Catalogue Generator Using Distributed Crawler Working Using Python and Redis

Sonia Bajaj [1],*, Shlok Bharuka [2], Danish Sheikh [3], Mechaiver Chherke [4], Rutika Barde [5]

Department of Computer Science and Engineering, G H Raisoni University, Saikheda, Chhindwara, India.

[1]sonia.bajaj@ghru.edu.in, [2]bharuka_shlok.ghrucs@ghru.edu.in, [3]sheikh_danish.ghrucs@ghru.edu.in, [4]chherke_mechaiver.ghrucs@ghru.edu.in, [5]barde_rutika.ghrucs@ghru.edu.in

* Correspondence with: sonia.bajaj@ghru.edu.in

Abstract - A price comparison website acts as a platform or medium between the consumers and the sellers. It allows consumers to see different lists of prices for the product chosen by user and it helps consumers to make an informed decision about which to choose in order to save money. It also act as a tool to help consumers increase their price consciousness so that they will not feel cheated by the advertisement from the retailers that claimed they are offering the cheapest price but the reality happened to be otherwise. On the other words, the price comparison website also has the role to promote the retailer/shop/hypermarket/supermarket to the customers. The pressures on time and money especially in the current economic situation where the living cost increases and there's only little time to do some shopping for household, a price comparison website will absolutely become a great help towards consumers.

Cite this article: Sonia Bajaj, Shlok Bharuka, Danish Sheikh, Mechaiver Chherke, Rutika Barde, "Catalogue Generator Using Distributed Crawler Working Using Python And Redis", *International Journal of Computational and Electronic Aspects in Engineering*, RAME Publishers, vol. 4, issue 3, pp. 58-62, 2023.
https://doi.org/10.26706/ijceae.4.3.20230901

1. Introduction

This website can be very useful for price comparison between products we purchase from e-commerce website. The large and growing industry of price comparison websites (PCWs) or 'web aggregators' is poised to benefit consumers by increasing competitive pricing pressure on -firms by acquainting shoppers with more prices. This website will help consumers in making decision to choose products that will save their money through online. Considering the customers' busy lifestyle especially those who are living in the city area, most of the consumers prefer to buy their needs through the internet because it save their time. Besides, consumers always go for the cheaper price in purchasing products therefore by using price comparison website, customers don't have to travel from shop to shop only to survey the price offered by different shops for the same product. They can just check it from the price comparison website itself and decide where they should buy the products they need [1,2].

Even though not all consumers are buying online, but it is one of the ways to help consumers increase their price awareness. Consumers have the right to know whether the price they are seeing in the shops are good deals as it is claimed or not. Thus it is an advantage for the consumers who are always aware about the current price of a certain product so that they are not cheated by the big words advertised by the shops.

2. Literature Survey

Moraga-Gonzalez J.L and Wildenbeest M.R in July, 2011 focused on price comparison sites and its connotation towards market efficiency and price competition. The price comparison sites attract all the involved parties no matter suppliers or the consumers to its platform as it has become the aggregator of product information. Nowadays people usually conduct first research on the internet regarding a particular product or service that they want which is very convenient as compared to what consumers have to do previously when there is no internet. Back then, consumers have to go through books, magazines, newspaper, encyclopedia and other medium to find needed information meanwhile yellow pages, directories, advertisements and others are useful to locate business and their products. Everything has to be done manually and requires a lot of time, energy and resources to find information without the help of internet. Comparison sites are known as the second type of search technologies as it helps users find products and services according to users' have chosen. It helps consumers to easily compare between the alternatives available in the market and reduce consumer search costs at the same time. Compared to the price comparison site from its first being established and act as online classifieds, nowadays comparison sites able to do a lot of things such as draw the products' cost, able to trawl sites and give the sites much greater control over the users with the help of current technology. It gathers and aggregate price, products and other relevant information Fromm third-party sellers and presents it to the consumers with its own way. There are several business models usually used by the price comparison website and the most widely implemented is the charge will be on the sellers if they want to be listed and users can access the sites for free. The fee usually based on a cost-per-click and is paid every time a consumer is referred to the seller's website from the comparison site. Sellers also have the alternative to pay fee if a consumer buys the product which is being implemented by Pricefight.com. Other than that is providing free services for both sellers and buyers and obtain revenue from the advertisers like what being practiced by Google Product Search. The less chosen business model is the sellers are listed for free meanwhile membership fee is charged towards the consumers such as AngeList.com. Whatever model is being used, the most important thing is the agreement from the both sides on the chosen model. Based on the business models explanation, the most suitable business model for PriceWar.com is the one that providing free services for both users and retailers and the revenues will be obtained from the advertisers who wants to advertise with PriceWar.com. The free services will attract more people to use PriceWar.com because it is free and when more people is using PriceWar.com, the online traffic will be increased as well. The high traffic will attract advertisers to advertise with PriceWar.com. By choosing the intermediation fees carefully to out-compete the search market, a price comparison site will be able to stay in business for a long run. Based on studies made by Yavas (1994), the intermediary such as comparison site will be able to obtain profit by attracting high valuation sellers and low valuation buyers. Comparison sites have the potential in providing abundant of data that might be useful for further analysis for example, analyze how exactly consumers search for products and provide indications about the best way to improve the current system [1-4].

