Assessing the E-commerce Websites for Performance using Automated Testing Tools

A.R.F. Shafana[1*], A.F. Musfira[2] & M.M.F. Naja[3]

[1]*Department of Information and Communication Technologies, Asian Institute of Technology, Thailand.*
[2]*School of Multimedia Technology and Communication, Universiti Utara Malaysia, Malaysia*
[3]*Department of Software Engineering, Universiti Malaya, Malaysia*
[1,2,3]*Department of Information and Communication Technology, South Eastern University of Sri Lanka, Sri Lanka*

[1*]arfshafana@seu.ac.lk, [2]ameermusfi@seu.ac.lk, [3]mmfnaja@seu.ac.lk

Abstract- *The number of users accessing an e-commerce website is generally high. The performance of websites to cater to the increased number of concurrent users is inevitable to manage the growing online business needs. The architecture of the company's website should be robust enough to manage the expected traffic on heavy load. The inability to support the growing customer needs would return frustration that leads to heavy business loss. Therefore, it is mandate for e-commerce websites to perform load testing to assess the robustness of their architecture to support scalability. This study features the assessment of e-commerce websites for its performance based on throughput, availability and response time. The study utilizes Apache JMeter to perform load testing on selected five e-commerce websites in Thailand by emulating customer behaviors at heavy load levels. The paper proposes a methodology that could help future testing practitioners and researchers to perform load testing efficiently.*

Keywords: Load Testing; Quality of Service; E-Commerce; Apache JMeter; Performance Metrics.

I. INTRODUCTION

E-Commerce is a platform that serves a multitude of online users concurrently and the usage increases even higher during extreme shopping periods (Janani and Krishnamoorthy, 2015). The outbreak of COVID19 has increased the momentum of people purchasing online significantly, giving rise to the number of concurrent online users accessing e-commerce websites. The Quality of Service (QoS) of such a large-scale system is typically measured in response time, throughput, and availability (Menascé and Mason, 2002). Thus, e-commerce web applications fail when they could not scale up to meet the growing concurrent accesses leading to lost business opportunities (Jiang and Hassan, 2015). Therefore, the infrastructure of such websites needs to be assessed timely and upgraded to obtain the required return on investment (ROI) for the growing business needs.

Load testing becomes crucial at places where hundreds of millions of users access a website simultaneously. This study applies load testing on popular e-commerce websites in Thailand to assess their quality of service in terms of concurrency. Measuring the performance of e-commerce websites has been a subjective topic since the past, and many scholars have suggested the use of many metrics (Ghandour, Benwell and Deans, 2010). However, this study is one-dimensional that aims to investigate the capability of such websites for growing customers alone and does not discuss the usability, accessibility, security, or other measures relevant to the target websites.

The primary objective of this study is to analyze the performance issues of e-commerce websites under heavy load. Five popular e-commerce websites namely Lazada, Shopee, JD Central, Power Buy, and JIB have been chosen to investigate their performance towards increased users. Apache JMeter has been used to simulate a test environment during a peak shopping period with thousands of clients sending requests to each of the websites. The tests on each website have been conducted simultaneously using the same machine. The websites have been assessed in terms of their throughput, response time, and availability.

Test results show that the suitable metrics to evaluate the performance of the website under heavy load are throughput, availability, and response time. The evaluation of websites across such metrics could help the websites rectify the issues encountered when many users access the same URL concurrently. Thus, load testing is an important phase during the development of large-

scale systems and testers should focus on load testing besides the testing carried out for functionalities and features.

The contribution of this paper is two-fold. First, the methodology presented here would be useful for load testing practitioners and researchers interested in load testing. Secondly, this study highlights the importance of load testing for an e-commerce application. It promotes the developers of such applications to include load testing and the conventional functional testing procedures. Although the case study is undertaken for e-commerce websites in Thailand, the findings are not subjective. This could be integrated to implement different business policies and tactics in any large-scale systems where concurrency is inevitable.

