Decision Making Process for B2C Model Using Behavior Analysis with Big Data Technologies

K. Kannan[1]* and K. Raja[2]

[1]Department of Computer Science and Engineering, Sathyabama University, Chennai - 600119, Tamil Nadu, India; kkannanap@gmail.com
[2]Department of Computer Science and Engineering, Alpha College of Engineering, Chennai - 600035, Tamil Nadu, India; raja_koth@yahoo.co.in

Abstract

Objectives: Business to Consumer (B2C) E-Commerce activities are developed with a large number through agent-based systems. Case Based Reasoning (CBR) has been applied in these systems by analyzing the consumer buying behavior to provide consumers, a support to the decision making process. **Analysis**: Current applications of CBR to E-Commerce are limited to fixed, unchangeable products. To make the environment support for configurable products, an interactive operator based customization approach from CBR can be applied. **Findings:** In this work, to make the process more reliable and efficient, real time data from provisional stores has been taken and the system is trained to predict the consumer buying behavior along with CBR to pave way for a consumer to make a better decision making process. **Applications/ Improvements:** This work also applies big data concepts in predicting the behavior of the consumers. It thereby also led the customers to mine about their preferences in purchasing necessary products.

Keywords: Big Data, Business to Consumer E-Commerce Activities, Case based Reasoning, Multi Agent Systems, Semantic Web

1. Introduction

E-Commerce is trading of products or services with the help of networks. The platform lets the people to buy or sell their products or goods. For retail environment online sales have become an important and negligible way for their business development. Software agents are programs to which one can delegate (aspects of) a task. The concept of SMAC has come into emergence in recent environment[1].

Retailers are trying to develop SMAC based environment, whereas to improve the performance of the system. Social networks, Analytics, Cloud environment and Mobility will play a major role in today's environment. These concepts have an inter relationship among them while implemented and is mandatory for a particular business sector have to succeed.

Social networks enable two way communications between enterprise and target audience. Analytics are highly recommended in each system in which predictive analytics are carried out to understand the behavior of the consumers. Mobility describes the transactions carried out using mobile platforms. Nowadays sectors improve their sales and target by developing mobility in their system by enabling sales through mobile apps.

The cloud based environment is booming in today's world. Each and every sector trying to host their services in a cloud based environment. Hosting in cloud zoom up the speed of the system compared to physical hosting. The environment provides auto scaling to handle unpredictable demand spikes. It scales up according to the volume of the traffic. Due to this adoption, infrastructure cost gets cut down and site availability is improved.

This paper is organized as follows, section 2 deploys

Author for correspondence

models and methods, section 3 describes about evaluation of the system, section 4 depicts about data analysis and findings, and the work is concluded and the future work is depicted in section 5.

[2]Depicts about the Integral based decision making by the possibility theory for the propagation Uncertainty.The E-Recommendation process is considered to carry out to illustrate the approach. According to different criteria, a generic mechanism based on the arithmetic mean of the members' evaluations is proposed to customers.

Recentdays, due to the tremendous increase in a number of E-Commerce websites[3], customers are not only unable to decide which websites are the most attracting and the most reasonable for their purchasing, but also no clarity is seen in differentiating the services offered by the shopping sites. So it paved a way for new generation of websites, that is, there commended websites. [4]States that increasingly large numbers of pharmaceutical drugs, paired with the ongoing research activities in the Medical field, make the task of discovering relevant information difficult.

[5]Uses Natural Language Processing to retrieve the uncooked XML data and applies data mining to obtain user preferences automatically. For the purpose of extraction it uses metaMarker, an automatic metadata generation tool based on XML. It uses both dependent and independent data sets for processing of the inputs and deliver the result as fast it can. Unlike other information extraction systems, metaMarker retrives and segrate only the business level instructions. The Key criteria for the product selection may be four segments, such as multiple functions, multiple parameters including shape, size, weight, and so on, a wide spectrum of colours, and a large range of price.

