Development of E-Commerce Product Ranking Using PageRank Technique Combined with Consumer Satisfaction Data

Mongkhol Rattanakhum1 and Pruet Boonma, Ph.D2 and Sumalee Sangamuang, Ph.D3

¹ Master's Degree Program in Data Science, Chiang Mai University, Chiang Mai, Thailand ² Department of Computer Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand

³ College of Arts, Media and Technology, Chiang Mai University, Chiang Mai, Thailand

Mongkhol r@cmu.ac.th

Abstract. This independent research aims to rank electronic commerce (E-commerce) commercial products by utilizing data from E-commerce platforms for study and investigation. It incorporates a tool known as "Page Rank" and combine with consumer satisfaction to reorganize the product ranking. The goal is to promote new or less-recognized products to improve their visibility. This point is the problem commonly faced by businesses that want to increase sales and maximize product distribution but struggle to showcase products that have not yet garnered significant sales, causing a vicious cycle of low visibility. In this research, have implement products same store as variables in calculating Page Rank scores. The testing results show that by combining Page Rank scores with consumer satisfaction ratings, products that were previously displayed in the middle or lower rankings on the E-commerce platform can be repositioned to appear at the top or on the first page more frequently and result from user acceptance test is not variant.

Keywords: Web Crawler, Page Rank, E-Commerce, Product search, Scoring and Ranking

1 Introduction

In the era of e-commerce, the trend of buying products online has seen significant growth, primarily due to increased internet accessibility and smartphone usage. According to data from E-Conomy SEA 2022, the Southeast Asian region has witnessed approximately \$131 billion in e-commerce sales. In Thailand, e-commerce sales and services have reached about 22 billion USD, accounting for roughly 17% of total sales [1]. The Electronic Transactions Development Agency (ETDA) found a continuous growth in Thailand's e-commerce sector, with a growth rate of approximately 6.11% between 2020 and 2021 [2]. This robust growth has prompted many sellers to embrace

e-commerce platforms, leading to intense competition among major players like Lazada and Shopee.

The e-commerce landscape in Thailand offers a wide variety of formats and sales channels to attract and engage consumers. One prominent channel is through e-commerce platforms such as Lazada and Shopee. According to news reports, in 2021, Lazada held an 11.8% market share, while Shopee had a 5.7% market share [3]. These platforms strive to entice consumers through promotions, discounts, and incentives. Furthermore, they attract sellers, both local and international, to expand their product offerings. These competitive strategies have led to a diverse selection of products, with each platform offering no less than 100,000 items from various sellers [3]. With consumers relying heavily on reviews and comparisons, businesses need to find ways to showcase their products effectively.

Product ranking on e-commerce platforms typically relies on a combination of factors, including consumer reviews and purchase frequency, to create a reliable ranking system. Advertising and special business privileges often play a role in determining the display order, with some products appearing prominently through paid promotions. However, these rankings occasionally lead to visibility issues for new products, which may not have received enough reviews to compete effectively. Consequently, businesses must invest in marketing efforts to increase their product visibility and achieve higher rankings.

The purpose of this research is to develop a product ranking system for e-commerce platforms. It will involve the collection of product data from Lazada through web crawling, including detailed product information, user ratings, and store details. The Page Rank algorithm will also be employed to assist in ranking. Additionally, customer satisfaction will be factored into the ranking system to ensure a balanced and effective product ranking approach, which will be displayed on a website for the benefit of both businesses and consumers.

2 Literature Review

Ranking and scoring in E-commerce has a been using and researching in many years. To understand history and methodology have study theoretical frameworks and relevant tools are explored to align the research with the specified requirements and provide guidance to researchers. The study involves examining the data collection process, as well as performing calculations using ranking methods through the utilization of the PageRank algorithm. Additionally, consumer satisfaction scores are considered as part of the evaluation process.

A Framework for Ranking Products Using Ranked Voting Method. [7] This research presents a framework for ranking products based on average review scores to determine product efficiency. It utilizes data from e-commerce websites and applies the TF-IDF (Term Frequency - Inverse Document Frequency) transformation to review data for

ranking products using the DEA (Data Envelopment Analysis) process. The results show that the ranking of products remains stable, as it calculates scores for efficient products. Changes in the efficiency of products do not significantly affect the ranking.