II. RELATED WORKS AND EXISTING LITERATURE

A. Assessment of e-commerce websites

The extant literature forwards several measures as Web Performance Metrics. However, the measure of performance for e-commerce websites has been one of the controversies since the past (Ghandour, Benwell and Deans, 2010). After an extensive review, the researcher suggests using termed usage, financial returns, and user satisfaction as key measures. Hamid, Bawany, and Zahoor (2020) assessed the usability and accessibility of e-commerce websites from Pakistan. The researchers emphasize that the level of satisfaction towards an e-commerce application would be vital to measure the success of such websites. The study also included design suggestions to improve the usability and accessibility of e-commerce sites to increase customers. Stefani and Xenos (2009) developed a framework of performance metrics to evaluate e-commerce websites specifically. They stated that the quality is subjective based on the stakeholder and mapped their identified metrics to the quality characteristics such as Functionality, Usability, Efficiency, and Reliability. The term efficiency in their work defined the capability of the web application to provide appropriate performance respective to the amounts of used resources under specified conditions. Zhao et al., (2019) conducted a behavior analysis of the system and the user of e-commerce trading systems. They analyzed the e-commerce systems using four types of network user behavior analysis methods such as the web browsing behavior of users, keystroke behavior of network users, network transaction behavior of users, and mobile terminal behavior. They concluded that this kind of analysis is important for building assessment standards. Usability is another important metric that e-commerce has been assessed upon. Scholarly works have proved that the usability and quality of service of e-commerce websites can have significant effects on the performance and satisfaction of online users (Wahyuningrum, Kartiko and Wardhana, 2020; Wijaya et al., 2021).

B. Load Testing

Menascé and Mason (2002) emphasized the importance of load testing in e-commerce applications. He affirms that the important metric in large-scale systems like e-commerce web applications is its performance. Further, he imposes that the load testing needs to be carried out at several business requirement levels. Jiang et al. (2009) put forwards that many of the field problems in relation with the web applications are the deficiency of systems to cater to the increased load. They further mention that load testing could uncover both functional and performance problems under load, and when the results are systematically analyzed, this can be used to uncover any underlying problems. Jha and Popli (2017) identified that the features and functionalities are not only the main concern of a website and describe the vital necessity of the systems' performance to work properly under its expected workload. They further mention that the application of load testing can eliminate performance bottlenecks greatly. Another work (Khan and Amjad, 2016) also describes the importance of load testing for website applications. The researchers suggested that metrics such as the end-user response time, response time of CPU, and memory statistics gathered from the load testing are inevitable for the critical analysis of application behavior. Yin et al. (2021) conducted a mass tourism data analysis API based on e-commerce platform. The system was tested for its performance using automated testing tools. The tool was preliminarily used to test the response speed of the system, concurrency, and stability index. Although the system adhered to the design principles, the test found out that the system was not able to process API requests of more than 4200 failing in stress testing of the system. Another work much closer to the methodology adopted in this study is from Musthafawi et al. (2020). Using the similar methodology utilized in this study, their research was conducted to explore the enthusiasm of online users and how it is affected by the COVID19

pandemic. The objective of the study was to compare the performance of an e-commerce website before and during the pandemic. To achieve this objective, the researchers used Apache JMeter in line with the questionnaire on enthusiasm assessment. The researchers claimed that this load testing is essential to identify the system deficiencies.

III. METHODOLOGY

A. Tools and Technologies

Apache JMeter is a desktop-based Java application developed by Stefano Mazzocchi of the Apache Software Foundation (Dhiman and Sharma, 2016). The software can be utilized for various performance testing. The load testing of HTTP requests of Apache JMeter has been utilized for this specific study. The application also allows multiple listeners to view the results. This study mainly uses the Graph Results to observe the deviation, the Response Time Graph to visualize the variation of the responses, and the Aggregate Report listener to obtain the performance report of each sample.

B. Experimental Design

Five e-commerce websites of Thailand based on user preference were chosen for the study. A short-structured interview from 50 participants on the popular websites they use for everyday shopping was used to choose the websites. The websites thus chosen are Lazada, Shopee, JD Central, JIB, and Kaidee. The Apache JMeter has been used to perform load testing on the target websites. A thread group of 1000 was created to simulate 1000 virtual users accessing the website simultaneously. The test plan of each website was similar and the tests were conducted in 05 iterations and the mean and median values were recorded. The results obtained were analyzed to get insights into the infrastructure of the websites and their capability to support concurrency under heavy load.

IV. RESULTS

The average time taken for a request, median time, the minimum and maximum amount of time taken for an HTTP request can be obtained from the listeners like an aggregate report. The results from 05 listeners are aggregated into Table 01 below.

