

Hyperlink Analysis of E-commerce Websites for Business Intelligence: Exploring Websites of Top Retail Companies of Asia Pacific and USA

Rathimala Kannan¹ and Marthandan Govindan²

Multimedia University, Faculty of Management, ¹rathimala.kannan@mmu.edu.my
²marthandan@mmu.edu.my

Received 1 June 2010; received in revised form 17 December 2010; accepted 3 March 2011

Abstract

Hyperlinks, which connect web pages on the World Wide Web, are rich sources of hidden information. E-Commerce Websites, which are created for different purposes from online sales to company promotion, would benefit if they receive more links from other websites as this would lead to increase the traffic to these websites. This paper analyses the structure of e-commerce websites using webometric approach to uncover any hidden information from the hyperlinks. The top 50 retail companies' e-commerce websites each from Asia Pacific and USA are chosen for this study. Our results found a positive relationship between the external inlinks count pointing to a retail company e-commerce website and one of its business measures, sales. But no association has been found between hyperlink metrics and business measure like revenue. However this conclusion does not hold good for all categories of companies. Comparing the web presence, US private retail companies are more visible on the Web than the Asia pacific retailers. Furthermore this study has found that counts of links pointing to a retail websites are positively correlated with the website age. That is older websites in English language received more external inlinks. Such a correlation does not exist for Japan, China and Korean language websites.

Keywords: Business intelligence, Hyperlink metrics, E-commerce websites, Web presence, Website age, World Wide Web

1 Introduction

In this decade the World Wide Web (WWW) has been extensively studied by many researchers in various dimensions. Webometrics, a new emerging field, is a study about investigating the nature and properties of the World Wide Web to get knowledge about the structure of the Web, usage patterns and various Web measures [2], [26]-[27]. Webometrics research includes page content analysis, web link structure analysis, web usage analysis and web technology analysis [3]. The research on the structure of WWW came out with a lot of interesting and potentially useful information about the online organizations, navigation patterns inside the website, and the characteristics of the website owner [17]. The website quality, user friendliness and smooth navigation can all be altered and improved when this dynamic information is available to the website designers and its owners.

The website is an infomediary between the enterprise and its customers. The online customers do conduct their transactions through the website only and its quality, user-friendliness and navigability play vital role in customer satisfaction, which leads to customer loyalty. More importantly the presence of a website should be known to the online users to enable them to make use of the website. The search engines, web advertisements, web infomediaries and other physical media help to provide the web links (Uniform Resource Locator) to shoppers. The virtual player like search engines use the hyperlinks to redirect the customer clicks to the respective company website. When the online users visit the home page of a company website, it is a requirement for the company to retain the user for some time make him/her to navigate the different pages to locate the product / service / information that he / she is looking for. If the company is successful in that attempt, it is bound to attract more repeated customers and new online users [9], [21].

If a website has more incoming links, for example, it is likely to be more visible to search engines and thus more easily found by buyers searching for a particular product or service [6]. A study analyzed web log data to find if there are any significant differences between the time spent on a websites and number of clicks with regard to the point of access that is from search engine and the URL of the website. The study found that there was a significant difference between the different origins of the visits. The search engines's users spent most time and did more clicks and concluded that search engines are the principal intermediaries to a website [15]. For popular search engine such as Google, high ranking is given to a website which receives most inlinks from other quality websites. The study [5] examined the impact of links on website traffic. They modelled the number of visits to individual internet retail stores as a function of the number of links each has with other sites and the information/product content of the internet retail store. The results demonstrate that the number of links to a particular internet retailer stores explains over 60% of the variance of site traffic between sites. Our previous study confirms that more hyperlinks lead to more traffic to a website [13]. Therefore more hyperlinks lead to more clicks for the website, is a necessary condition. Also more traffic demands more hyperlinks. The bottom line of this approach is that with the increase in customer base, the online sales will go up and profitability of the company will shoot up. All these things may be possible if and only if the quality of the website, navigation friendliness etc are to the taste of the online users.

This research paper aims to investigate possible associations between the hyperlink counts of retail companies' websites and their business performance, the difference in Web presence among USA and Asia Pacific retailers, and the relationship between the age of the website and the hyperlinks pointing to the retail company's website. To answer these research questions, websites of the top 50 retail companies of Asia Pacific Region and top 50 largest private retail companies of U.S.A are considered in this research.

2 Related Research

Earlier studies about the hyperlink structure found that mining the hyperlink structure gave a global understanding of the way the independent users had built connections between themselves [7], [23], [32]. Extensive studies have been done about the hyperlink structure in academic domain and only a few studies are available in government websites and commerce websites.

Studies about the number of hyperlinks pointing to a university website found significant correlation between the total number of links targeting to the university's websites and the research productivity of the university [18], [22], [25]. For example a study about websites of European life sciences research teams found association between the site inlink counts and relevant owner characteristics such as web presence size and research group size [1]. Another study [29] found significant correlation between the number of external inlinks and the journal impact factor for Library and Information Science (LIS) journals' websites. A study about electronic journals and New Zealand university websites found positive correlation between the extent of metadata use by a site and website's Impact Factor [20].

Although hyperlink parameters have been studied widely, very few studies have shed light on e-commerce domain in contrast to academic domain. Earlier studies of commercial websites showed that hyperlinks that are pointing to a company website contain useful business information [26]-[28], [30]-[31]. A study about Amazon.com indirectly revealed the relationship between hyperlinks and mutual trust among online consumers [14]. Another study

examined e-commerce websites using hyperlink method, found that the number of incoming links was strongly correlated with the use, prominence of Trusted Third Parties (TTP) and privacy statements [16].

Studies about China's top 100 information technology (IT) companies' websites found that link count to a company's website correlates with the company's revenue, profit, and research and development expenses [18]. But the link count study about China's top 100 privately owned companies found no relationship between link count and the business performance due to its heterogeneous nature of the companies. Earlier studies [26]-[27], [30] explored the websites of IT companies of China, US, and Canada and reported that there existed a significant correlation between links count and the company's revenue and profit. Also the study found that the correlation coefficients for the USA and Canada were similar although the two groups of websites had very different characteristics.

