Using R language and open source software architecture to build a high efficient enterprise market integration platform

Tung-Shou Chen, Jeanne Chen, Li-Hsuan Lai[*]
Department of Computer Science and Information Engineering National Taichung, University of Science and Technology Taichung, No.129, Sec. 3, Sanmin Rd., Taichung City 404, Taiwan
[*]Corresponding author: Li-Hsuan Lai, graduate student, s1810532017@gmail.com

Abstract

Currently, there are many types of software for website information analyses. However, there also are some problems which include security issues using third party software, expansion weaknesses and operational complexities of software. The marketing integration platform proposed in this research, has resolved the third party security problem. The platform is developed in R language and open source software architecture environment. Its functionality is expandable and can be arbitrarily invoked to add more features. The platform operation page can be custom designed by user. This reduces the difficulties of user's interface.

Key words: *R language ; website analyses; open source software; data mining; customer relationship management*

1 Introduction

In recent years, e-commerce and online sales are rapidly increasing. Business management attitude is different from the past. The online marketing market has become a trend. The most important first step in marketing is to be able to arouse the interest of consumers visiting the page. There is need for a website analysis system to record the interactivity details. Web analytics is the best tool to help businesses operating on the e-commerce sites. Analysis on a website is the core of any research firm that has an online business (*H. Pakkala et al.*[1]). Information on a site is analyzed through web analytics. Where the analyzed results can be shown on a visualization chart. Enterprises can then quickly understand the current situation of the site, and have the opportunity to find more sales opportunities. Website Analysis is analysis of enterprise-oriented interactivity with main items that include clicks, page browsing and user sessions (*S. Jana et al.*[2]). This study is based on these three interactivity. Based on

the results from the website analysis an enterprise is able to forecast if a surfer will eventually becomes a customer and whether if the customer will become a long-term customer and to establish Customer Relationship Management (CRM). There are many definitions on customer relationship management, but the core idea is that the customer is an important asset to an enterprise. The purpose is to help enterprises manage the sales cycle, where new customers are continuously recruited and old customers remain loyal. In the process is to provide customer service and enhance the relationship between enterprises and customers to enhance the business in the customer point of view, as well as the opportunity to introduce new customer sources. A successful introduction and implementation of customer relationship management business is expected to be able to obtain customer loyalty for the enterprise and to maintain profit (*I. J. Chen et al.*[3]).

The most popularly used, on the market site analysis and customer relationship management software, is the website analysis software Google Analytics (GA) (Google Inc.[4]). Google Analytics planning is based on the requirements of the public as a priority and cannot meet the customized requirements of many. Therefore, the individual customer is unable to plan use of applications. Google Analytics handles web data in a way that is processed by a third party. Since the third party platform process data collection and analysis on site, there will be privacy issues and there is no way data leak or unlawful use. The platform for this study can improve Google Analytics privacy issues using R language and open source software architecture.

2 Method

The platform diagram of this study is as shown in Fig. 1, which consists of two blocks. The first block is the website analysis system. The system mainly analyzes user behavior on the site model and observe the site traffic and many other web information. This allows the enterprise to be in control of the current situation of the page. The second block is customer relationship management system. The system helps companies to target different groups of customers and to manage individual customer, to help enterprises to target different groups in order to design marketing strategies. Businesses are able to conduct marketing through the system and to manage the web visits of every customer or potential customers. The analysis tool uses the open source software R language.

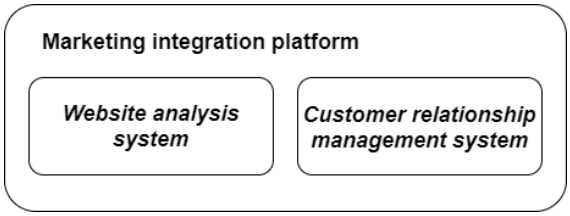

Fig. 1 – Platform architecture diagram

2.1 Website analysis system

To carry out the website analysis system, first is data collection. Data collection is a useful feature in the collection page for use as a source for subsequent analysis of data. The flow chart of the data collection is as shown in Fig. 2. When a user visits a page, the system collects the user's behavior. The data is then uploaded to the database for use as a follow-up analysis. Most of the enterprise side does not have a complete collection and use of the page information on the site. The website analysis system proposed in this paper must obtain the data of these website pages. This study uses Page Tagging to insert a tracking code into the webpage to help the enterprise to collect information about the page (*M. Khoo et al.*[5]). Collected content includes time information, user messages, and source messages.

