

Exploring Web Co-link Patterns for Business Intelligence: The Case of Two Chinese Industries

Liwen Vaughan

lvaughan@uwo.ca

Faculty of Information and Media Studies
University of Western Ontario, Canada

Juan Tang, Jian Du

jtang@libnet.sh.cn, jdu@libnet.sh.cn

Institute of Scientific and Technical Information of
Shanghai, Shanghai, China

Abstract

The method of Web co-link analysis for business intelligence that was successful in the international environment in previous studies was applied to the Chinese environment to analyze the business competition positions of companies in two Chinese industries. The study was able to map companies into various sectors and show their competitive positions. The mapping result was better when data collection was restricted to the Chinese Web pages only. The finding of the study suggests that the method of co-link analysis may be applicable to a variety of industries and in different countries. The study also found that the Chinese Web pages are very consumer oriented, a phenomenon that was not seen in previous studies of international companies. Websites of companies whose products gear directly toward consumers attracted significantly more hyperlinks than Websites of companies whose products and services are not directly consumer oriented.

Background of the Study

It has been well established that Web hyperlinks contain useful information that can be explored for various purpose. For example, Google (Page, Brin, Motwani & Winograd, 1998) and other major search engines (Sullivan, 2007) all use sophisticated algorithm to rank Web pages based on the numbers and ways that Web pages linking to each other. Patterns of Web page linking have been used to identify Web communities (Flake et al, 2002), to obtain scientific research information (Thelwall, 2005), to get business information (Reid, 2003; Thuraisingham, 2003), and to carry out counter terrorism research (Mooney et. al., 2004). In the case of business world, it was found that Websites of business competitors tend not to link to each other (Vaughan, Gao & Kipp, 2006) but the Website

of a third party such as a customer or a retailer will link to a pair of business competitors, i.e. the Websites of two business competitors will be co-linked by a third party. Analyzing co-link patterns of international telecommunications companies, Vaughan and You (2005) found business information in the form of overall business competition scene. The current study applied the method developed in that earlier study to the Chinese environment to determine if the method is applicable to an environment that has a different language, culture, and business practice. This testing of the method is needed because different cultures may have different Web linking patterns and different business practice may result in different business competition scenes. An added benefit of the study is that it would inform us on the special nature and characterizes of Chinese Websites, which would enrich our knowledge of the Web linking phenomenon in general and guide our future research in particular.

The overall approach of the study is: select a group of companies for the study; locate company Websites; for each possible pair of companies in the study, use a search engine to find all Web pages that have hyperlinks that point to the pair of company Websites (i.e. find the number of co-links of these two companies); construct a co-link matrix of all the companies in the group; and analyze the co-link matrix using a statistical method called Multidimensional Scaling (MDS). The result of the MDS analysis is a map that positions all companies according to their similarities with similar companies positioned closer to each other. The similarity of a pair of company is measured by the number of co-links they have. The more co-links pointing to the two companies, the more similar they are. This is based on the idea that the more two companies are related, the more likely that they will be co-linked. For example, the Website of a computer retail store is likely to link to two computer companies that are their suppliers but it is unlikely that the Website will have links to a computer company and a food company. As similar or related companies are likely to be competitors (two computer companies are competitors but a computer company and a food company are not competitors), the MDS map will effectively place competing companies together. Thus MDS map will effectively show the business competition landscape.

China's two major industries, electronics/information technology industry and chemical industry, were chosen for the study. China's electronics/IT industry

contributes 7.5% to the country's GDP while China's chemical industry is the third largest of the world by 2005 (following U.S.A. and Japan). These two Chinese industries are major international players in these two markets in that they are both major exporters and importers. They are a main focal point of their respective international competitors. These two industries are also very different so they can be used as a contrast to test the co-link analysis method. Each of these two industries has thousands of companies which were too many to be included in the study. It was decided to study the top companies because reliable business information on these companies is more readily available for the study. Details on the company selection are reported in the Methodology section below.

China was chosen as the country of the study for several reasons. First, there is a tremendous interest in business intelligence in China as shown by the existence of a national organization, Society of Competitive Intelligence of China (<http://www.scic.org.cn>), and several commercial portals on competitive intelligence such as China's Network of Competitive Intelligence (<http://www.chinaci.com>). Second, an examination of research literature showed a lack of research on business intelligence methods in the Chinese environment. In China, there is plenty of discussion on business intelligence but relatively little has been done on developing original research methods for business intelligence. Although methods used in producing business intelligence reports are very rigorous, they are traditional methods without the use of Web resources (China's Network of Competitive Intelligence, 2007). Third, we cannot assume that methods of analyzing Web hyperlinks for business information that are developed at the West environment will fit the Chinese environment by default. Not only China has a different language, culture, and business practice, its history of Web development is also different from the rest of the world. Web development started later in China but the pace of development in recent years is astounding. By 2004, China's Web user population size was second only to the U.S. (Chua, 2004). However, a recent study (Vaughan & Zhang, 2007) comparing coverage of Websites of different countries by major search engines found that Chinese sites received lower rate of coverage relative to their U.S. counterparts. The discrepancy was more pronounced in the areas of commercial sites, the type of sites that the current study examines.