Among various categories of websites, the dot com domain numerically dominates the WWW and it will be interesting to know the valuable information that is hidden in the hyperlink structure of e-commerce domain. To explore this fact, earlier hyperlink studies have considered only IT [26]-[27], [30] and Telecommunication companies' websites [31]. However it is unlikely that these findings would be appropriate for other companies from different categories such as retail business and other commercial establishments. Therefore this study tries to fill that gap by studying different commercial establishment websites to explore for possible association between web structure parameters and the business performance.

3 Methodology

This study has been carried out in five stages. The identification of the retail companies among the total companies was the first task. In this process, the first 50 retail companies of USA are obtained from the America's top 500 largest private retail companies list published in 2008 by Forbes magazine. First 50 retail companies of Asia Pacific have been selected from Asia Pacific's Top 500 Retail companies of 2008 published by Euromonitor International. In the second stage the official e-commerce websites of the selected retailers were identified and verified manually. In this step the home page URL (Uniform Resource Locator) of all chosen retail companies were collected. For the USA retail companies, the URL is available from Forbes magazine. For Asia Pacific companies, the URL is collected from Google search engine and verified manually. The third stage was to gather business information of those companies. For the USA retail companies, the annual revenue and total number of employees were obtained from Forbes magazine. For the Asia Pacific retailers, total annual sales were obtained from Euromonitor magazine. In the fourth stage we collected hyperlink data using search engines. Yahoo Site Explorer from Yahoo! has been used to collect the hyperlink data such as total number of web pages, total number of inlinks and total number of external inlinks for all the websites. In the last stage, we calculated correlation coefficients using SPSS software to determine the possibility of significant relationship between the hyperlink data and the companies' business information.

This study quantitatively analyzes the hyperlink structure of commerce websites using webometric measures such as counts of inlinks, external inlinks, number of pages in a website and two standard web impact factors. Using these measures, this study investigates whether these measures have any association with the business information of the retail companies and their visibility on the web.

Reliability of a search engine count estimates can be tested by parallel-test reliability which is comparing the quantitative results from different search engines. Earlier research [12] about the quality of the Search Engine Count Estimates (SECEs) examined whether and to what degree SECEs meet the goodness criteria of objectivity, reliability, and validity strived for in scientific investigations. He suggested that reliability of a search engine count estimates can be tested by comparing the SECEs of different search engines and confirmed that, among leading search engines the parallel-test reliability of SECEs is high. This finding was again supported by another study [24] which quantitatively compared search engine results, found that even though the hyperlink counts are different from various search engines; the results are consistent among them. Thelwall recommended Yahoo! search engine to collect data for webometric studies. Our previous study about hyperlink analysis of the commerce websites also confirmed the above finding [13]. Hence the hyperlink data for this study is collected from Yahoo site explorer, which ensures the reliability of hyperlink data. Validity concerns the degree to which a measure expresses a phenomenon it is taken to reflect. Earlier literatures show that hyperlink studies have been conducted across many disciplines and topics and used webometric measures to analyze the hyperlink structure [5]-[6], [8], [11], [13], [17]. As a result of these studies, it can now be concluded that hyperlink counts could be used to study the website hyperlink structure to extract hidden information and thus the validity is ensured.

3.1 Top Retail Companies of Asia Pacific and USA

To explore e-commerce websites' hyperlink structure, two different sets of companies have been selected as explained above. The first set is from Euromonitor International which is the world's leading independent provider of business intelligence on industries, countries and consumers. The companies are ranked based on their sales generated in 2007. In the absence of any other business information, we used the total sales amount as the business performance measure.

The second data set included America's first 50 retail companies from Forbes magazine (a primary competitor for Fortune business magazine). Since additional information is available for these companies, they were ranked based on their revenue and the total number of employees. And these two measures were used as business information here.

The official websites of Asia Pacific and USA retail companies are obtained from search engines and verified manually for their existence. For the Asia Pacific retailers, websites with the regional language are considered. While assigning ranks, group-companies were considered as one company and were assigned the same rank. By this means, such company's individual websites are considered with the equal ranking. For example the Kojima and MaxValu supermarkets are under Aeon Group and are ranked 17 in the list. Their respective Websites got the rank 17. In cases like The Metcash Company which has a grocery store, liquor retail and a wholesale, a different approach is adopted. Treating the supermarket and liquor business as two separate businesses, different rankings are given to them even though these operations are done by the same company. Thereby they got rank 20 for its supermarket and 39 for liquor business. But there is only one website that exists for this company. In order to have a fair deal, we have taken the larger of the two rankings.

3.2 Hyperlink Data Collection

Earlier studies have used popular search engines such as Google, Alta vista, Live Search, and Yahoo! for hyperlink data collection. Vaughan found that, though different search engines reported varied inlink counts for the same type of query, they were all highly correlated [27]. Another study about quantitative comparisons of search engine results [24] concluded that quantitative results from three search engines Google, Yahoo!, and Live Search were mostly consistent. Also this study strongly recommended Yahoo! for the hyperlink study. As the popular search engines give similar results with consistency, we want to fall in line with other researchers and therefore we used Yahoo Site Explorer to collect the total number of web pages, total number of inlinks and total number of external inlinks for all the websites in our study.

Here inlinks count refers to the number of web pages which have one or more hyperlinks pointing to a website. The total number of external inlinks is the counts of links which are pointing to the website from other websites. That is external inlinks count is the difference between the total number of inlinks and the total number of website' self links and hence must be less than the total inlinks for any website. Self links are created for the smooth navigation purpose within the website and therefore are not included in hyperlink analysis. But few websites showed higher external inlinks than the total number of inlinks and for this reason they are not included in further analysis. Also websites having only one webpage are removed from the data set. Finally our data set contains 47 retailers' websites from Asia Pacific and 41 retailers' websites from USA, totalling 88 websites.

