	0.096552	39
480	North suburban Detroit	0.095238	28
600	North Chicago	0.086022	26
152	Pittsburgh	0.076923	32
146	Rochester	0.068182	26
	Ann Arbor & South vicinity of		
481	Detroit	0.058608	36
913	San Fernando	0.057971	41
190	Suburban Philadelphia North west	0.045455	25

Graph: network of blog links in America

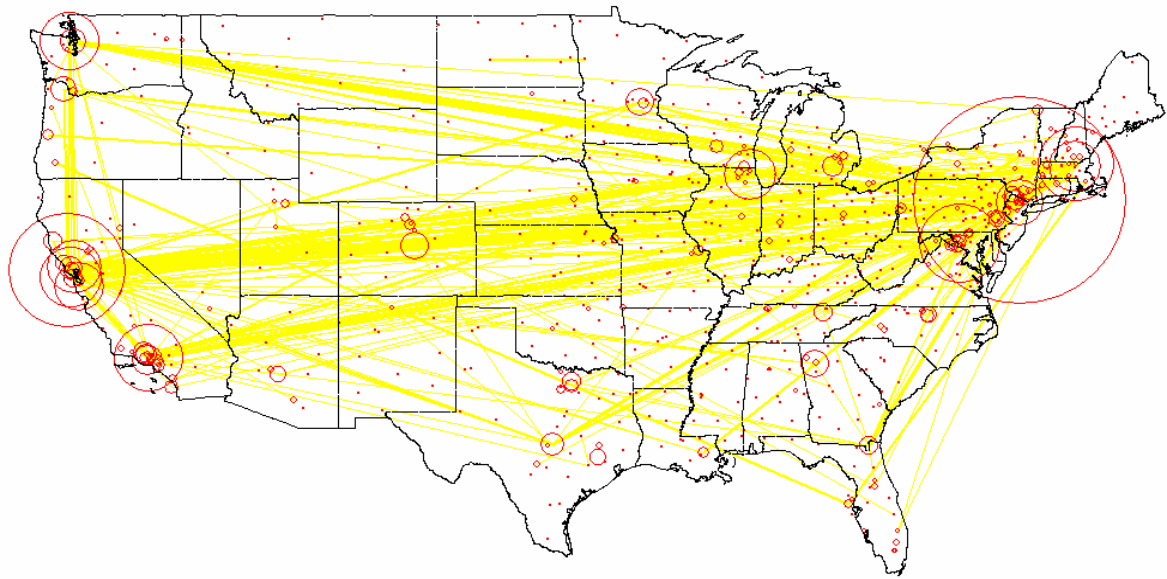


Table 3: Correlation of SSC&O and blog topics (significant at .05 level)

Topics	Correlation with SSC& O
Politics	.302*
Art and letters (high art & books)	.429*
Lifestyle	.337*
Technology	.209