The Use of Price Comparison Sites The research entitled 'The Use of Price Comparison Sites in the UK General Insurance Market' which written by Emily Knight, a strategist for Consumer Intelligence has reported the current performance, media coverage, usage and marketing activity of price comparison sites in the UK General Insurance sector. The results shows that there is increased on the advertising spend and competition and it gives adverse effect on the financial performance of the price comparison sites. Meanwhile the number of consumers using price comparison sites for quotes has remains high and its average number of sites used are increasing over time. Based on the research, 8 out of 10 people are likely to get quotes from the price comparison sites in the future which show good sign of price comparison sites to stay in business in future. To compete with other brands through the internet, big brands usually have to face big advertising spends. Based on Nielsen advertising data, back in 2006 it requires £35m has increased to £85m in 2009. The amount of spend needed to compete with the big player with the big name has make it tough for other competitors unless they have a good marketing budget. Some of the comparison sites have launched various campaigns to promote their sites and get more traffic for example 'Compare the Market' that creates a personality designed to appeal to consumers and improve their impression towards the brand at the same time increase its web traffic and boost conversion rates [4-6].

Consumers Plan More Web Research Before Buying This article was written by Reuther T. – Senior Editor of internetretailer.com, a portal related to e commerce. It is based on the findings of a survey made by Deloitte LLP, concludes that a fifth of online consumers plan to conduct more web research this year (2012) as compared to year 2011 before buying. Delaoitte LLP, a consulting, auditing and financial services firm, has conducted the online survey on 5-

12 July 2012 towards 1, 314 parents of children in kindergarten through the 12th grade. Around 20% of the survey respondents plan to shop online this year but the web is playing the important role in giving influence for the purchases either it is done online or offline (going to the shop to purchase). The result shows that a third of the parents who responded plan to visit e-commerce sites, retail blogs and other web locations to learn before buying the items. Not only that, the survey also resulted that 57% of the respondents owned smart phones and use it as shopping tools for seeking price information with 63%, getting coupons and discounts with 45%, and finding store locations with 38% were among the most popular mobile shopping activities. From this result, it shows that there is big number of people who are using smart phone to check on the price information of a certain product. Therefore, based on this article, it can be conclude that most of the customers, as well as potential customers of a product will make some research online before making a purchase [7,8].

The large and growing industry of price comparison websites (PCWs) or 'web aggregators' is poised to benefit consumers by increasing competitive pricing pressure on firms by acquainting shoppers with more prices. However, these sites also charge firms for sales, which feeds back to raise prices. I investigate the impact of introducing PCWs to a market for a homogeneous good. I find that introducing a single PCW increases prices for all consumers, both shoppers and non-shoppers. More generally, in the most profitable equilibrium for competing PCWs, prices tend to rise with the number of PCWs. (JEL: L11, L86, D43) [8]..