3.3 Hyperlink Metrics

This study makes use of hyperlink metrics of a website such as total number of web pages, the total number of inlinks, and total external inlinks to further the analysis. Using these metrics, we calculated standard Web Impact Factor (WIF), a measure for the website visibility and ratings using the formula suggested by Ingwersen [10]. Two standard WIFs named as WIF_{Tin} (WIF of Total inlinks) and WIF_{Te} (WIF of Total external inlinks) are calculated as shown in the equation 1.

$$\begin{aligned} WIF_{Tin} &= \frac{\text{Total inlinks}}{\text{Total number of web pages}} \\ WIF_{Te} &= \frac{\text{Total external inlinks}}{\text{Total number of web pages}} \end{aligned} \quad (1)$$

To measure the impact of websites, Web Impact Factor (WIF) was developed by Ingwersen and defined as the ratio of links made to a website to the number of pages at the websites. Among the two standard WIFs, external WIF appears to be the best valid measure of impact factor. This is very much similar to Google's concept of page rank. Also WIF is the extended concept of impact factor, introduced by Eugene Garfield, who pointed out that, WIF is analogous to the Citation Impact Factor (CIF) proposed by him [11].

Though WIF has been developed a decade ago, it is still used to measure the impact of a website in webometric study. In view of the fact that, this type of study involves data set which is restricted with only the hyperlinks counts, as of now we can use only WIF to measure the impact of the websites. A number of recent webometric studies have also used WIF as a metric to measure the impact of a website. For example to quantitatively evaluate the visibility of a Website, [6] proposed a method which defines a plan of action for the improvement of Web visibility and therefore, the assurance of greater visitor numbers. Their method accompanied by a set of measures which have been identified from the authors own experience and are based on actual requirements. The authors have also implemented the proposed method in a public tool called Know Your Visibility (Site 1), which calculates the visibility of a website. WIF is one of the topology measures proposed by the authors to compare the visibility of a website. They also suggest that to increase the visibility of a website, it is important to attract more external inlinks to the website.

A study about the Internet Visibility of Localized Websites of Twelve Global Brands in Japan, Korea and Hong Kong used WIF to evaluate the visibility of those websites [8]. Another study about the characteristics and link structure of a national scholarly Web space comprising top ranking universities and government supported research institutions in South Korea used WIF as one of the webometric measure to study the scholarly communication activity and linking behaviour among the web space [4]. Our previous study about the hyperlink structure of the Top 50 Small USA companies used WIFs as measures and investigated the hyperlinks pointing to those commerce companies on the web to determine any association with the business performance of those companies [13].

4 Results and Discussion

Using the hyperlink data, the two WIFs; WIF_{Tin}, WIF_{Te} are calculated for both USA and Asia Pacific Retailers' websites. Preliminary statistical analysis of the dataset shows that the frequency distribution of the hyperlink data and the business measures are very much skewed thus are not normally distributed. Therefore the non-parametric Spearman correlation coefficients tests were conducted in this study.

To determine the association between the hyperlinks counts pointing to retail company website and its business information, Spearman Correlation coefficients test is conducted using SPSS software for both data sets. Table 1 and Table 2 show the correlation coefficients between counts of hyperlinks and business information for Asia Pacific Retailers and USA respectively. In these tables, Tp refers total number of pages in a website, Tin refers total inlinks and Te refers total external inlinks.

Table 1: Correlation coefficients of Asia Pacific retailers (47 websites)

	Tin	Te	WIF _{Tin}	WIF _{Te}
Sales r	0.214	0.239	0.096	0.097
p	0.149	0.105	0.520	0.515

From Table 1 it is clearly seen that none of the correlation coefficients are significant at 0.01 levels. Therefore the result indicates that there is no correlation between the hyperlink metrics; total number of inlinks, total number of external inlinks and the two WIFs with the business measure sales for Asia Pacific Retailers.

Table 2: Correlation coefficients of America's largest private retail companies (41 websites)

	Tin	Te	WIF _{Tin}	WIF _{Te}
Revenue r	0.247	0.242	0.160	0.137
p	0.119	0.127	0.319	0.392
Employees r	0.305	0.304	-0.015	-0.020
p	0.053	0.053	0.928	0.903

Similar results are seen for US companies also (Table 2) where the results found no significant relationship between the business performance measure revenue, the companies' total number of employees and the hyperlink metrics for retail companies' e-commerce websites. This may be due to fact that the retailers' websites in our study consist of various categories, and this might have a significant role in identifying the relationship between the hyperlink metrics and the total sales of the retail companies. This result is in line with the findings of the earlier study [30] which found no relationship between link counts and the business performance measure in the heterogeneous nature of websites. Similar results are found in a earlier study [19] which analyzed 138 US Web based businesses to find the Internet positioning and performance of e-tailors'.

However, the revenue of a company could be affected by other variables such as company size which can be measured by assets, properties or number of employees working in a company etc. For example, even when all other variables are equal between the two retail companies, the company with many employees will have more revenue than its rival company. Therefore in this study a measure named revenue per employee is calculated and tested for any relationship with the hyperlink metrics. Company size is measured by the number of employees. Nevertheless this measure also found no correlation with hyperlink metrics of the American retail companies' websites.

We further wanted to investigate the e-commerce websites category-wise and conducted the Spearman correlation coefficients test. Among Asia Pacific retailers, there were three categories of websites; department stores (21 Websites), supermarket/hypermarket (15 Websites) and hardware stores (11 Websites). Table 3 shows the result for Asia Pacific Supermarket companies. It provides evidence that there is a strong correlation between the business information sales and hyperlinks counts; total number of inlinks (0.652 p=0.008) and total number of external inlinks (0.627 p=0.012) of the supermarket websites. Tin and Te correlation coefficients are significant at 0.05 levels. This result indicates that a retail company with high sales attracts more external inlinks to its company website. This does not imply however, that there is a cause-and-effect relationship between the two.