The fuzzy decision system was implemented using shopper's preferences to provide recommendations in the E-Commerce systems[6]. To manage the collection of trust related items, one should understand about trust. Trust can be categorized into two namely indirect and direct trusts. Direct trust illustrates the trust moulded from straight involvement or negotiations, whereas indirect trust is acquired through recommendations from trusted third party. Uncertainty prevails through the indirect trust environment. To handle the uncertainty fuzzy logic is used.

Research on Customer loyalty based upon activity in B2C E-Commerce was conducted. Relational trust is considered important while handling customer loyalty. In online purchasing, the quality of the product, quality of service, value of customer, satisfaction, trusts and loyalty have a direct or indirect effects. A bounded relationship exists between customer and the seller based on the network services.

Product recommendation is provided to infrequently purchased products like car, trucks etc[7]. This is obtained based on user behavior and opinions. Adaptive collaborative filtering is used by utilizing user and product profiles to recommend products. ARM was used to create designs and association rules, to apply among the values in the attribute of products. ROSETTA, a tool that can extract the association rule from the rough sets in the information system is used.

Customer reviews are summarized based on product reviews[8]. Association rules & LDA model are joined to retrieve product features and cache words of an item. Corpus formats are prime focused so as to suggest the asset of opinion words.

2. Methods for Deployment

2.1 Case based Reasoning

Case Based Reasoning (CBR), the name itself suggests that it is a kind of reasoning. CBR is a significant tool in Neural Networks & Artificial Intelligence, which it mainly focused on mathematical concepts like logic predicate and propositional logic[9]. There has been a prominent quest for new paradigms and roads for increasing the efficacy of CBR systems for decision support[10].

It is a reasoning system principally based on the history used to solve new hitches by analyzing solutions of
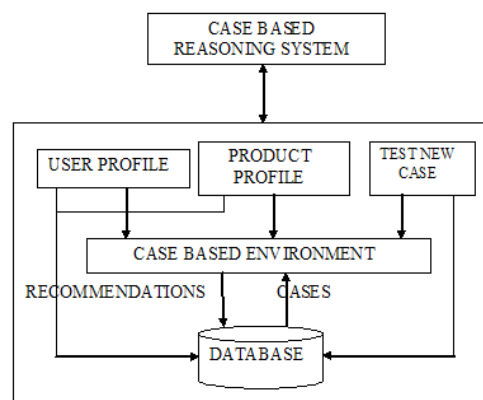

Figure 1. Case based environment.

the solved problems. CBR framework mainly consists of two phases, adding the existing cases and testing the new cases. The proposed work analyses the product purchase by the customer and predicts the likelihood of the product being purchased by the customer, which is depicted in the Figure 1.

The system uses user profile and product profile to test the new cases. In the case based environment, analysis of profiles is carried out. After the process gets completed, recommendations based on the previous cases are stored in the database. The CBR system can retrieve recommendations.

2.2 SMAC Environment

Retailers had been troubled by the gap between supply chain management and the solutions provided to the customers. E-Commerce environment leaves way for them to bridge the gap. E-Commerce integrates the intra and inter-company functions[11]. It provides a way for the retailers identify the gaps between different levels of supply chain, helps to develop ERP systems to manage operations between customers and suppliers, also helps to face political barriers or cross country changes.

Social impact of E-Commerce has been emerging and creating a great impact on the individuals and also in the business environments. Customisation of Web pages Paved the way for many critical phenomenon which can be met using Ontologies[12]. Semantic Web (SW), a modern methodology can be referred as a network of ontologies and other web resources[13]. The primary aim of the SW is to supplement the current internet by a tier of machine understandable content[14].

[15]Demonstrates that ontology-based literature mining is a prevailing approach for examining interaction networks and producing new systematic assumptions. Information Science, a domain which is built for sharing of knowledge by the building blocks Ontology. Domain language is made meaningful to both humans and machines[16] with the help of ontologies which defines the domain concepts and the relationships between them. Domain ontologies are used as the knowledge base to understand the meanings of the concepts. Ontologies are a central building block of the semantic web[17].