Methodology

Some terms need to be defined before discussing the details of the study. Inlinks (also called back links) are links coming into (or pointing to) a Web page. Two different types of inlinks need to be distinguished, total inlinks and external inlinks. Total inlinks include all links pointing to a particular page or a site while external inlinks include only links coming from Websites outside the site in question. In other words, external inlinks do not include links within the site itself, such as the "back to home" type of navigational links within the site. If page X and page Y are both linked to by page Z (i.e. page X and page Y both have inlinks from page Z), then X and Y are co-linked.

Companies in the Study

China's Ministry of Information Technology conducts an annual ranking of top 100 electronics and IT companies. The ranking is based on revenues and other financial measures and ranking results are published at the Website <http://www.ittop100.gov.cn>. The Website also contains other information such as company profiles, industry trend, and business research reports. As this is an official government Website, the site content is considered authoritative and reliable. So it is used as the main information source of the study. At the time of the data collection (winter 2007), the most current ranking was 2006 (China's Ministry of Information Technology, 2006) so this ranking was used. The chemical companies in the study were from the ranked list of China's top 50 chemical enterprises compiled by SRI Consulting, Beijing Office (Alibaba.com, 2006). SRI Consulting is a business research service for the global chemical industry. It publishes research reports and conducts client-sponsored research (SRI Consulting, 2007). It has almost sixty years of history and their reports are considered fairly authoritative and reliable. The ranking of the companies was based on 2004 revenue data. When we started the project in late 2006, this was the most current list publicly available.

The list of electronics and IT companies does not contain company URLs. URLs of these companies were manually searched and carefully verified. The list of chemical companies contained URLs of these companies. Each URL was manually checked and corrected if there is a change of URL at the time of our data collection. Not all 100 electronics/IT

companies can be included in the study as some companies do not have sufficient number of co-links with other companies in the study for a proper MDS mapping. So the companies were first ranked by inlink count and then the top 60 were selected for the study. As was expected, the Websites of chemical companies had on average fewer inlinks than that of the electronics/IT companies so only the top 30 chemical companies were used for the study. See Appendix 1 and 2 for the list of companies in the study.

Search Engine and Query for Data Collection

MSN's Live Search was used for collecting inlink data as two other two major search engines in the market, Google and Yahoo!, could not serve the purpose of the study. Google can only search for total inlinks, i.e. it cannot filter out internal links in the search result. Google (2006) documentation on back link search query states "No other query terms can be specified when using this special query term". The "link" query needs to be combined with the "site" query in order to filter out internal links. Yahoo! was the search engine used in inlink data collection in recent years (e.g. Ortega et al, 2006; Vaughan & You, 2005) after it acquired AltaVista and AllTheWeb. However, at the time of data collection for this study (February and March 2007), we found that Yahoo! did not support co-link search although inlink search still seemed to work. A co-link query such as (link:www.uwo.ca –site:uwo.ca) AND (link:www.ubc.ca –site:ubc.ca) was interpreted by Yahoo! as a keyword search because the word "link" and the URL in the query were bolded in the search result screen as they would be in a keyword search. Many retrieved pages did not have the actual links that the query was meant to search for but had the word "link" on them. So Yahoo! could not be used for this study. MSN China (<http://cn.msn.com/>) could not perform link search either as it also interpreted a link search as a keyword search.

Assume that we are searching for co-links between Websites www.abc.com and www.xyz.com, the query in MSN would be:

(link:www.abc.com –site:abc.com) (link:www.xyz.com –site:xyz.com)

Note that MSN, like other major search engines, adds the Boolean operator AND by default in between query terms so the AND operator is omitted in our data collection. In other words, the query used is effectively:

(link:www.abc.com –site:abc.com) AND (link:www.xyz.com –site:xyz.com).

Data collection was carried out using MSN's API (Application Programming Interfaces) which usually returns the same results as its Web interface (Thelwall, in press). Although data collected from commercial search engines have limitations (Barllan, 2004; Mayr & Tosques, 2005; Thelwall, in press), in practice, there is no way an individual researcher or even a large group of researchers can crawl the whole Web to build up a search engine large enough to rival commercial search engines. For this reason, we used a commercial search engine for data collection. Our results as reported below reflect the industry reality fairly well, which is to some extent an assurance of the validity of our approach.

The "link" command of MSN finds Web pages that link to a particular URL (in this study, links to a company homepage rather than all pages of the company Website). The "linkdomain" command of MSN will search for Web pages that link to all pages of a site including the homepage. We collected data using both the "link" command and "linkdomain" command and compared the results. Since all companies in the study are Chinese companies, we wanted to find out what difference, if any, it would make to restrict data collection to Web pages in the Chinese language. So we collected data in two scenarios, one to include all Web pages regardless of the language and the other to include pages in the Chinese language only. Altogether, there were four sets of data: "link" command and include all pages; "link" command and Chinese pages only; "linkdomain" command and include all pages; "linkdomain" command and Chinese pages only. The outcomes from these four data sets are reported in the "Results" section below.