A Recommender System for E-Commerce Using Multi-Objective Ranked Bandits Algorithm. [6] The research focuses on developing a recommender system for e-commerce using the Multi-Objective Ranked Bandits Algorithm. The system takes user interactions with product links and computes rankings based on four components: Scalarization functions, Recommendation Quality Metrics, Weighting Scheme, and Multi-Arm Bandits Algorithm. The study suggests that the developed system can be used to analyze user behavior and market trends in e-commerce.

Best Seller Rank (BSR) to Sales: An Empirical Look at Amazon.com. [8] This research examines the Best Sellers Rank (BSR) on Amazon, a key variable for displaying product rankings on the platform. The study collects data using web scraping and applies machine learning techniques, particularly regression analysis, to predict sales. The findings indicate that BSR changes when products are sold within three days, and these changes can be predicted based on sales frequency and review details.

Evaluation of Iterative PageRank Algorithm for Web Page Ranking. [9] The research evaluates the efficiency of various PageRank algorithms, including Jacobi, Gauss Seidel, and Induced Dimension Reduction (IDR), compared to the original Power Method for ranking web pages. The results show that Jacobi and IDR have faster computation times than Power Method and Gauss Seidel, respectively.

Product Rank Based Search Engine for E-Commerce. [10] This study addresses the challenges of comparing product prices and discounts across different e-commerce websites. It utilizes web crawlers to gather real-time search data and uses PageRank to rank products. The system helps users make informed decisions about purchasing products by providing real-time price and discount information.

Ranking Web Search Results from a Personalized Perspective. [11] The research focuses on personalizing web search results using the Personalized PageRank (PPR) algorithm. It calculates web page relevance weights based on user clicks and interest grouping, aiming to provide users with more relevant search results.

A Crowdsourcing and Gamification-Based Product Ranking Method for E-Commerce. [12] This research incorporates gamification and crowdsourcing to improve product ranking. Data is collected from the Digikala e-commerce website, and a new ranking system is proposed. The results show that customers tend to trust rankings based on review scores more than computer-generated recommendations.

3 Data and Methodology

The method of "Developing an E-Commerce Product Ranking System Using PageRank and Consumer Satisfaction" involved planning and preparing research activities in a sequential manner as follows.

3.1 Design of Research Tools

Based on the research findings and review, a plan for designing the research tools for ranking electronic commercial products using the PageRank technique combined with consumer satisfaction was developed. This plan encompasses the entire workflow, from data collection to ranking results. It includes the following details.



Fig 1. Flowchart system design

1) Design of Data Collection Structure

Data collection for research was studied, considering different possible sources of data. The research utilized a data collection method involving the search for commercial electronic products on the e-commerce platform called Lazada. The search query "monitor" was chosen as the product is often competitively priced. The data was collected using a web crawler tool to automate the process of gathering content from web pages. Selenium was selected as the tool to access the website due to its compatibility with Lazada's web crawling restrictions and its frequent anti-scraping measures. Data collection was structured into two parts: product details from search results and products from the same store.



Fig 2. Example raw data from Lazada

2) Design of Data Storage Structure

To make collected data usable for subsequent analysis and display, a relational database structure was designed to store product data. PostgreSQL was chosen as the database management system to store product details, including product ID, name, URL, pricing information, ratings, store information, and shipping details.

3) Design of PageRank Calculation Structure

The structure for PageRank calculation was designed based on the collected data, utilizing the relationship between products from the same store and product brands to calculate PageRank scores. Data was processed to establish the weightings required for graph relationships. The calculated PageRank results were stored in a relational database for efficient retrieval and application in the next steps.

4) Design of the Display System Structure

A system for displaying product rankings was designed to visualize the results of the research. This system needed to support ranking based on PageRank scores and consumer satisfaction. Django, a popular web framework in Python, was chosen to develop this system. It integrates with the database, processes the data, and displays the rankings of electronic commercial products.