Over the past two decades a new industry of price comparison websites (PCWs) or 'web aggregators' has emerged. The industry has enabled consumers to check prices of many firms selling a particular service or product simultaneously in one place. The sites are popular in many countries, and in many markets including utilities, financial services, hotels, flights and durable goods.1 These sites command billions of dollars of revenue annually.2 In the UK, PCWs for utilities and financial services have been particularly successful. There are roughly 48 such PCWs in the UK, where over 70% of internet users have used such a site. The largest four aggregators generated approximately £800m ($1.2bn) in revenue during 2013, with average annual profit of the group increasing by 14% that year.3 The internet has altered search costs, allowing consumers to compare prices across firms in a matter of clicks, intensifying competitive pricing pressure between firms. While a consumer may not know of all the firms in a market, a PCW can expose the full list of market offerings, maximizing inter-firm pressure. However, underlying this increased competition are the fees paid by firms who sell their products through the websites. In the UK these are understood to be between £44-60 ($69-95) for a customer switching gas and electricity provider.4 These fees, in turn, act as a marginal cost faced by producers, affecting their pricing decisions [9-10].

Two ShopBots were used to determine high-to-low price dispersion for identical models of 25 consumer durables, in 2007 and again in 2011, revealing substantial but declining price dispersion ratios. A survey of 1,135 American online shoppers revealed their dependence on ShopBots and frequency of other online shopping actions. Typical respondent reported they "very often" used search sites to locate what they wanted. Nearly 30 percent used the most often named price comparison site, Yahoo! Shopping, in the past year, suggesting substantial potential for future price rationalization. Several customer relationship management tools online merchants might use to avoid the resulting direct price competition are discussed. Finally, the impact of mcommerce, tablets, and apps on online price comparison behavior is explored. annual online shopping in North America reaches the $200 Billion mark (Lumry, 2011), the emergence of ShopBots such as Yahoo! Shopping and Google Product Search has facilitated consumers' ability to gather price and product information, virtually mitigating search costs. It was therefore initially assumed that as ShopBots became increasingly prevalent on-line merchants would come under increased price competition, eventually leading to price convergence. With ShopBots now the third most popular online shopping option after eBay and Amazon, we reassess the impact of ShopBots on price as they reach their tipping point. Finally, we further expand the discussion to explore the impact of m-commerce, tablets, and apps on online price comparison behavior [11].

Consumers increasingly rely on Internet price comparison sites (PCS) to gain knowledge about the market. The prices generated by a PCS search can act as contextual reference prices and influence the attractiveness of prices encountered later as consumers shop offline at local stores. This paper demonstrates that both PCS retailer ratings and the shape of the PCS price distribution influence the impact of PCS search results on later price evaluations. A favorable PCS retailer rating increases the perceived validity of the price associated with that retailer, enhancing the impact of that PCS price on offline price evaluations (Study 1).The shape of the PCS price distribution can also influence later price evaluations, however this effect depends on the information provided by the PCS retailer ratings. When PCS

retailer ratings are similar, implying similar validity for the associated prices, low PCS prices and those appearing more frequently in the PCS price distribution have more impact (Studies 2 and 3). When PCS retailer ratings are variable (some high and some low), the PCS price distribution effect occurs only when the PCS retailer ratings provide congruent information about price validity — that is, the most frequent price is offered by retailers with more favorable ratings. Study 3 shows that price validity inferences do mediate this result. Finally, we depart from the offline shopping context to show that when consumers choose a retailer directly from the PCS search results, the effect of PCS retailer ratings is stronger for high-priced retailers and for consumers who rely less on the retailer price as a heuristic to infer retailer service level. Based on our findings we offer insights for online and offline retailers when considering strategic responses, such as price matching guarantees.

3. Research Methodology

Stage 1: Project Planning

This is the stage where the problem related to the project is identified and the significance of the study is determined. The objective and also scope of study are outlined and the feasibility of the project work is ensured to be developed within the time frame given. The solution to the problem statement is studied and the types of system to be develop and tools used for developing the system are also identified through literature reviews.