Mobile penetrations in our country have been developing tremendously. E-Commerce companies are also seen at tremendous up lift in transactions using mobile platforms. It is noticed that 95% of the internet traffic is generated through mobile. On owing to this, it is very much important for a retailer to accomplish mobility, for them to be succeeded.

Analytics are generally termed as the discovery and communication of meaningful patterns in data. Marketing organizations and retailers determine the outcomes of efforts and provide decisions for further investments through analytics.

Emerged Cloud Computing and Storage Access Network (SAN) solutions provide consumers and industry with various abilities to record and retrieve their data in data warehouses. This, in turn reduces the hosting costs for the organizations. By using the cloud environment, retailers can avoid huge investments in infrastructure and can concentrate on projects which distinguish their commerce. Cloud environment provides characteristics like agility, cost, device and location independence, maintenance, multi-tenancy, performance, reliability, etc.

The cloud computing platform can be deployed mainly through private, public and hybrid environment. Private cloud infrastructure that is completely operated by a single body can be managed within or by a third-party. In public cloud, the facilities are provided openly for the use of public. Hybrid cloud is a combination of two or more same or different clouds. The entities present in the cloud will not collide, each other protected by guard region serves well in all scenarios.

[18]States users availing cloud services intended to receive service from more clouds rather than from a single cloud. That is, a drastic change in the environment of variabes from single environment to multi valued environment. The consumer who requests for a complex service that cannot be fulfilled with single service, but has to take help of two or more services that are provided by various peers to satisfy the client's request[19].

Services are tiny compartments present on the web that cooperatively creates anentire application environment. Web Services are the passive platforms designed to produce results once it has been triggered with an input[20]. Web services are independent components that can be located, shifted, transferred and accessed through XML based Internet Protocols[21].

All the available services in the web interact with each other, for fulfilling tasks and for multi-tasking, in return, carry out parts of complex transactions or workflows[22]. This canbe categorized in to three: manual composition, semi-automatic, and automatic composition[23].

Automated Composition is one of the prime practice of Web Services. A surplus volume of work has been done for computerized web service configuration, but still there is a planetary to bulk up for particular requirements[24]. It is the sole responsibility of the retailers to choose between cloud platforms for them to adopt according to their requirements.

2.3 Big Data Lifecycle

Life cycle of big data can be visualized as a three phased approach viz., acquire, analyse, and actions which result in continuous optimization. The first phase in moving through the economy is to acquire data. In this case, the retailer can be able to understand the macro and micro environment trends, consumer behavior, their requirements, reactions, etc. Data from social media like blogs twitter etc., provides retailer a humongous amount of data regarding the customer.

The data dealt with Big Data is not in a regular form; it is unstructured and cannot be integrated directly into the traditional analytical tool.

In the third phase of actions, the retailer converts the data acquired from the analysis phase to actionable business decisions.

3. Evaluation of the System

Figure 2 describes the architecture of the scheme. Prediction of recommendations is carried out based on the user ratings with the help of user and product profiles. Contents are segregated based on the feature values, and the contents with similar features are recommended to the users. Figure 3 depicts about the evaluation of the
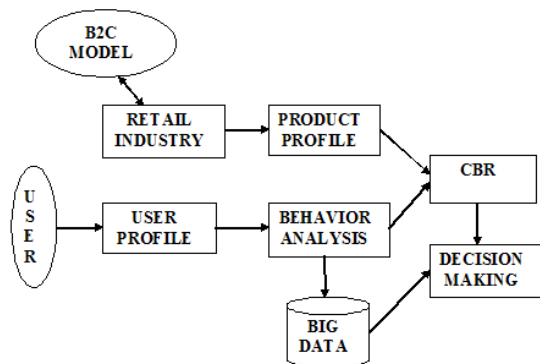

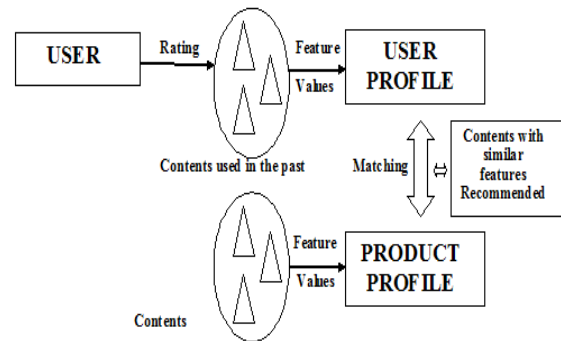**Figure 2.** System architecture.