Table 01: Response Information of the 5000 samples

E-Commerce Websites	Median (ms)	Mean (ms)	Minimum (ms)	Maximum (ms)
Lazada *(https://www.lazada.co.th/)*	3624	5033	332	60682
Shopee *(https://shopee.co.th/)*	5870	11591	0	51256
JD Central *(https://www.jd.co.th/)*	17703	20467	137	178578
JIB *(https://www.jib.co.th/web/)*	24722	122142	10	128737
PowerBuy *(https://www.powerbuy.co.th/en/)*	5457	101790	0	566761

The graph results listener of Apache JMeter showed many deviations of HTTP requests for each of the websites. Thus the responses are not normally distributed. Therefore, the median is considered as a metric to evaluate the time taken by the sample as the median value represents that 50% of the samples did not take more than this time. The median value is represented in Figure 01.

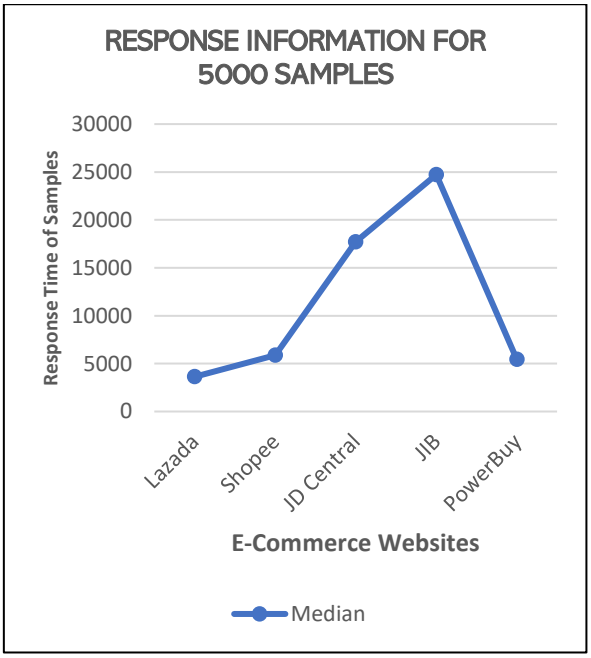

Figure 01: Response Information for 5000 samples (Metric: Median)

From the graph, it could be seen that the time taken to get the HTTP request from Lazada is the minimum and JIB consumes the maximum time. Further, the performance of Shopee and PowerBuy is comparatively better than that of JD Central. It could be inferred that the response time taken by the two least performing websites on heavy load is almost five times higher than that of the average performing websites.

The websites were also assessed in terms of their throughput during a heavy load of 5000 HTTP requests simultaneously. Throughput measures the number of requests catered by the website per second. The results are depicted in Figure 02.

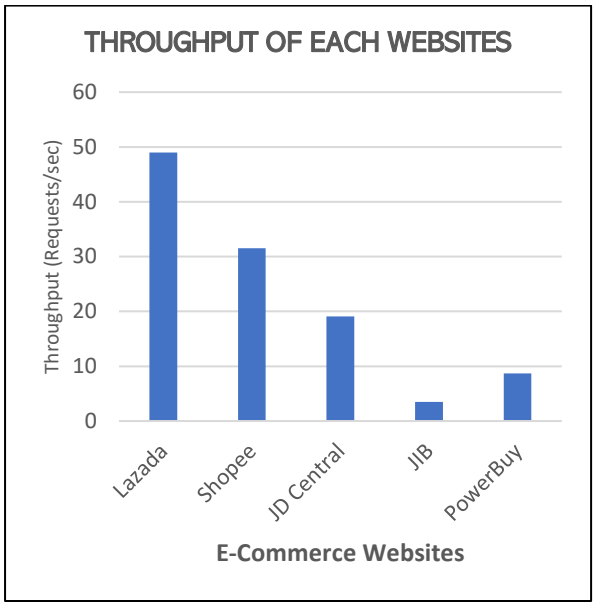

Figure 02: Throughput of each website (Requests/second)

The graph in Figure 02 shows that the throughput of Lazada is very high even at a heavy load of about 49 requests per second and JIB was able to support only around 3.5 requests per second. The throughput of PowerBuy is still lower (8.7 requests/second) when compared with Shopee (31.5 requests/second) and JD Central (19.1 requests/second).

The study also considered the use of error rate in evaluating the performance of websites during heavy load. Thus, the percentage of requests with errors was also recorded for analysis. The graph in Figure 03 gives a concise view of the error rate of the selected e-commerce websites.