The website analysis system records and the behavior of surfers after they surf a page. The web data quickly transfers the current page status, page views, bounce rate and dwell time to be stored. The analysis of the pageviews is presented in year, month and day, so that the number of surfers and the number of long-term surfers can be quickly understood by way of pageviews. Bounce rate means that users who browse the page just look at the page after the jump, there is no entry into any page within the next site. High bounce rate means that the page is not enough to attract consumers. Actually, to retain the eyes of consumers creates a chance for marketing opportunities. So the company can use the high bounce rate to assess whether there is need to re-plan the design page and to achieve the purpose of bounce rate decline. It can be used to observe the user for the entire site operating procedures by analyzing the user model. Button events is to collect the click rate for each button. Duration analysis is a way to help us understand the more detailed information about a surfer’s page, including how long it takes for surfers to stay on the site. Source analysis is the analysis of a surfer source of surfing and the devices used. Advertisement performance analysis allows enterprise to understand the effectiveness of all kinds of advertising. Through the analysis the marketing strategy can accurately target customers. The flow chart of the website analysis is as shown in Fig. 3. The website analysis system is used to select the function to be analyzed and to adjust

the parameters (date, time, page, etc.) to be analyzed. The next step is using the R language operators to derive the information you need from the database. R language allows data to be visualize and to create a suitable visualization chart for different data types. Finally, the results is displayed as a graph on the web page.

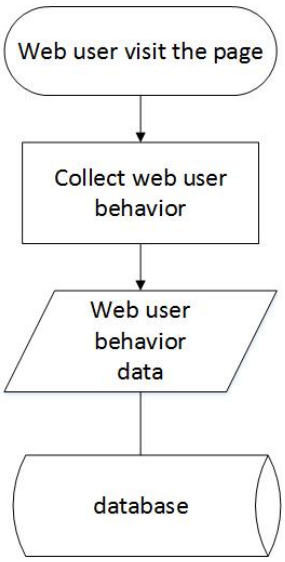

Fig. 2 – Data collection flow chart

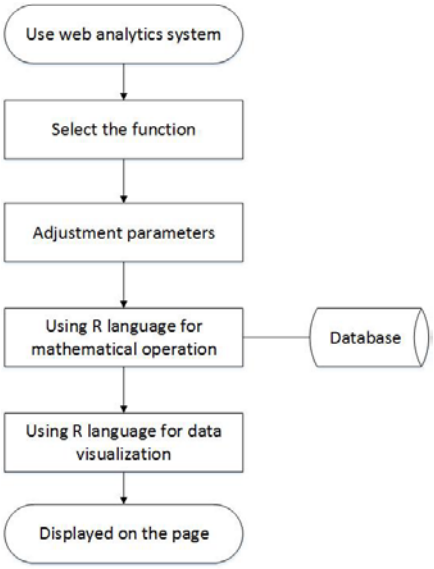

Fig. 3 – Website analysis flow chart

2.2 Customer relationship management system

The proposed system provides different stages of enterprise management customers. It is expected to be divided into two blocks to perform customer relationship management. The first block is to help integrate customer information, including consumer behavior patterns, transaction records, demand types and transaction credit. This part must be integrated with the site analysis system to achieve a complete integration of customer information. The second block is the use phase. This research emphasizes the design and develop, to meet the needs of the enterprise side of the customized needs for the purpose. From the large number of customers in the integration of information, customers similarity will be divided into different customer base. The management and marketing of customers will be used in various marketing strategies to achieve the purpose of increasing revenue and reducing costs.

3 Results and discussions

The experimental data from this study are from the data collected from an ecommerce website (Skyschool web.[6]). The data to be analyzed is from April to August 2017. The platform proposed in this paper analyzes the amount of web pages, bounce rate, surf time, device information and advertising effectiveness through the website analysis system. The