Table 3: Correlation coefficients of Asia Pacific supermarket companies (15 Websites)

	Tin	Te	WIF_Tin	WIF_Te
Sales r	0.652**	0.627*	0.097	0.038
p	0.008	0.012	0.732	0.894

Our result substantiates the previous studies by Vaughan who found strong relation between the inlinks count pointing to a company website and its business information [26]-[27], [30]-[31]. From our results we can say that hyperlink metrics have a positive relationship with the business information of retail companies of the same group. This result explains that the companies' with higher sales have more links to their websites' and thus popular on the WWW. However this study found no correlation between sales and the WIFs.

To further investigate the usefulness of hyperlink metrics in determining the business information of a company, we classified the websites of the top US retail companies' based on the business type. Since there were more than 12 categories with each category having less than 5 websites, we were constrained to calculate the Spearman correlation coefficients only for the Food and Grocery companies having 12 e-commerce websites. The results are depicted in Table 4.

Table 4: Correlation coefficients of America's top food and grocery retail companies (12 websites)

	Tin	Te	WIF_Tin	WIF_Te
Revenue r	0.147	0.133	-0.105	0.091
p	0.649	0.681	0.746	0.779
Employees r	0.880**	0.894**	-0.880**	-0.246
p	0.000	0.000	0.000	0.440

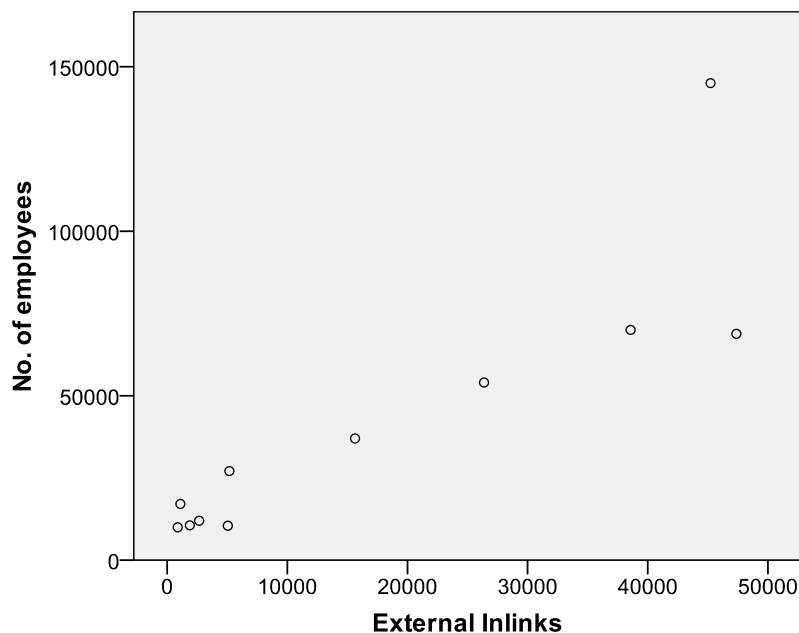


Figure 1: External inlinks Vs number of employees (US top food and grocery retail companies)

Websites of food and grocery companies from US retail companies' exhibit strong positive correlation between total number of inlinks, total number of external inlinks and total number of employees. The Spearman correlation coefficients are 0.880, and 0.894 respectively significant at 0.01 levels. This is possible when the employees in companies are allowed to have a personal webpage and a link to the working company. Thus, increase in the number of employees in a company subsequently will boost the number of inlinks to the company's website and this is revealed in Figure 1. Also in this way the company's visibility on the Web will subsequently increase. This result confirms a similar study by Vaughan who explored websites features of all IT companies in USA and Canada for business information [26]. Therefore to increase the Web visibility, our findings recommend companies to encourage their employees to have personal website and a link to the company's website.

However, this study found no relationship between revenue and hyperlink metrics for America’s top retail companies. This result contrasts with earlier studies about IT and Telecommunication companies [26], [31]. This might be due to the nature of IT and Telecommunication companies. Basically these two industries are more dependent on the usage of Internet and technology which might have caused a positive relationship between the links count to the business information profit and revenue. Whereas being a food and grocery companies, more of its customers prefer to choose the physical stores rather than an e-commerce website and thus no association between link counts and revenue. This finding shows that the relationship found between the number of inlinks to a commercial website and the company’s business performance measures do not hold for all companies regardless of the nature of the company.

To shed more light on the association between sales and external inlinks in country-wise, among Asia Pacific Supermarket retailing companies’ websites, we plotted the graphs which are shown in Figure 2 and Figure 3. Country-wise study showed China and Korea supermarket websites have linear positive relationship between external inlinks and sales. Japan and Australian supermarket companies’ websites did not exhibit any association between sales and external inlinks. However we need further large scale study to confirm this statement.

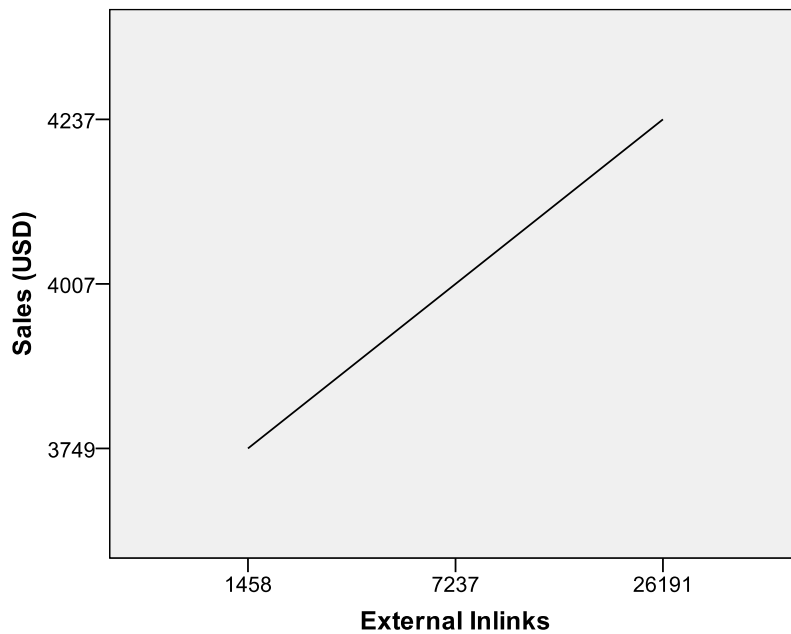


Figure 2: External inlinks Vs total sales (China – supermarket retail company websites)