Stage 2: Data Gathering and Analysis

Series of studies has been performed to gain further knowledge on the energy and electricity consumption. Also, readings was done to get better understanding on what comparison site is all about, how it helps people to solve problem before buying home groceries product, and example of existing comparison sites to check on the competitors.

Stage 3: Research on any Existing Similar Systems

Next is the study performed to check if any similar system exists. The main objective of doing research on similar existing system is to know how it works, what concept is being applied in the system, what is being computed by the system and how the system helps solving the problem.

Stage 4: Drafting the Main Components of the System

After the research on similar existing systems, the next step needed is to identify what will be the main component that made up the system to be developed. In order for the visitors of this website to use it services, they have to register with their basic information such as name, email, etc. The registered users will be subscribed automatically to this website's newsletter.

Stage 5: Develop System Architecture

The next phase to develop the architecture on how the system will works. This will give the clear picture and understanding on how the system will operate.

4. Discussion

Using this crawler we intend to make the process of data collection easier. In ways like needing to change only the yaml file and give input in redis and not the code itself. We have coded the crawler in a generic way so that new crawlers are not required for each website. Using this we have made the process to extract data simpler.

5. Conclusion

This document comprises all details related to the project. The abstract, background and objectives can be considered as the overview of the project meanwhile the methodology part explains in details what is the suitable method chosen for this project to be completed. It is also important to get reference from the previous research paper that is related to the project. The literature review shows the research that has been done previously by other researchers. It is very important to get more knowledge in this topic in order to develop a successful and well functional price comparison website. More research needs to be done in order to get the desired results because different site uses its own formula to determine how, when and from where it gathers prices. Therefore, it is very important to get the correct way to retrieve the desired data according to what the user wants. The website provides users with useful information that will help them making

informed decision. With this price comparison website, it solves the problems of the working people to check on the price before buying products. It provides platform for sellers to promote new products and advertise any promotion or sales going on.

References

[1] Moraga-Gonzalez, Jose Luis and Wildenbeest, Matthijs R., Comparison Sites (July 1, 2011). IESE Business School Working Paper No. 933, http://dx.doi.org/10.2139/ssrn.1947292

[2] Berger, P. D., & Nasr, N. I. (2019). Customer lifetime value: Marketing models and applications. Journal of Interactive Marketing. https://doi.org/10.1002/(SICI)1520-6653(199824)12:1%3C17::AID-DIR3%3E3.0.CO;2-K

[3] Brynjolfsson, E., Dick, A., & Smith, M. D. (2021). Search and Product Differentiation at an Internet Shopbot. https://doi.org/10.2139/ssrn.450220

[4] Burdett, K., & Judd, K. L. (2019). Equilibrium Price Dispersion. Econometrica (pre-2019) https://doi.org/0012-9682(198307)51:4%3C955:EPD%3E2.0.CO;2-J

[5] Chevalier, J., & Goolsbee, A. (2021). Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com. Quantitative Marketing and Economics. **DOI** 10.3386/w9085

[6] Choi, D. H., Kim, C. M., Kim, S.-I., & Kim, S. H. (2019). Customer Loyalty and Disloyalty in Internet http://hdl.handle.net/10203/6148

[7] Retail Stores: Its Antecedents and its Effect on Customer Price Sensitivity. International Journal of Management (2021). http://dx.doi.org/10.24912/jm.v24i3.679

[8] Clay, K., Krishnan, R., & Wolff, E. (2018). Prices and price dispersion on the Web: Evidence from the online book industry. http://dx.doi.org/10.1111/1467-6451.00161

[9] Dawkins, P. M., & Reichheld, F. F. (2021). Customer Retention as a Competitive Weapon. Directors and Boards. http://dx.doi.org/10.1108/02634500210428003

[10] Degeratu, A. M., Rangaswamy, A., & Wu, J. (2021). Consumer choice behavior in online and traditional. http://dx.doi.org/10.1016/S0167-8116(00)00005-7

[11] Supermarkets: The effects of brand name, price, and other search attributes. International Journal of Research in Marketing,(2021).http://dx.doi.org/10.1016/S0167-8116(00)00005-7