JD Central website outperforms the other websites in terms of error rate, where it was error-free. The website Lazada also gives a tough competition with that of only 0.12%. On average, Shopee was producing an 18.52% of error rate while the error rate of JIB and PowerBuy was huge compared to others reaching almost 61.84% and 75.96%, respectively.

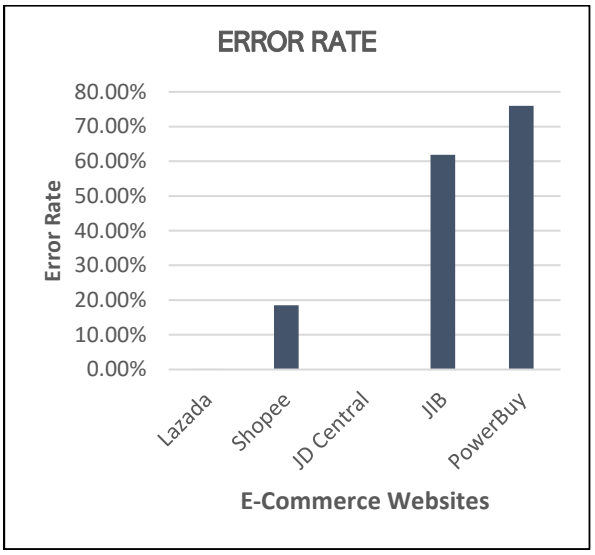

Figure 03: Error Rate (Percentage)

The listener "*Response Time Graph*" of Apache JMeter was used to generate the Response time graph of each website. Figure 04 provides a comparative view of the response time graphs. The response time graph of JIB is highly fluctuated. Comparatively, the response times of PowerBuy and Shopee remain high with the increase of load. Although the response time of JD Central was high initially, it has reduced greatly with time. The response time graph of Lazada is quite outstanding, where the time was high initially and turned low considerably quite earlier and reaching almost 100 milliseconds finally.

V. DISCUSSION

The performance of five websites for increasing load has been tested with the standard Apache JMeter tool. The websites were assessed based on the time taken to respond to an HTTP request, throughput, error rate. The above metrics were chosen based on the guidelines from the work of Menascé and Mason (2002). As in the same work, response time is calculated as the median time taken by the particular website to respond, availability is measured via error rate, and the requests per second are considered as the throughput. The comparison of results and the ranking provides an insight into the performance of the websites under heavy load. Figure 05 below

summarizes the findings from the study. The scores are given based on the ranking of the websites for a given metric.