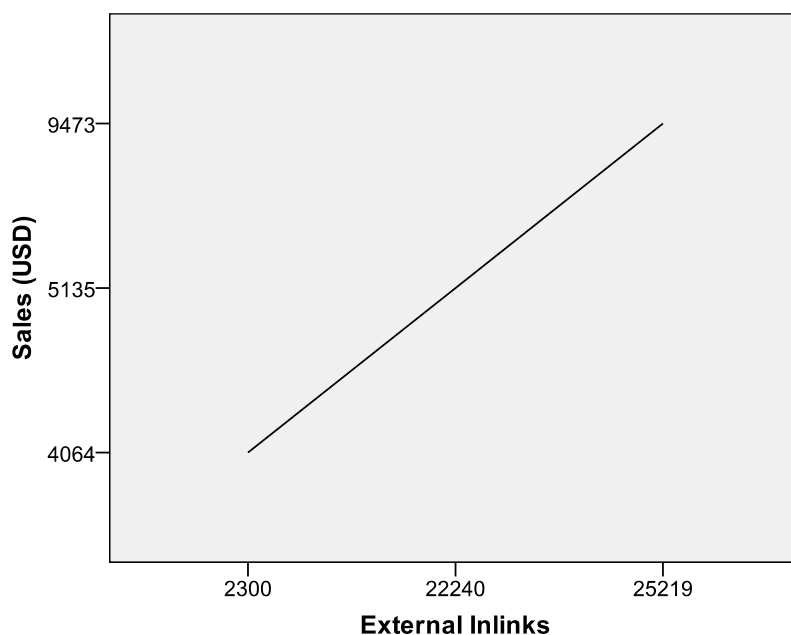


Figure 3: External inlinks Vs total sales (Korea – supermarket retail company websites)

This finding is very important, especially for supermarket companies' because attracting more inlinks, means stronger Web visibility which subsequently increases potential traffic to the website. For a huge percentage of people going to the supermarket to buy the same stuff every week is tiresome. Due to the availability of the technology at a much more affordable price and the comfort of online shopping, they prefer to buy the known products which they need regularly through online. Hence it is crucial for online supermarkets to make its website visible to the buyer. This can be achieved by designing a search engine friendly website, as search engines are the most common information source used by Internet shoppers to find any kind of information on the WWW. A study on the impact of links on website traffic reveals that the number of links to a particular internet retailer stores explains over 60% of the variance of site traffic [5]. Therefore more hyperlinks lead to more traffic to the website [13]. Conversely e-Commerce websites of the department stores and the hardware stores did not show evidence of any significant relationship between the sales and the hyperlink metrics. This result indicates that the association between the inlinks count to a commerce company website and its business performance measure differs according to the nature of the company.

To study whether there is any difference between hyperlink metrics of retail company website and its business performance measure among different countries, we further probed the e-commerce websites country-wise. In the e-commerce websites dataset, there were 35 websites from Japan, 7 from Australia, 5 from China, and 3 from Korea and one each from New Zealand and Taiwan. Except Japan websites, all other country websites are less in number and hence not sufficient to do further analysis. Amongst the total Asia Pacific retailers' websites considered in this study, 74% of the websites are from Japan. Therefore Japan retail companies' websites are considered for the country-wise hyperlink analysis. The result shows that there is no significant relationship between the Japan e-commerce websites and its hyperlink metrics. This again may be due to the diverse nature of the retail companies. Thus our research confirms that there is an association between the hyperlink counts of retail companies' websites and its business performance, thereby answering our research objective.

To study the Web presence of USA and Asia Pacific retailers, WIF of external inlinks (WIF_Te) is used as a metric to measure the impact of the website [6], [8], [10]-[11], [13], [17]. Figure 4 shows that, over all America's top retail companies are more popular thus visible on the Web than Asia Pacific retailers. Based on the category-wise Asia Pacific departmental stores websites are greatly visible on the Web than the rest of its categories. Also the figure shows that Web presence of food and grocery companies of America and supermarket companies of Asia Pacific are nearly equal.

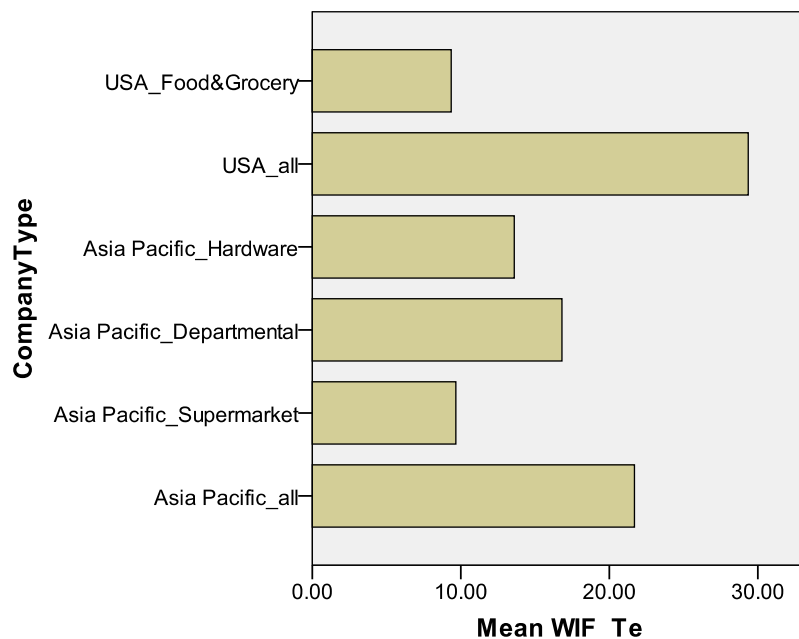


Figure 4: Comparison of web visibility of retail companies of Asia Pacific and USA