experimental results are as shown in Fig. 4 and Fig. 5. Fig. 4 shows the monthly data for June, where it Days 3 to 6 shows a large amount of data. Data volume is from 500 to 600 for a single-day volume. These days is the date of advertising. Advertising data can be used to attract surfers to visit the website is one of marketing methods. Through the report enterprises can clearly understand the current situation of the web page. The enterprise can quickly be updated on the current status of the page and whether the product has attract customers. The analysis of the surfing behavior is as shown in Fig. 5. Fig. 5 shows that the number of surfers through the Line link and QRCode scanning is low. The results show that most of the surfers use the google search directly to surf. The analysis provides information to the enterprise to help the enterprise to understand the source of marketing and to give the enterprise follow-up decision-making marketing methods to promote products. The platform proposed in this paper through the customer relationship management system helps enterprises understand the relationship between customers. The experimental results of the customer relationship management system are shown in Fig 6. The results show that the number of visits to the customer and the progress can be through the customer's visits and the progress of the customer group which then grouped into management and marketing. Experiments show that the difference in the degree of customer relationship management between Google Analytics. Google Analytics is not a complete combination with the enterprise, so the customer management can only be representative. The customer relationship management system and the enterprise are integrated to achieve the purpose of data synchronization. This allows enterprises to operate their own operation to analyze the details.

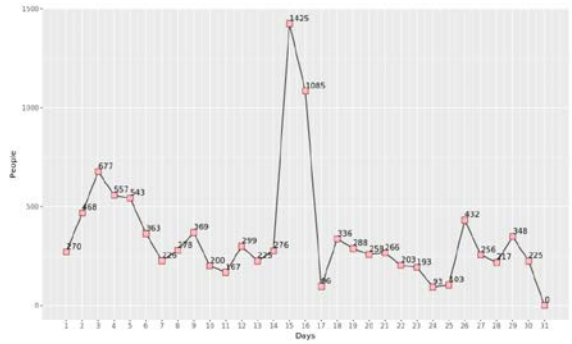

Fig. 4 – Report results graph

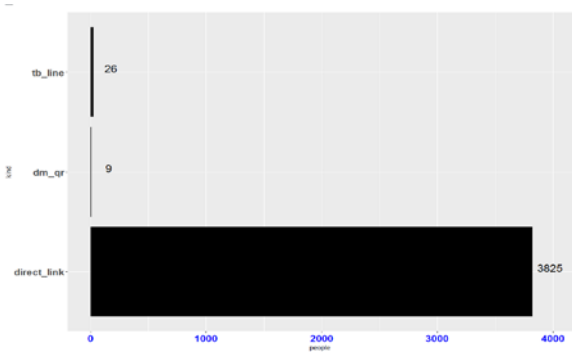

Fig. 5 – Analysis of the results of the visit

User_id	School_name	Number of visits	Page progress
5577	Kaohsiung City private love Kyrgyz children kindergarten	28	Registered
5384	Kaohsiung City Xinxing District seven Yin national primary school attached kindergarten	5	EDM page
2871	Hsinchu County Private Shihlin Kindergarten	3	EDM page
201	Taipei City Private Love Kindergarten	1	Registration system
1046	New North City Private Education Kindergarten	1	Management system
1392	New North City Private Happy Jing Siyuan Kindergarten	1	skyschool index

Fig. 6 ─ Analysis of the results

4 Conclusions

This paper presents the use of R language and open source software architecture to build a marketing integration platform for related research. The purpose is for making sound e-commerce marketing strategies to increase sales and profits. Experimental results show that the combination of the two systems are the most direct and effective way for the enterprise to assist the marketing planning system. The system analyzes the project to provide the most reliable data on the marketing plan. The use of R language and open source software architecture developed by the platform resolves the information security, and enhance the functionality of the expansion and ease of operation. In the future, the system could be extended to the entire commodity market trend analysis, so that the function of the marketing platform is more complete.

References

1. *H. Pakkala, K. Presser, and T. Christensen*, Using Google Analytics to measure visitor statistics: The case of food composition websites, International Journal of Information Management, 32(6), (2012) 504-512.
2. *S. Jana, and S. Chatterjee*, Quantifying Web-site visits using Web statistics: an extended cybermetrics study, Online Information Review, 28(3), (2004) 191-199.
3. *I. J. Chen, and K. Popovich*, Understanding customer relationship management (CRM) People, process and technology, Business process management journal, 9(5), (2003) 672-688.
4. Google Inc, Google Analytics, URL: http://www.google.com/analytics/, 2009.
5. *M. Khoo, J. Pagano, A. L. Washington, M. Recker, B. Palmer, and R. A. Donahue*, Using web metrics to analyze digital libraries, Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries, ACM, (2008) 375-384.
6. Skyschool web, URL: https://www.skyschool.com.tw/skyschool/skyschool/.