5) Design of Product Ranking Structure

A structure for displaying product rankings was designed to accommodate ranking from PageRank, Lazada, and price. The ranking system was designed to support different sorting options for comparison, and it involved PageRank, Lazada ranking, and price sorting.

3.2 Tool Development for Research

The development of tools for testing and research includes the creation of tools for various tasks. It also involves the development of the system structure for ranking electronic commerce products using PageRank in conjunction with consumer satisfaction. This encompasses the entire process from data collection to the presentation of rankings.

1) Data Collection System Development

For the development of the data collection system, a tool called Selenium was used to collect product data from Lazada. During the data collection process, additional tools such as Chrome Driver were employed to enable Selenium to interact with the website. Data extraction was performed by specifying positions on the webpage using class names and XPath found in the HTML. When the system identifies data at those locations, it stores the information in variables for subsequent accumulation and storage in the database.



Fig 3. Example of Using XPath and Class Name Commands

The system's working process is divided into two parts: collecting product details and collecting products from the same store. When collecting store data, commands are executed to store data according to the database design. To gather data about products from the same store, an ID for referencing products from the same store that were collected is added to the product details.



Fig 4. Program Code for Scraping Storing Product Details

2) Database Development

Once data collection is complete, the data is stored in a database using PostgreSQL as the database management system. This database comprises two tables: the product details table and the products from the same store table. Each row in the database represents a single product.

In the database structure, various columns have been created following the data storage design plan. The data types for input are adjusted to ensure data cleanliness before use, such as converting product prices to numerical values and storing scores as decimals.