The co-link data collected need to be normalized to obtain a relative measure of the number of co-links because a co-link count of 5 is very high if the number of links pointing to each Website is 6 while it will be low if the number of links pointing to each Website is 100. The normalization is done through Jaccard Index as follows:

$$\text{NormalizedColinkCount} = n(A \cap B) / n(A \cup B)$$

Where

A is the set of Web pages which links to Website X
B is the set of Web pages which links to Website Y
 $n(A \cap B)$ is the number of pages which link to both Website X and Website Y, i.e. the raw co-link count

$n(A \cup B)$ is the number of pages which link to either Website X or Website Y.

Multidimensional Scaling (MDS) is then applied to the normalized co-link matrices using SPSS. The MDS output is a map that positioned companies according to their normalized co-link counts. The higher the normalized co-link counts between a pair of companies, the closer they will be placed in the MDS map. Essentially the map will cluster competing companies together as competing companies have higher co-link counts.

Results

Link command vs. linkdomain command

For both the electronics/IT companies and the chemical companies, the map from the "linkdomain" command has clearer clusters so the results reported below are based on the data collected using the "linkdomain" command. An earlier study (Vaughan & You, 2005) that examined world major telecommunications companies also compared the results these two commands and found that the result from the "link" command reflects the industry reality better. A follow up study (Vaughan, Kipp & Gao, in press) that qualitatively examined the reasons of co-linking found that links to homepage were more likely to be business related than links to non-homepage which explains why the "link" command had better result. The fact that the "linkdomain" command works better for the Chinese companies in this study suggests that the Chinese pages may have a different reasons for co-linking. Further qualitative study is needed to gain a better understanding of the Chinese co-link pattern.

All Web pages vs. Chinese Web pages

The MDS maps from data collected in these two scenarios are similar but the maps generated from the data of Chinese pages are better. This is true for both industries in the study. To examine this phenomenon further, the two sets of data were compared in terms of inlinks counts. For electronic/IT companies, the global links (links from all pages without the language restriction) were only 12.3% more than the links restricted to Chinese pages only. For chemical companies, this ratio was 11.7%. Given that the search engine used for data collection (MSN) is a global search engine that indexes Web pages around the world, the small number of non-Chinese language pages retrieved in

this study cannot be attributed to a possibility of under representation of non-Chinese pages in this search engine. In other words, the only explanation of the small number of non-Chinese pages is that these company Websites do not attract inlinks from international Websites. We considered the possibility of collecting data excluding Chinese pages, i.e. only including non-Chinese pages and then map this data set to find out how the companies are positioned in the international market. However, this idea is not feasible because excluding Chinese pages will result at very few data points, i.e. a very sparse co-link matrix, which is not appropriate for a MDS analysis.

Business Competition Maps

The MDS maps reported below are based on the "Chinese pages only" data collected with the "linkdomain" command because this data set generates MDS maps that best reflect the industry reality as reported above. Fig. 1 is the MDS mapping result of the electronic/IT industry. In both Fig. 1 and Fig. 2, data points were labeled with abbreviated company names. Please refer to Appendix 1 and 2 for the full names and other information of the companies. Fig. 1 can be divided into two parts: the right side are companies that produce consumer products such as televisions, refrigerators, and home computers while the left side are companies whose products or services (e.g. cables, software outsourcing, and integrated circuits) are not geared directly toward the consumer market. The right side can be further divided into two groups: companies in the upper circle produce home appliances including televisions, home air conditioning machines, and refrigerators. The lower circle encompasses companies that produce personal computers. Unlike the right side, the left side of Fig. 1 does not have clear clusters of companies by products or services. For example, cable companies (e.g. Changfei at the far left, jszt at the lower middle, and hengtong at the left center) are mixed in with other types of companies rather than clustered together.