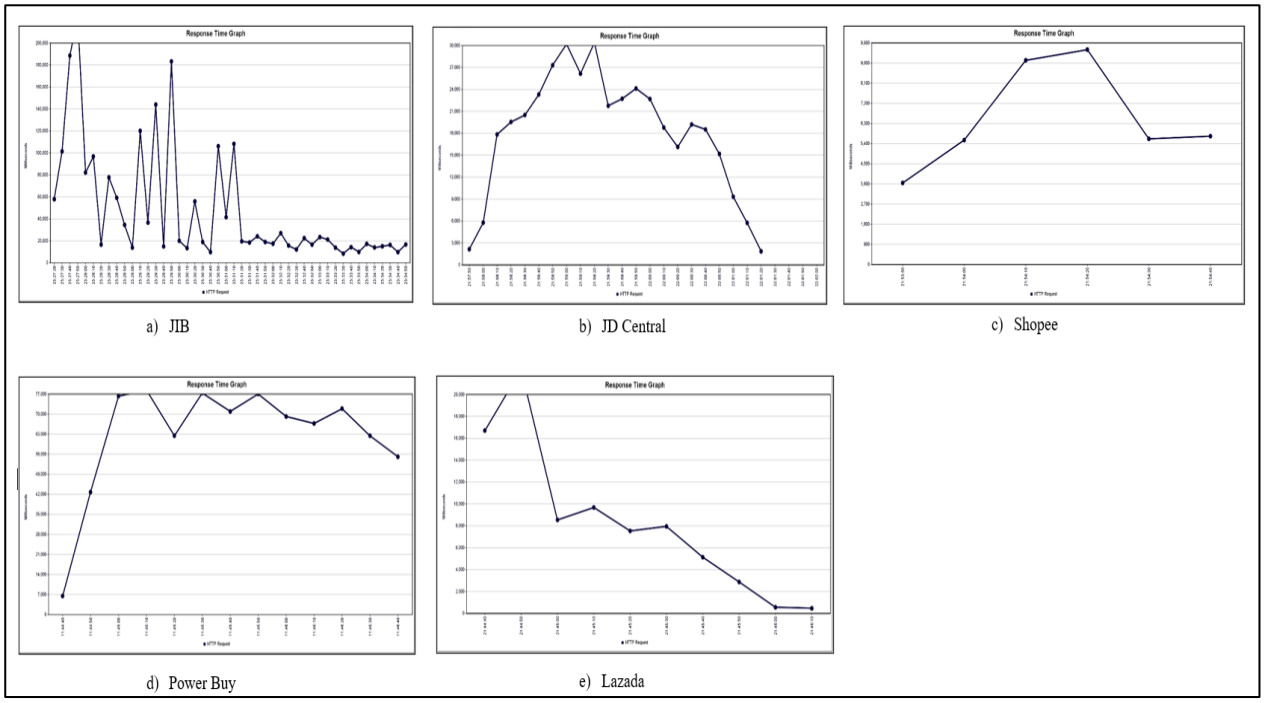

Figure 04: Response Time Graphs of websites

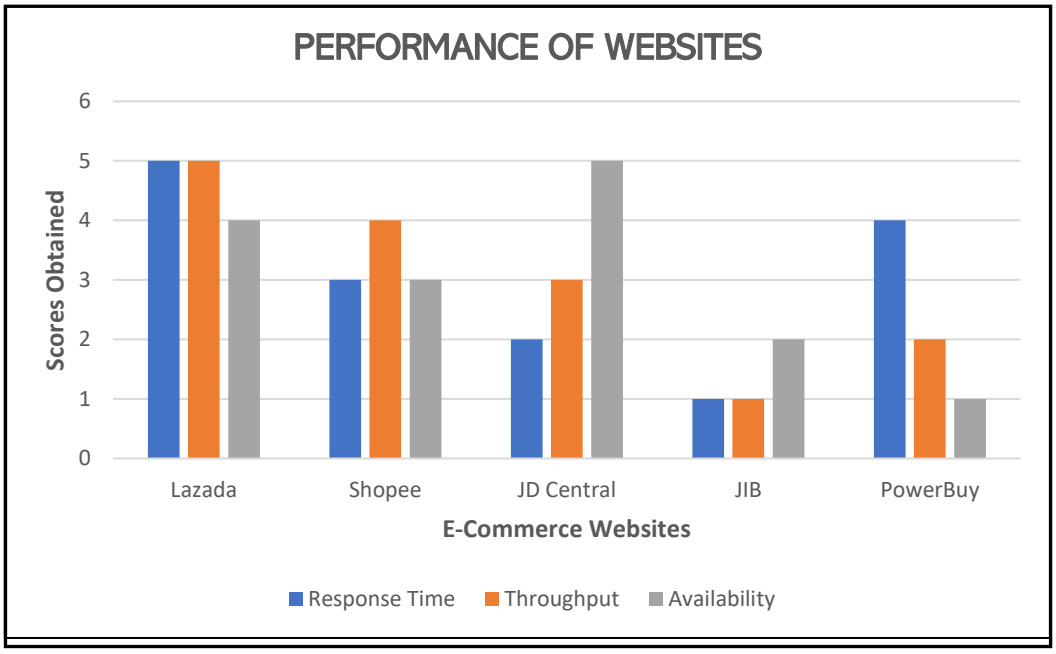

Figure 05: Comparison of websites under heavy load based on performance metrics.

An e-commerce website can be measured on various metrics such as its performance, usability, security, accessibility, and many others. However, performance is an essential measure since the outreach of a website depends on its performance. And for a large-scale system like e-commerce websites, the performance is a key measure since there is a high chance of getting hundreds or thousands of users simultaneously and the system should support concurrency. Thus, this paper serves as a model for performing load testing on e-commerce websites. The methodology of this paper can be utilized by developers and testing practitioners for testing their websites for their capability to support multiple concurrent users. The methodology can serve many large-scale software systems irrespective of their genre. The study is not intended to criticize the performance of websites rather use the selected websites to devise a methodology for load testing.