Figure 5 exhibits the Web presence of Asia Pacific retail companies according to their country wise. Unsurprisingly Japanese companies have larger visibility compared to other countries that are included in this study. Since there was only one company each from New Zealand and Taiwan in the top 50 list those are not included in the graph. Thus our research confirms that American retail companies' websites have stronger Web presence having more visibility on the web than the Asia Pacific retailers.

We furthered our research to find the relationship between the age of the website and counts of inlinks to a retail company's website. For this study, the age of the retail companies' websites are collected from Internet Archive website. Website age is calculated in months from the date seen by Internet Archive. We calculated the Spearman

correlation coefficient between the website age and the hyperlink metrics for both data sets. The outcome reported some interesting information. For Asia Pacific websites the results showed that there is no association exists between the website age and the total number of links to a retail company website. Conversely American retail websites showed strong positive relation between them. We tested the impact of the website age in both category-wise and country-wise for Asia Pacific retail companies' websites. The supermarket/Hypermarket retail companies' websites revealed a positive significant correlation between the website age and sales ($0.571 p=0.026$) and the total number of pages ($0.525 p=0.044$) at 0.05 level of significance. This is possibly due to the fact that people have trust on the retail companies who exist in the industry for a long time compared to the new players in supermarket / hypermarket industry.

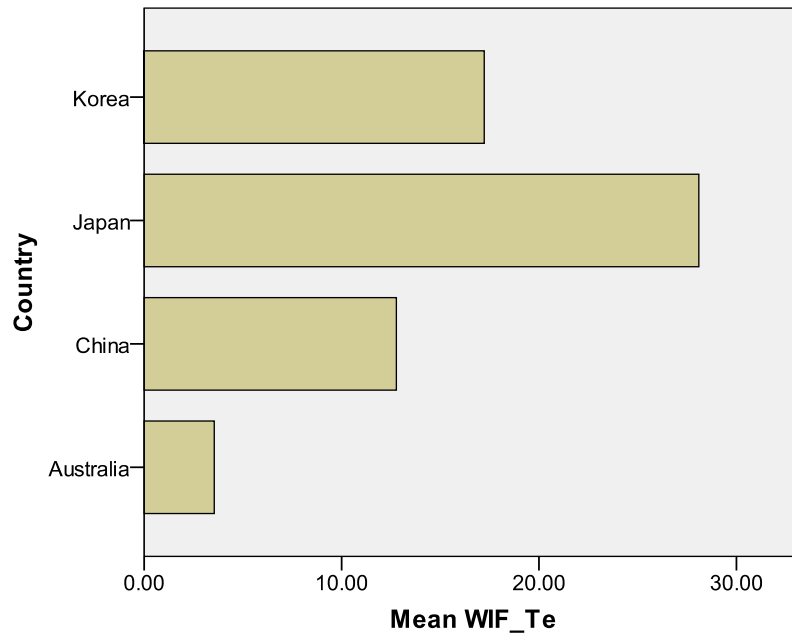


Figure 5: Comparison of web visibility of retail companies among Asia Pacific countries

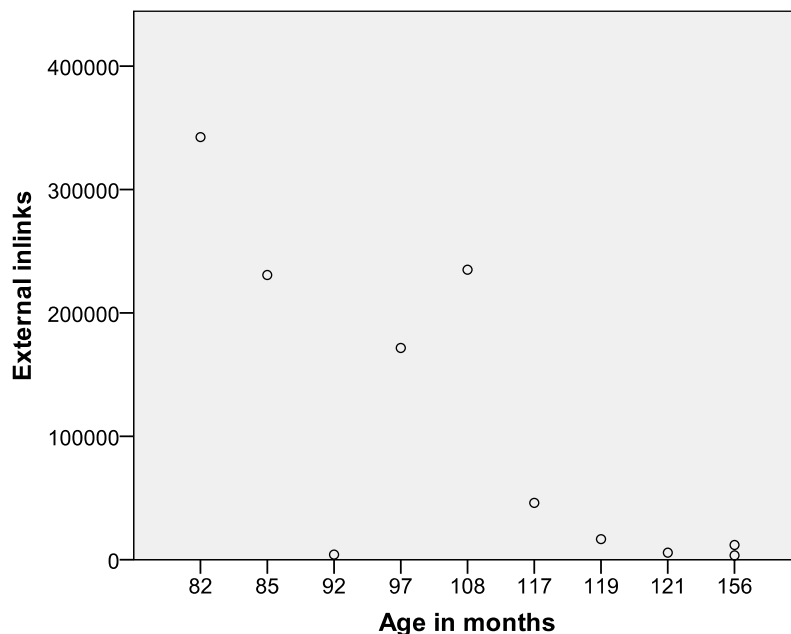


Figure 6: Age of website Vs external inlinks (Asia Pacific hardware retailers)

Unexpectedly, the Spearman correlation test showed evidence of a strong negative correlation coefficients between the website age and sales ($-0.710 p=0.021$) and total external inlinks ($-0.669 p=0.035$) for the Hardware retail companies. Figure 6 illustrates the negative correlation between website age and its external inlinks count. This shows that the most recent hardware companies are doing better sales and attract more links to its websites than the companies who have existed in the field for longer periods. The country-wise study about the relation between the

website age and the hyperlink metrics revealed that Australian retail companies' websites have statistically significant correlation (0.852 $p=0.015$) at 0.05 level of significance. Thus the result indicates that older Australian retail companies receive higher inlinks to its website from other websites. The study found no correlation between website age and hyperlink metrics in countries such as Japan, China and Korea.