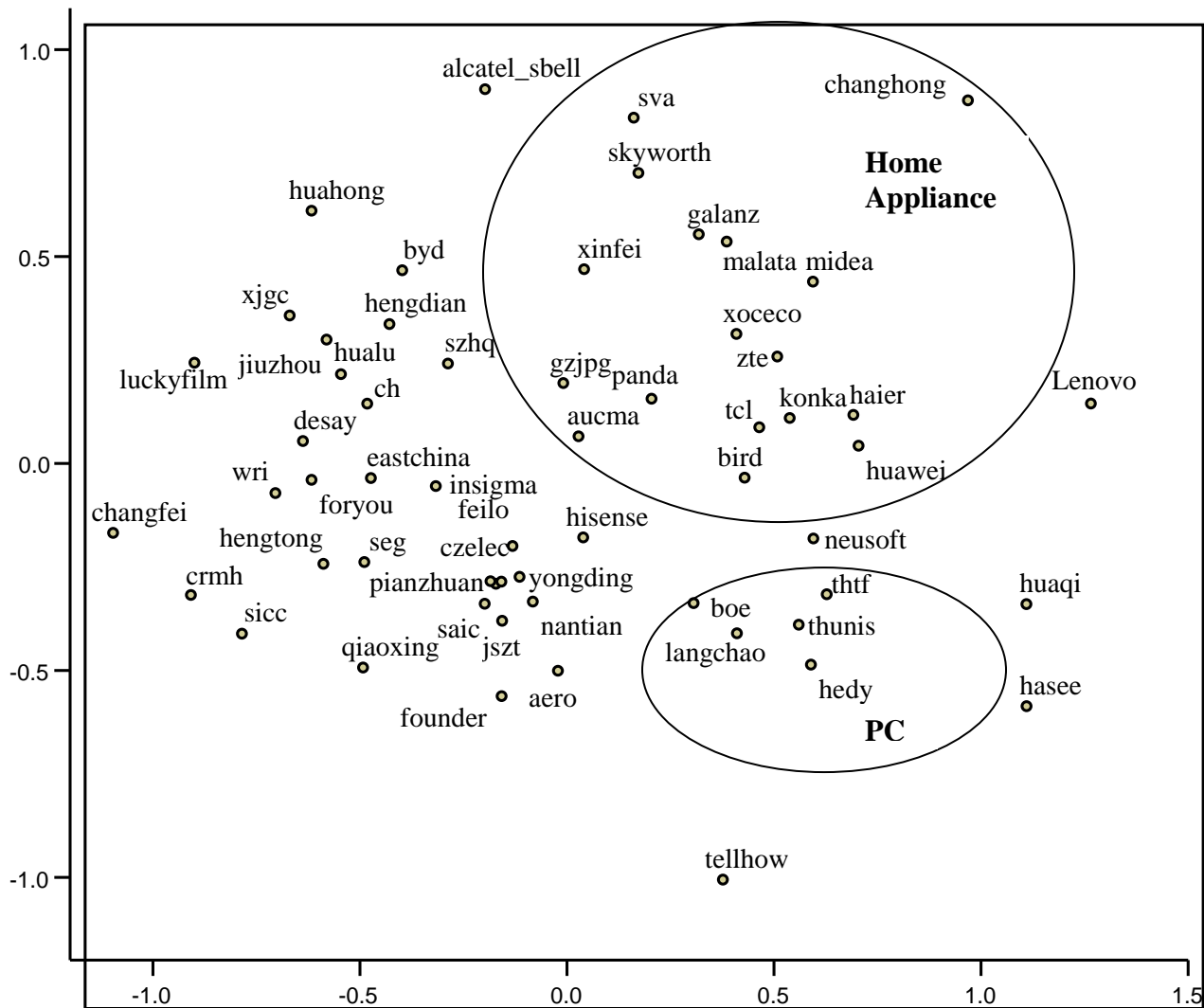


Fig. 1 MDS mapping result of the electronic/IT industry

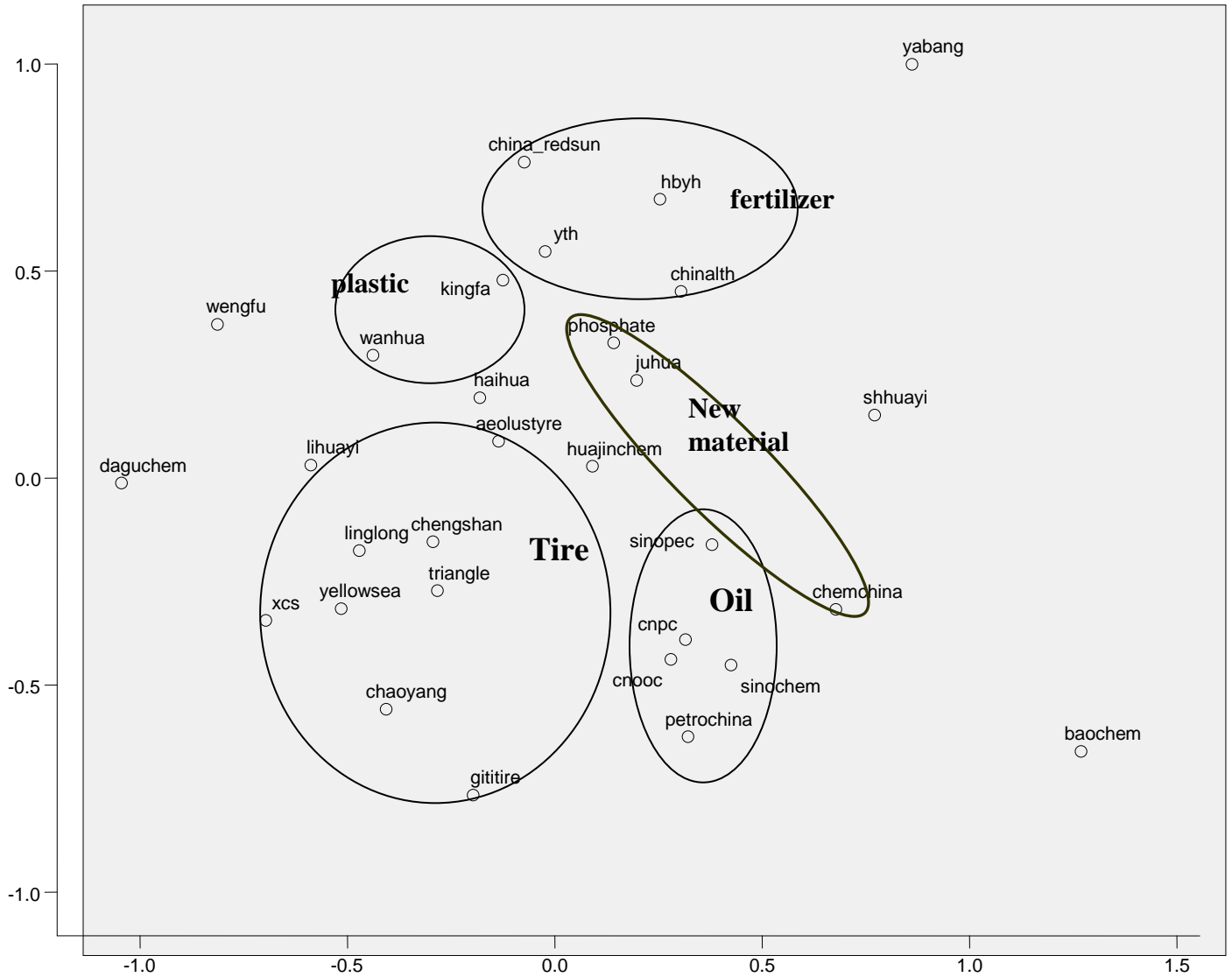


Fig. 2 MDS mapping result of the chemical industry

We hypothesized that the lack of clear clusters on the left side of Fig. 1 is caused by the lack of sufficient number of inlinks to show a statistical pattern of clustering. A large portion of the Chinese Web pages are consumer related (i.e. targeting the consumer market). Companies on the left do not attract links from consumer Web pages because their products and services are not consumer oriented. To test this hypothesis, we compared the number of inlinks to the companies on the left side with that on the right side. A t-test shows that the inlink counts on the two sides are indeed significantly different ($p < 0.05$). The average number of inlinks to the right side is 8579 while that to the left side is only 983! It is clear that companies whose products and services are consumer oriented attracted far more inlinks. For example, Huaqi Information Digital Technology Co., Ltd, attracted 10974 inlinks while Alcatel Shanghai Bell received about half of that number of inlinks, 5609 to be exact. The former is a much smaller company than the latter by various measures. For example, Huaqi is ranked 84 among the top 100 companies while Alcatel Shanghai Bell is ranked 18. The revenue of the former is RMB 1.8 billion compared with RMB 12 billion of the latter. The products of these two companies explains the strong inlink contrast between the two: the main products of Huaqi are MP3 player, digital camera, USB flash drive etc. while the main products of Alcatel Shanghai Bell are telecommunications networks such as voice networks, mobile networks, and broadband networks.