ld [PK] bigint	name text	uri text
1	Acer Nitro Gaming LED 27 VG270bmiix (IPS Panel) input Signal VGA,HDMI x 2,SPK, Audio out (จอคอมพิวเตอร์ , จอคอม , Monitor)	https://www.lazar
2	ACER 21.5" MONITOR SA220QBbix, LED 21.5, 16.9, IPS Panel VGA: 1,920 x 1,080 @ 60 Hz, HDMI: 1,920 x 1,080 @ 75 Hz, Panel type : IPS (งอดธมพิวเตยร์)	https://www.lazar
3	HP Monitor M22f 21.5 FHD IPS 75Hz sRGB 99%	https://www.lazar
4	HP V20 HD+ Monitor Warranty 3 Years Onsite By HP	https://www.lazar
5	Acer Monitor 21.5' SA220QBbix (IPS, VGA, HDMI) 75Hz vaavaliumar, Salestore	https://www.lazar
6	[อุปการณ์ครบ หรือมสาย] LG 29 Ultrawide 29WK660-W, 75Hz [HDM, DP] IPS Monitor ขนาด 29 นิ่ว จอคอมพิวเตอร์ จอมอนิเตอร์ (ประเท้น 3 ปี) Computer-SaleStore	https://www.laza
7	🗲 🗲 สินตักNewราคาพิเศษ 🗲 🗲 01เMSI MONITOR (ขอมอนิเตอร์)MSI PRO-MP271P/LCD 27Wide Screen (16:9) IPS /5MS/ 75H2 /16:9/1,000:1/1920 x 1080 (Full HD)/มีล่าโพง ขอวรับชิ้น-ลงได้/3Year	https://www.lazar
8	ACER Gamisg LED 21.5 VG220Qbmilk (IPS Panel VGA HDM) (งอดอมพิวเตอร์ งอดอม Monitor)	https://www.lazar
9	(ฟักแขทรับดูปอง) GIGABYTE MONITOR G24F - 23.8 IPS FHD 165Hz ขอมอนิเตอร์เกมนั้ง (ประกับสูบย์วิบี)	https://www.lazar
10	Gaming Monitor Nitro 21.5 VG220Qbmix 75Hz IPS เครื่อง ไหม่ประกันฐมย์	https://www.lazar
11	(ฟักแขทรับดูปอง) Philips Monitor 19.5 รุ่น 203V5LS82 TN 60Hz FHD รับประกันดุบมิโทย 3 ปี (Onsite)	https://www.lazar
12	จอมอนิเตยร์ SAMSUNG MONITOR LF24T350FHEXXT (IPS 75Hz) by Banana IT	https://www.lazar
13	MONITOR (ขอมอนิเตอร์) 23" DELL (\$2340L+HDMI) B IPS Panel 60Hz ขอตอมพิวเตอร์ 23 นั่ว	https://www.laza
14	Dell P2422H 23.8 Monitor Warranty 3 Years Onsite By Dell	https://www.lazar
15	[484 0% 10 n.]ACER MONITOR C8292CUbmiliprx/Jtsrifu3Y/BY TOP COMPUTER	https://www.lazar
16	ACER LED Monitor 27" KA272bi (IPS Panel) LED 27 Inch IPS 75Hz (ຈອກສາສົງເສຍຕົ້)	https://www.lazar
17	LG GAMING MONITOR 27GL850-B 27Inch 2K (2560 x 1440) 144Hz ประกัมชุมส์ LG 3 ปี	https://www.lazar
18	MONITOR MSI OPTIX 624C4 23.6 CURVED 144Hz (by Pansonics)	https://www.lazar
19	ให้กแสทาวันชุปชิง] Xiaomi Monitor Gaming Curved 34 21:9 144 Hz WQHD 3440*1440	https://www.lazar
20	JIB MONITOR (ขณะมนิมกตว์) LG 29WK600-W 29 IPS 75Hz	https://www.lazar
21	MONITOR (ขอมฉบังสะวั) ACER VG220QBMIX 21.5 IPS 75Hz (PORT AUDIO / VGA / HDMI) By Speed Gaming	https://www.lazar
22	ชาธิตจอดอะพิวเตอร์ Single Computer Monitor Desk Mount Stand	https://www.lazar
23	MONITOR (100405/005) LG 24MK430H-B 23.8 IPS 75Hz	https://www.lazar
24	MONITOR (valuaDurad) MSI MODERN MD241P 23.8 IPS SPEAKERS USB-C 75Hz By Speedcom	https://www.lazar
25	HP E24 G4 FHD Monitor Warranty 3 Years Onsite by HP	https://www.lazar
26	ACER MONITOR Nitro Gaming VG240YSbmilpx (IPS 165Hz) by Banana IT	https://www.lazar
27	โท้กแขทรับอุปอง) Philips Monitor 21.5 รุ่น 223V5LS82 TN 60Hz FHD วับประกับอุบมโทย 3 ปี (Onsite)	https://www.lazar
28	Philips 2G3V5LS82 LCD Monitor 19.5 ขอคณาที่วงคอร์ ของณาที่ ประกับคุณที่ 3ปี	https://www.lazar

Fig 5. Example of Data Stored in the Database

3) PageRank Score Calculation System Development

The PageRank calculation system was developed using tools to compute weighted scores. This involved using data from products from the same store to find web page connections after successfully calculating linkages. The results of these calculations are then stored in the database for use in ranking products in the next step.

	id [PK] bigint	link_id character varying (10) 🖍	store_id integer	brand character varying (50)	pagerank double precision
1	1	E1	61	ACER	0.025676418
2	2	E2	62	ACER	6.02e-05
3	3	E3	63	ACER	6.02e-05
4	4	E4	64	ACER	3.46e-05
5	5	E5	65	ACER	3.46e-05
6	6	E6	66	ACER	3.46e-05
7	7	E7	67	ACER	3.46e-05
8	8	E8	68	ACER	3.46e-05
9	9	E9	69	ACER	3.46e-05
10	10	E10	70	ACER	3.46e-05
11	11	E11	71	HP	0.016815742
12	12	E12	72	HP	6.02e-05
13	13	E13	73	HP	6.02e-05
14	14	E14	74	HP	3.46e-05
15	15	E15	75	HP	3.46e-05
16	16	E16	76	Intel	0.003745485
17	17	E17	77	Dell	0.02564698
18	18	E18	78	Dell	6.02e-05
19	19	E19	79	HP	3.46e+05
20	20	E20	80	Dell	6.02e-05
21	21	E21	81	AOC	0.012105747
22	22	E22	82	Philips	0.002303027
23	23	E23	83	ACER	3.46e-05
24	24	E24	84	LG	0.015608806
25	25	E25	85	Neffos by TP-LINK	0.000625
26	26	E26	86	ACER	3.46e-05
27	27	E27	87	AOC	6.03e-05
28	28	E28	88	ACER	3.46e-05