The study of relation between the age of a retail company website and number of external inlinks to its website for American retail companies show some valuable information. Figure 7 illustrates that older retail companies' website receive higher external inlinks than the newer. The Spearman correlation coefficient test shows that there is a statistically significant positive correlation (0.442 $p=0.006$ at 0.01 level of significance). This finding was true when tested for each category. For the food and grocery retail companies the Spearman correlation coefficient is 0.582 $p=0.047$ at 0.05 level of significance. From this result, we found that all the retail websites which are in English language have positive relation with the age of the website and its external inlinks.

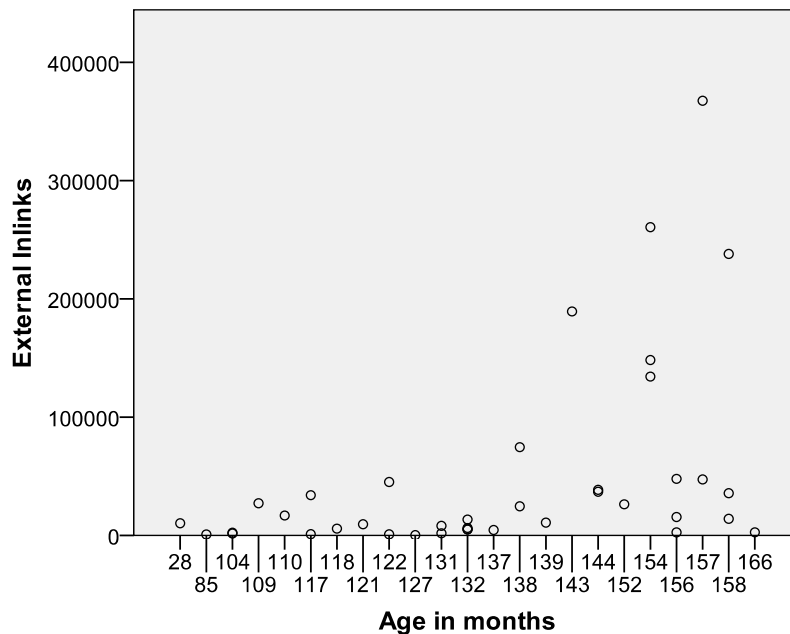


Figure 7: Age of website Vs external inlinks (American retail company websites)

According to the Internet world stats (Site 2) as on 31st December 2009, Asians are the largest Internet users with 42.4% of overall internet users and North America accounts only 14.4%. But the Internet penetration rate by geographic regions statistics shows North America has the highest penetration 76.2% and Asia is only 20.1%. Though many factors affect the Asia pacific companies' performance on the Internet, there are many opportunities on the Web that these companies can explore to attract more potential online customers.

5 Conclusion

This paper tried to answer three research questions in terms of verifying the relationship between the hyperlink counts of retail company Website and its business performance, higher visibility on the web between USA and Asia Pacific retailers and the relationship between the age of the website and the external inlinks to the website. This research considered the hyperlinks counts of top retail company websites of Asia Pacific and USA.

Our study results show that there exists a relationship between the hyperlink counts of Retail Company's website and its business performance. So in a competitive business, if one company attracts more links to its website than its competitors, it is an indication of the popularity of that company among its online shoppers. Therefore we can say that the hyperlink metrics, specifically external inlinks count tends to reflect business information and hence could be used as a measure to estimate the performance of an e-commerce company. Our findings confirm the earlier studies results by Vaughan who found strong relation between the inlinks count pointing to a company's website and its business information [26]-[27], [30]-[31]. However this finding is limited to specific sectors of retail industry websites. We also found that American retail companies' websites have stronger Web presence thus are more visible on the web than the Asia Pacific retailers. Among industry-wise, departmental store companies' websites have greatest visibility. To increase the Web visibility, it is recommended that companies to encourage their employees to have personal website and link to the company's website. For Asia Pacific websites the results showed that there is no association exists between the website age and the total number of links to a retail company website. Conversely American retail websites revealed strong positive relation between them. In addition to this, we noticed that native

English speaking countries' websites have positive correlation with its website age. This study found no correlation in Japanese, Chinese and Korean retailers' website.

This study has some limitations. First of all, the hyperlink data was collected using only one search engine and we understood that no one search engine could index the entire Web. Secondly, the sample size is too small to generalise the results and hence in the future we have planned to include more websites to study the hyperlink structure of commerce companies with many business performance measures. Though there were extensive hyperlink studies in different disciplines with different topics, very few researchers have focused the company's e-commerce websites and studied only IT and telecommunication companies' websites. Our research filled up this gap by studying e-commerce websites both from different countries and categories thus contributed in better understanding of the hyperlink structure of e-commerce websites.