VI. CONCLUSION

The study utilizes the load testing feature of Apache JMeter to investigate the performance of 05 selected e-commerce websites in Thailand. The metrics used for evaluation are throughput, availability, and response time. The study was able to rank the e-commerce websites based on their performance metrics, where Lazada outperform the other websites. It could be seen that the architecture of Lazada supports multiple concurrent users. The study also highlights the importance of load testing on large-scale systems like e-commerce websites and proposes a methodology that testers and developers could utilize to assess their websites. The methodology presented here would be readily available for those interested in load testing without the need to review the literature every time. The study concludes that the appropriate metrics to assess the performance are throughput, availability, and response time. The e-commerce website architecture is considered robust if it has higher throughput, higher availability, and lower response time.

REFERENCES

Dhiman, S. and Sharma, P. (2016) 'Performance Testing: A Comparative Study and Analysis of Web Service Testing Tools', *International Journal of Computer Science and Mobile Computing*, 5(6), pp. 507–512.

Ghandour, A., Benwell, G. and Deans, K. (2010) 'Measuring the performance of ecommerce websites', *PACIS 2010 - 14th Pacific Asia Conference on Information Systems*, (January 2014), pp. 284–295.

Hamid, S., Bawany, N. Z. and Zahoor, K. (2020) 'Assessing ecommerce websites: Usability and accessibility study', *2020 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2020*, (December), pp. 199–204. doi: 10.1109/ICACSIS51025.2020.9263162.

Janani, V. and Krishnamoorthy, K. (2015) 'Evaluation of cloud based performance testing for online shopping websites', *Indian Journal of Science and Technology*, 8(35). doi: 10.17485/ijst/2015/v8i35/76200.

Jha, N. and Popli, R. (2017) 'Comparative Analysis of Web Applications using JMeter.', *International Journal of Advanced Research in Computer Science*, 8(3), pp. 774–777. Available at: http://libproxy.wustl.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=asf&AN=122961322&site=ehost-live&scope=site.

Jiang, Z. M. *et al.* (2009) 'Automated performance analysis of load tests', *IEEE International Conference on Software Maintenance, ICSM*, (September), pp. 125–134. doi: 10.1109/ICSM.2009.5306331.

Jiang, Z. M. and Hassan, A. E. (2015) 'A Survey on Load Testing of Large-Scale Software Systems', *IEEE Transactions on Software Engineering*, 41(11), pp. 1091–1118. doi: 10.1109/TSE.2015.2445340.

Khan, R. and Amjad, M. (2016) 'Performance testing (load) of web applications based on test case management', *Perspectives in Science*, 8, pp. 355–357. doi: 10.1016/j.pisc.2016.04.073.

Menascé, D. A. and Mason, G. (2002) 'of Web Sites', *Ieee Internet Computing*.

Musthafawi, A. Z. *et al.* (2020) 'Performance testing on the shopee website in the pandemic period of COVID-19', *ACM International Conference Proceeding Series*, pp. 195–199. doi: 10.1145/3427423.3427457.

Stefani, A. and Xenos, M. (2009) 'Meta-metric Evaluation of E-Commerce-related Metrics', *Electronic Notes in Theoretical Computer Science*, 233(C), pp. 59–72. doi: 10.1016/j.entcs.2009.02.061.

Wahyuningrum, T., Kartiko, C. and Wardhana, A. C. (2020) 'Exploring e-Commerce Usability by Heuristic Evaluation as a Compelement of System Usability Scale', *2020 International Conference on Advancement in Data Science, E-Learning and Information Systems, ICADEIS 2020*, pp. 1–5. doi: 10.1109/ICADEIS49811.2020.9277343.

Wijaya, I. G. N. S. *et al.* (2021) 'E-commerce website service quality and customer loyalty using WebQual 4.0 with importance performances analysis, and structural equation model: An empirical study in shopee', *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, 7(2), pp. 107–124. doi: 10.26594/register.v7i2.2266.

Yin, X. *et al.* (2021) 'Mass Tourism Data Analysis API Based on E-Commerce Platform', *Journal of Physics: Conference Series*, 1881(2). doi: 10.1088/1742-6596/1881/2/022065.

Zhao, P. *et al.* (2019) 'Behavior analysis for electronic commerce trading systems: A survey', *IEEE Access*, 7, pp. 108703–108728. doi: 10.1109/ACCESS.2019.2933247.