A few outliers on the right side of Fig. 1 are companies that are not competing directly with the companies in the two clusters. The most noticeable one is Lenovo, positioned on the far right. Lenovo is formed as a result of the acquisition by the China's Lenovo Group of the IBM Personal Computing Division. Although its main products are microcomputers (both desk top and notebook), it is not grouped into the PC cluster of Fig. 1. This reflects the fact that other companies in the PC cluster are not really at the same level of competitiveness as Lenovo, a leader in the PC market not only in China but around the world. Lenovo has the largest number of external links among all 100 companies in the study, which matches its number one position in the overall ranking of these companies.

MDS analysis clustered the chemical companies into five distinctive sectors as shown in Fig. 2. The five major oil companies are located in a very small

area reflecting a very competitive oil market. Tire companies and chemical fertilizer companies are well positioned within the respective circles. There are two plastic product companies and they are clustered in a small circle. If were not for ChemChina, the "new material" circle would be much smaller. The reason that ChemChina is somewhat distant to the other "new chemical material" companies is that it is a chemical enterprise that consists of several companies with various products. The main products of ChemChina are new chemical materials but ChemChina also has oil and other chemical products. The few companies that are positioned on the outer skirts of Fig. 2 are those that are not main competitors of the companies within the five circles because they have different products. For example, the main products of Yaband Goup, located on the top right corner of Fig. 2, are dyestuff and paints. The main business of BaoChem, located at the bottom right corner of Fig. 2, is coal tar processing. Huayi Group, located at the middle right edge of the MDS map, consists of several chemical companies with a very broad spectrum of products ranging from chemical devices to biomedicine. MSD map correctly positioned Huayi outside the five main circles.