Fig 6. Example of Stored PageRank Data

4) Data Presentation System Development

The data presentation system was developed using Django, allowing the web pages to support testing and product ranking. The system has a search feature for product comparisons between PageRank and Lazada.



Fig 7. Search System Interface

5) E-Commerce Product Ranking System Development

Ranking based on Page Strength and Satisfaction: The ranking of products will be determined by combining the page strength scores with consumer satisfaction to display results. The primary criteria for ranking will be the page strength scores, and in cases where scores are equal, the subsequent ranking will be determined based on other satisfaction aspects of the products.

 Table 1. Shows the priority of sorting using the PageRank technique combined with consumer satisfaction.

Detail	Score
PageRank score	5
Star	4
Reviewer	3
Pricing	2
Store score	1

4 Results and Discussion

Results and experimental procedures for ranking e-commerce products using a combination of PageRank techniques and consumer satisfaction are discussed as follows.

4.1 Experimental procedure

In the experimental, the testing formats are divided into two types.

- 1) the search query selection process for testing product ranking and the comparison between Lazada's ranking and the ranking using page strength techniques, combined with consumer satisfaction.
- Enter the search term and confirm the search. The system will display ranked results.
- Select one product for testing.
- Examine the ranking order of the selected product when using a ranking system based on page rank combined with consumer satisfaction.
- Examine the ranking order of the selected product when using the ranking system from the Lazada platform.
- Record the ranking results in a table along with the details of the selected products.

Order	Keyword
1	MSI
2	Acer
3	Gaming monitor
4	Dell
5	งอคอม

Table 2.	Searching 1	keyword i	for testing
			(1

Keyword	Product Name	Pag-	Lazada
		erank	
MSI	[ทักแชทรับคูปอง] MSI Monitor 27 รุ่น OPTIX	1	155
	G273QF IPS 2K 165Hz+แถมฟรี Mouse Gaming		
	GM11		
Acer	JIB MONITOR (จอมอนิเตอร์) ACER	18	129
	VG240YBMIIX 23.8 IPS		
Gaming	LG GAMING MONITOR 27GL850-B 27Inch	12	60
monitor	2K (2560 x 1440) 144Hz ประกันสูนย์ LG 3 ปี		
Dell	[ทักแชทรับคูปอง] Dell Monitor Touchscreen	46	157
	23.8" รุ่น P2418HT IPS 60Hz		
งอคอม	[จอกอมพิวเตอร์] Dell E2220H 21.5 Monitor War-	20	65
	ranty 3 Years by Dell		

Table 3	Order	number	Page	rank	compare	with	Lazada
	~				••••••••••		

2) testing involves randomly selecting a sample group of users interested in purchasing computer monitors from electronic commerce platforms. The goal is to inquire about satisfaction levels regarding product rankings on Lazada and to assess the extent of the differences when employing page strength techniques.

In the testing phase with the sample group, the researchers have created a product search website specifically for the computer monitor category. Users were able to experiment with the search functionality, including ranking results based on page rank and Lazada rankings. After the testing phase concluded, the researchers planned to have participants respond to a questionnaire to gather feedback for the subsequent processing and analysis.