Websites List

Site 1: Know Your Visibility

<http://kyv.webportalquality.com>

Site 2: Internet world stats

www.internetworldstats.com/stats.htm

References

- [1] F. Barjak and M. Thelwall, A statistical analysis of the web presences of European life sciences teams, *Journal of the American Society for Information Science and Technology*, vol. 59, no. 4, pp. 628-643, 2008.
- [2] L. Bjerneborn and P. Ingwersen, Perspectives of webometrics, *Scientometrics*, vol. 50, no. 1, pp. 65-82, 2001.
- [3] L. Bjerneborn and P. Ingwersen, Toward a basic framework for webometrics, *Journal of American Society for Information Science and Technology*, vol. 55, no. 14, pp. 1216-1227, 2004.
- [4] Y. M. Chung, S. Y. Yu, Y. K. Kim, and S. Y. Kim, Characteristics and link structure of a national scholarly web space: The case of South Korea, *Scientometrics*, vol. 80, no. 3, pp. 595-612, 2009.
- [5] C. Ennew, A. Lockett, I. Blackman, and C. P. Holland, Competition in Internet retail markets: The impact of links on web site traffic, *Long Range Planning*, vol. 38, no. 4, pp. 359-372, 2005.
- [6] J. Espadas, C. Calero, and M. Piattini, Web site visibility evaluation, *Journal of American Society for Information Science and Technology*, vol. 59, no. 11, pp. 1727-1742, 2008.
- [7] D. Gibson, J. Kleinberg, and P. Raghavan, Inferring web communities from link topology, in *Proceedings of 9th ACM Conference on Hypertext and Hypermedia*, Pittsburgh, PA, USA, 1998, pp. 225-234.
- [8] S. He, Internet visibility of localized websites of twelve global brands in Japan, Korea and Hong Kong: An evaluation with web impact factor (WIF), in *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing*, Shanghai, 2007, pp. 3799-3802.
- [9] J. Holland and S. M. Baker, Customer participation in creating site brand loyalty, *Journal of Interactive Marketing*, vol. 15, no. 4, pp. 34-35, 2001.
- [10] P. Ingwersen, The calculation of web impact factors, *Journal of Documentation*, vol. 54, no. 2, pp. 236-243, 1998.
- [11] S. K. Jalal, S. C. Biswas, and P. Mukhopadhyay, Web impact factor and link analysis of selected Indian universities, *Annals of Library and Information Studies*, vol. 57, no. 2, pp. 109-121, 2010.
- [12] D. Janetzko, Objectivity, reliability, and validity of search engine count estimates, *International Journal of Internet Science*, vol. 3, no. 1, pp. 7-33, 2008.
- [13] R. Kannan and M. Govindan, Exploring hyperlink structure of electronic commerce websites: A webometric study, *International Journal of Electronic Business*, vol. 8, no. 4-5, pp. 391-404, 2010.
- [14] V. Krebs, Working in the connected world book network, *International Association for Human Resource Information Management Journal*, vol. 4, no. 1, pp. 87-90, 2000.
- [15] J. L. Ortega and I. Aguillo, Difference between web sessions according to the origin of their visits, *Journal of Informetrics*, vol. 4, no. 3, pp. 331-337, 2010.
- [16] J. W. Palmer, J.P. Bailey, and S. Faraj. (2000, March) The role of intermediaries in the development of trust on the WWW: The use and prominence of trusted third parties and privacy statements. *Journal of Computer - Mediated Communication*. [Online]. vol. 5, no. 3. Available: <http://jcmc.indiana.edu/vol5/issue3/palmer.html>.
- [17] H. W. Park and T. Mike. (2003, July) Hyperlink analyses of the World Wide Web: A review, *Journal of Computer-Mediated Communication*. [Online]. vol. 8, no. 4. Available: <http://jcmc.indiana.edu/vol8/issue4/park.html>.
- [18] N. Payne and M. Thelwall, A statistical analysis of UK academic web links, *International Journal of Scientometrics, Informetrics and Bibliometrics*, vol. 8, no. 1, pp. 1-13, 2004.
- [19] C. Serrano-Cinca, Y. Fuertes-Callen, and B. Gutierrez-Nieto, Internet positioning and performance of e-tailers: An empirical analysis, *Electronic Commerce Research and Applications*, vol. 9, no. 3, pp. 237-248, 2010.
- [20] A. G. Smith, Does metadata count? A webometric investigation, in *Proceedings of International Conference on Dublin Core and Metadata for e-Communities*, Florence, Italy, 2002, pp. 133-138.

- [21] M. Spiliopoulou, C. Pohle, and L. C. Faulstich, Improving the effectiveness of a web site with web usage mining, in Proceedings of the International WEBKDD'99 Workshop, San Diego, California, USA, August 15, 1999, pp. 142-162.
- [22] M. Thelwall, Extracting macroscopic information from web links, Journal of American Society for Information Science and Technology, vol. 52, no. 13, pp. 1157-1168, 2001.
- [23] M. Thelwall, Interpreting social science link analysis research: A theoretical framework, Journal of the American Society for Information Science and Technology, vol. 57, no. 1, pp. 60-68, 2006.
- [24] M. Thelwall, Quantitative comparisons of search engine results, Journal of the American Society for Information Science and Technology, vol. 59, no. 11, pp. 1702-1716, 2008.
- [25] M. Thelwall and G. Harries, The connection between the research of a university and counts of links to its web pages: An investigation based upon a classification of the relationships of pages to the research of the host university, Journal of the American Society for Information Science and Technology, vol. 54, no. 7, pp. 594-602, 2003.
- [26] L. Vaughan, Exploring website features for business information, Scientometrics, vol. 61, no. 3, pp. 467-477, 2004.
- [27] L. Vaughan, Web hyperlinks reflect business performance: A study of US and Chinese IT companies, Canadian Journal of Information and Library Science, vol. 28, no. 1, pp. 17-31, 2004.
- [28] L. Vaughan, Y. Gao, and M. Kipp, Why are hyperlinks to business websites created? A content analysis, Scientometrics, vol. 67, no. 2, pp. 291-300, 2006.
- [29] L. Vaughan and K. Hysen, Relationship between links to journal Web sites and impact factors, Aslib Proceedings, vol. 54, no. 6, pp. 356-361, 2002.
- [30] L. Vaughan and G. Wu, Links to commercial websites as a source of business information, Scientometrics, vol. 60, pp. 487-496, 2004.
- [31] L. Vaughan and J. You, Mining web hyperlink data for business information: The case of telecommunications equipment companies, in Proceedings of The 1st International Conference on Signal-Image Technology & Internet - Based Systems, Yaoundé, Cameroon, 2005, pp. 190-195.
- [32] I. Wormell, Informetrics and webometrics for measuring impact, visibility, and connectivity in science, politics, and business, Competitive Intelligence Review, vol. 12, no. 1, pp. 12-23, 2001.