Discussion and Conclusions

The method of Web co-link analysis for business intelligence that has been employed successfully to international companies in previous studies was used in analyzing the business competition positions of two groups of Chinese companies, the electronics/IT companies and the chemical companies. Statistical method Multidimensional Scaling (MDS) was applied to the co-link matrices of these two groups and the resulting two MDS maps showed the business competition positions of the two industries by clustering companies into sectors of related and thus competing companies. This demonstrates that the Web co-link analysis is applicable to the Chinese Websites, particularly when the data collection is restricted to the Chinese Web pages only. Excluding Web pages of other languages in data collection resulted at a more homogenous collection of pages which is probably the reason of the better result.

Previous studies on international companies found that "link" command (finding links that point to homepage only) generated better results while the current study of Chinese companies found that the "linkdomain" command (finding links that point to all pages of a site not just the homepage) was better.

This suggests that the linking pattern of Chinese Websites is different. It is also worth noting that the two industries in the study are quite different in their inlinks. The median number of inlinks to Websites of electronics/IT companies was 1434 while that of the chemical companies was only 302. The difference is statistically significant ($p < 0.05$) as shown by a Mann-Whitney Test (The Mann-Whitney Test rather than the t-test was used because the frequency distributions of the inlink counts are very skewed). The fact that the co-link analysis method is applicable to Websites of two different industries is reassuring in that the method is probably fairly widely applicable.

An interesting and somewhat unexpected finding of the study is that the Chinese Web pages are very consumer oriented, a phenomenon that was not seen in our previous studies of international companies. Websites of companies whose products (e.g. home appliances) gear directly toward consumers attracted significantly more inlinks than Websites of companies whose products and services (e.g. telecommunications networks) are not directly consumer oriented. This also explains why the Websites of chemical companies attracted much fewer inlinks than that of the electronics/IT companies. Products of the chemical industries are generally not sold directly to consumers and these companies' Websites are not consumer oriented. The consumer orientation of Chinese Websites is probably the result of the huge consumer market of China, the consequence of a huge population size. This consumer market is developing rapidly along with the rapid development of China's economy. Meanwhile, e-commerce is gaining momentum in China as well. Thus competitive intelligence using Web data will be a fruitful area to explore and we will pay particular attention to sites that are consumer oriented in our future research. Specifically, we will qualitatively examine the nature and characteristics of the Chinese Web space, particularly consumer related sites, and compare them with Western Websites. Another direction of the project is to apply more advanced Web data mining techniques, such as combining Web co-link analysis (a type of Web structure mining) with Web keywords analysis (a type of Web content mining), to the Chinese Websites in order to gain more in depth business intelligence information.

Acknowledgement: This study is part of a larger project funded by the Initiative on the New Economy (INE) Research Grants program of the Social Sciences and Humanities Research Council of

Canada (SSHRC). Research assistants Karl Fast and Gloria Liu helped with data collection. Senior chemical engineer Lei Xu of Science and Technology Information Institute of Shanghai Chemical Industry helped with the interpretation of the results of the chemical industry.

References

Alibaba.com (2006). *China's Top 50 chemical enterprises among Asian's top chemical enterprises ranked by revenue*. Retrieved Jan. 8, 2007 from <http://info.china.alibaba.com/news/detail/v4-d5871708-p2.html#newsdetail-content>.

Bar-Ilan, J. (2004). Search engine ability to cope with the changing web. In Levene, M. & Poulouvasilis, A. (Eds.), *Web dynamics*. Berlin: Springer-Verlag.

China's Ministry of Information Technology. (2006). *The 20th Annual Ranking of the Top 100 Electronic and Information Technology Companies*. Retrieved Dec. 17, 2007 from <http://www.ittop100.gov.cn/detail?record=1&channelid=1270&presearchword=ID=188766>.

China's Network of Competitive Intelligence. (2007). *Our work procedure*. Retrieved Aug. 1, 2007 from <http://www.chinaci.com/about/flow.htm>.

Chua, C. H. (2004). *China: Number of Net users in China soars to 87 million*. Retrieved July 31, 2007 from <http://www.asiamedia.ucla.edu/article.asp?parentid=12909>.

Flake, G.W., Lawrence, S., Giles, C.L. & Coetzee, F.M. (2002). Self-Organization and Identification of Web Communities, *IEEE Computer*, 35(3), 66-71.

Google (2006). *Google SOAP Search API Reference*. Retrieved Aug. 1, 2007 from http://www.google.com/apis/reference.html#2_2.

Mayr, P. & Tosques, F. (2005) *Google Web APIs – an Instrument for Webometric Analyses?* Retrieved Aug. 1, 2007 from http://www.ib.hu-berlin.de/~mayr/arbeiten/ISSI2005_Mayr_Tosques.pdf.

Mooney, R. J., Melville, P. Tang, L. R., Shavlik, J., de Castro Dutra, I., Page, D. & Costa, V. S. (2004). Relational data mining with inductive logic programming for link discovery. In Kargupta, H. et.

al. (Eds). *Data mining: next generation challenges and future directions*. Cambridge, Mass.: MIT Press, 239-254.