No.	Question	Answer	Per-
			cent
1	Do you have interest in to buy PC monitor on E-	Interested	100 %
	commerce?	not inter-	0 %
		ested	
2	Following a experimental	Excellent	13 %
	(https://topthesis.mininggarden.com/)	Good	43.5 %
	Do you have satisfaction for searching result?	Fair	26.1 %
		Poor	17.4 %
		No opinion	0 %
3	From the previous experiment (No. 2) can help	Matched	73.9 %
	you to find product that does not have review in-	Not	26.1 %
	formation but matching with your requirement?	Matched	
4	Following a experimental,	Excellent	8.7 %
	(https://topthesis.mininggarden.com/?sorting=3)	Good	52.2 %
	Do you have satisfaction for searching result?	Fair	34.8 %

Table 4. Questionnaire and result

No.	Question	Answer	Per-
			cent
		Poor	4.3 %
		No opinion	0%
5	From the previous experiment (No. 4) can help	Matched	73.9 %
	you to find product that does not have review in-	Not	26.1 %
	formation but matching with your requirement?	Matched	

5 SUMMARY

As the results of the experiment on ranking products in electronic commerce using the page strength technique combined with consumer satisfaction, it was found that there is a significant difference in product rankings compared to rankings through the Lazada platform. This difference arises from the use of different display factors. The results of product rankings showed a higher increase rate when employing the page strength technique along with consumer satisfaction, with an average increase rate of 70%, or at least 70 positions out of 100 rankings.

For the test results from the user sample group, it was observed that using page strength scores for ranking products did not differ much from the results obtained from Lazada. The findings can be summarized as follows.

- The survey participants, with an age range of 25-30 years and an interest in purchasing computer monitors, consisted of a total of 23 individuals.
- The satisfaction results from using the page strength technique combined with consumer satisfaction and Lazada showed that satisfaction levels were generally higher when using the page strength technique for ranking products.
- In the test results, it was found that the search outcomes for products without scores using the page strength technique combined with consumer satisfaction and Lazada were consistent with each other, meeting the desired criteria without significant differences.

6 **DISCUSSION**

In the process of ranking, variables were used to calculate the PageRank scores based on products within the same store. If there are changes in the variables used for a new ranking, such as review scores and review results, adjustments in the scoring and ranking process would be necessary to ensure accuracy and relevance.

References

- 1. e-Conomy SEA, "Google e-conomy SEA," 2022. [Online]. Available: https://economysea.withgoogle.com/.
- สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์, "สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์," 2021. [Online]. Available: https://www.etda.or.th/th/pr-news/ETDA-Reveals-the-Value-of-e-Commerce-in-2021.aspx.
- อมรินทร์ เทเลวิชั่น, "Amarintv," [Online]. Available: https://www.amarintv.com/spotlight/insight/detail/24463.
- 4. PwC, "PwC Global," 2023. [Online]. Available: https://www.pwc.com/gx/en/industries/consumer-markets/consumer-insights-survey.html.
- 5. e. .ee, "Marketeer Online," 2021. [Online]. Available: https://marketeeronline.co/archives/207221.
- A. Hidayatullah and M. A. Anugerah, "A Recommender System for E-Commerce Using Multi-Objective Ranked Badits Algorithm," in Second International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, India, 2018.
- R. Kumar, A. Sharan and P. Biswas, "A framework for ranking products using ranked voting method," in International Conference on Computing, Engineering, and Design (ICCED), Bangkok, Thailand, 2016.
- A. Sharma, H. Liu and H. Liu, "Best Seller Rank (BSR) to Sales: An empirical look at Amazon.com," in IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C), Macau, China, 2020.
- F. U. Zambuk, A. Y. u. Gital, S. Boukary, F. Jauro and H. Chiroma, "Evaluation of Iterative Pagerank Algorithm for Web Page Ranking," in 4th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT), Mysuru, India, 2019.
- U. Rathod, P. Pavate and V. Patil, "Product Rank Based Search Engine for E-Commerce," in 3rd International Conference for Convergence in Technology (I2CT), Pune, India, 2018.
- P. Wen-Chih and L. Yu-Chin, "Ranking Web Search Results from Personalized Perspective," in The 8th IEEE International Conference on E-Commerce Technology and The 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services (CEC/EEE'06), San Francisco, CA, 2006.
- M. Hajarian and S. Hemmati, "A Crowdsourcing and gamification based Product Ranking Method for E-Commerce," in 7th International Conference on Web Research (ICWR), Tehran, Iran, 2021.