Ortega J. L., Aguillo, I., Cothey, V., & Scharnhorst, A. (2006). Maps of the academic web in the European Higher Education Area - an exploration of visual web indicators. In *Book of Abstracts, 9th International Conference on Science & Technology Indicators*, pp. 107-110, Leuven, Belgium, Sept. 7-9, 2006.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). *The PageRank citation ranking: Bringing order to the Web*, retrieved Sept. 21, 2006 from <http://citeseer.ist.psu.edu/page98pagerank.html>.

Reid, E. (2003). Using Web link analysis to detect and analyze hidden Web communities. In D. Vriens (Ed.), *Information and communications technology for competitive intelligence*, Hilliard, OH: Ideal Group, 57-84.

SRI Consulting. (2007). *About Us*. Retrieved Aug. 21, 2007 from <http://www.sriconsulting.com/SRIC/Public/Aboutus.html>.

Sullivan, D. (2007). *How Search Engines Rank Web Pages*. Retrieved Aug. 1, 2007 from <http://searchenginewatch.com/showPage.html?page=2167961>.

Thelwall, M. (in press). Extracting accurate and complete results from search engines: Case study Windows Live. To appear in *Journal of the American Society for Information Science and Technology*.

Thelwall, M. (2005). Scientific Web Intelligence: Finding relationships in university webs. *Communications of the ACM*, 48(7), 93-96.

Thuraisingham, B. (2003). *Web Data Mining and Applications in Business Intelligence and Counter-Terrorism*. Boca Raton, Florida: CRC Press.

Vaughan, L., Gao, Y. & Kipp, M. (2006). Why are hyperlinks to business Websites created? A content analysis. *Scientometrics*, 67(2), 291-300.

Vaughan, L., Kipp, M., & Gao, Y. (in press). Are co-linked business Websites really related? A qualitative study. To appear in *Online Information Review*.

Vaughan, L. & You, J. (2005). Mapping business competitive positions using Web co-link analysis. In *Proceedings of ISSI 2005 – the 10th International Conference of the International Society for Scientometrics and Informetrics*, (P. Ingwersen & B. Larsen eds.), pp. 534-543, Stockholm, Sweden, July 24-28, 2005.

Vaughan, L. & Zhang, Y. (2007). Equal representation by search engines? A comparison of Websites across countries and domains. *Journal of Computer-Mediated Communication*, 12(3), available at <http://jcmc.indiana.edu:80/vol12/issue3/vaughan.html>.

APPENDIX 1 Electronics/IT Companies in the Study

Rank	Company Name	URL	Label in Fig. 1
1	Lenovo Group	http://www.lenovo.com.cn	Lenovo
2	Haier Group	http://www.haier.com	haier
3	BOE Technology Group Co.	http://www.boe.com.cn	boe
4	TCL Corporation	http://www.tcl.com	tcl
5	Huawei	http://www.huawei.com.cn	huawei
6	Midea Group	http://www.midea.com.cn	midea
7	Hisense	http://www.hisense.com.cn	hisense
8	Sva Group	http://www.sva.com.cn	sva
9	Panda Electronics Group Co., Ltd.	http://www.chinapanda.com.cn	panda
10	Founder Group	http://www.founderpku.com	founder
11	ZET Corporation	http://www.zte.com.cn	zte
12	Changhong	http://www.changhong.com.cn	changhong
13	Huaqiang Holdings Ltd	http://www.szhq.com	szhq
15	Galanz Group Co. Ltd.	http://www.galanz.com.cn	galanz
16	Skyworth Group Co., Ltd.	http://www.skyworth.com	skyworth
17	Inspur Group	http://www.langchao.com.cn	langchao
18	Alcatel Shanghai Bell	http://www.alcatel-sbell.com.cn	alcatel_sbell
19	Desay Corporation	http://www.desay.com	desay
20	Konka Group Co., Ltd	http://www.konka.com	konka
22	Foryou Group	http://www.foryougroup.com	foryou
23	Tsinghua Tongfang	http://www.thtf.com.cn	thtf
24	Ningbo Bird Company	http://www.chinabird.com	bird
26	Shenzhen Electronics Group CO., Ltd	http://www.seg.com.cn	seg
27	Xiamen Overseas Chinese Electronic Co., Ltd.	http://www.xoceco.com.cn	xoceco
28	Shanghai Feilo Co., Ltd	http://www.feilo.com.cn	feilo
30	Qindao Aucma Co., Ltd	http://www.aucma.com.cn	aucma
31	BYD Company Limited	http://www.byd.com.cn	byd
32	Cosun Group	http://www.qiaoxing.net	qiaoxing
33	Hengtong Group	http://www.hengtonggroup.com	hengtong
34	Wanlida Group Co., Ltd.	http://www.malata.com	malata
36	XJ Group Corporation	http://www.xjgc.com	xjgc
37	East China Electronics Group Co., Ltd	http://www.hdeg.com	eastchina
39	ETERN Group Ltd.	http://www.chinayongding.com	yongding

40	IRICO Group Corporation	http://www.ch.com.cn	ch
41	Frestech	http://www.xinfei.com	xinfei
43	Insigma Technology	http://www.insigma.com.cn	insigma
44	Huahong Group	http://www.huahong.com.cn	huahong
47	China Hualu	http://www.hualu.com.cn	hualu
49	Zhongtian Technologies Co., Ltd.	http://www.jszt.com.cn	jszt
53	Tsinghua Unisplendour Corporation Limited	http://www.thunis.com	thunis
58	Neusoft	http://www.neusoft.com	neusoft
59	Hengdian Group	http://www.hengdian.com	hengdian
60	FiberHome Technologies Group	http://www.wri.com.cn	wri
61	HEDY Holding	http://www.hedy.com.cn	hedy
62	Jiuzhou Electric Group Co., Ltd	http://www.jiuzhou.com.cn	jiuzhou
63	China Silian Instrument Group Co. Ltd	http://www.sicc.com.cn	sicc
64	China Zhenhua Electronics Group	http://www.czelec.com.cn	czelec
65	Aisino Corporation	http://www.aero-info.com.cn	aero
66	Yangtze Optical Fibre and Cable Company Ltd.	http://www.changfei.com.cn	changfei
67	Shanghai Jinling Co., Ltd.	http://www.jin-ling.com	jin-ling
70	China Lucky Film Corporation	http://www.luckyfilm.com	luckyfilm
71	Pianzhuan Group	http://www.pianzhuan.com.cn	pianzhuan
73	China Resources Microelectronics (Holdings) Ltd	http://www.crmh.com.cn	crmh
82	Hasee Computer Co., Ltd	http://www.hasee.com	hasee
83	Jinpeng Group Co., Ltd.	http://www.gzjpg.com	gzjpg
84	Huaqi Information Digital Technology Co., Ltd.	http://www.huaqi.com	huaqi
86	Nantong Fujitsu Microelectronics., Ltd	http://www.fujitsu-nt.com	fujitsu
90	Nantian Electronics Information Corp, Ltd.	http://www.nantian.com.cn	nantian
92	Tellhow Sci-tech Co., Ltd.	http://www.tellhow.com	tellhow
94	Shanghai Automation Instrumentation CO., Ltd.	http://www.saic.sh.cn	saic

APPENDIX 2 Chemical Companies in the Study

Rank	Company Name	URL	Label in Fig. 2
1	China Petroleum & Chemical Corporation (Sinopec)	http://www.sinopec.com.cn	sinopec
2	China National Petroleum Corporation (CNPC)	http://www.cnpc.com.cn	cnpc
2	PetroChina Company Ltd	http://www.petrochina.com.cn	petrochina

3	Sinochem Corporation	http://www.sinochem.com	sinochem
4	China National Offshore Oil Corporation	http://www.cnooc.com.cn	cnooc
5	Shanghai Huayi (Group) Company	http://www.shhuayi.com	shhuayi
7	Shandong Haihua Group Co.Ltd.(SHG)	http://www.haihua.com.cn	haihua
9	GITI Tire	http://www.gititire.com	gititire
11	Shandong Chengshan Group Co., Ltd.	http://www.chengshan.com	shengshan
12	Liaoning Huajin Chemical Industry Group Co., LTD	http://www.huajinchem.com	huajinchem
13	Hangzhou Zhongce Rubber Company Limited	http://www.chaoyang.com	chaoyang
16	Jiangyin Chengxing Industrial Group Co. , Ltd	http://www.phosphatechina.com	phosphate
19	Yuntianhua Group	http://www.yth.com.cn	yth
21	Triangle Group	http://www.triangle.com.cn	triangle
22	Juhua Group Corporation	http://www.juhua.com.cn	juhua
23	Baochem Group	http://www.baochem.com	baochem
25	Tianjin Dagu Chemical Corporation Ltd	http://www.daguchem.com	daguchem
28	Shandong Linglong Rubber Co., Ltd	http://www.linglong.cn	linglong
31	Red Sun Group Corporation	http://www.china-redsun.com	china_redsun
32	Lihuayi Group Co. Ltd	http://www.lihuayi.com	lihuayi
34	Yantai Wanhua Group	http://www.wanhua.com.cn	wanhua
36	Lutianhua Group Inc.	http://www.chinalth.com	chianlth
37	Aeolus Tyre Co., Ltd.	http://www.aeolustyre.com	aerolus
38	Qingdao Huanghai Rubber Group	http://www.yellowsea.com.cn	yellowsea
42	Guizhou Wengfu Chemi-Phos Imp. & Exp. Corp.	http://www.wengfu.com	wenfu
43	Yabang chemical industry Corporation (Group)	http://www.yabang.com	yabang
44	ChemChina Group Corporation	http://www.chemchina.com.cn	chemchina
46	Hubei Yihua Chemical Industry Co., Ltd.	http://www.hbyh.cn	hbyh
47	Guangzhou Kingfa Sci. & Tech. Co., Ltd.	http://www.kingfa.com.cn	kingfa
50	Xiamen Cheng Shin Rubber Ind.,Ltd (XCS)	http://www.xcs.com.cn	xcs