Figure 3. Performance of the system.

system. The process that takes place inside the system is elaborated.

4. Data Analysis and Findings

4.1 Percentage Analysis

The data sets are collected from a leading provisional store of a city in Tamilnadu. It is found that the number of female respondents is more compared to the male respondents. Also, the age group of consumers falls at a majority of 25 to 50. Married couples visit the store more often than unmarried persons. It is inferred that the average monthly income of the respondents is above 15000. Consumers are visiting the stores based on word of mouth marketing rather than other means of marketing. The survey is processed using tools like percentage analysis, chi-square test, and correlations.

Table 1. Relationship between consumers and Product quality

Age group of consumers	Product Quality			Total
	HS	S	NS	
Below 25	14	32	26	72
25 - 50	30	68	70	168
50 - 60	22	54	38	114
Above 60	20	74	52	146
Total	86	228	186	500

Table 1 represents the relationship between consumers based on different age groups and product quality. Product quality is compared based on the products available in the stores (HS – Highly satisfied, S – Satisfied, NS – Not Satisfied).

Table 2 represents the chi-square test results, the calculated value of 0.625 is greater than 0.05. Hence, the null hypothesis is accepted.

Table 3 denotes the correlations between the people who respond to the survey and the offers/discounts provided by the stores. Pearson correlation coefficient value, R=0.133. The numbers are discreetly correlated. So, there exists a relationship between the gender of respondents and the offers/discounts provided by the provisional stores. It is inferred that the nature of products bought by different groups of people varies accordingly.

Table 2. Chi-square test results

Particulars	Value	Degree of freedom	Significance
Pearson chi-square	7.118	9	0.625
Likelihood ratio	7.157	9	0.621
Linear - by -linear association	0.857	1	0.355
No.of Valid Cases	500		

Table 3. Correlation between respondents and offers in stores

Variables	Particulars	Gender	Offers/ Discount
Gender	Pearson correlation	1	-0.067
	Significance		0.133
	Number of respondents		500
Offers/ Discount	Pearson correlation	-0.067	1
	Significance	0.133	
	Number of respondents	500	

5. Conclusion

The work deals with decision making process for business to consumer model with the help of behaviour analysis. This process along with the SMAC environment to provide retailers with a better environment for them to improve their business. Recommendation system makes recommendations to the consumers regarding user and product profiles based on feature values. The future work of this paper is to handle big data concepts more effectively and efficiently so as to increase the demand of customers worldwide.

6. References

1. Guttman RH, Moukas A, Maes P. Agents as mediators in electronic commerce. Intelligent Information Agents, Springer-verlag: Berlin Heidelberg; 1999. p. 1–6.
2. Denguir-Rekik A, Mauris G, Montmain J. Propagation of uncertainty by the possibility theory in choquet integral-based decision making: application to an e-commerce website choice support. IEEE Transactions on Instrumentation and Measurement. 2006; 55(3):721–28.
3. Yu Z, Zhang L, Zhang C. Web service composition method based on FAHP and TOPSIS. Journal of Communications and Information Sciences. 2012 Apr; 2(1):75–84.
4. Sun H, Wu H, Li S, Liu M. The customer loyalty research based on B2C ecommerce sites. 2010 IEEE International Conference on E-Business and E-Government (ICEE), Guangzhou; 2010 May. p. 3156–59.
5. Zhang S, Gu M. Applying data miming to automatically acquire user preferences in commercial web sites. IEEE International Conference on Information Engineering and Computer Science, Wuhan, ICIECS09; 2009. p. 1–4.
6. Wu Z, Wu H. A fuzzy decision system using shoppers preferences for recommendations in ecommerce applications. Ninth IEEE International Conference on Intelligent Systems Design and Applications,ISDA09, Pisa; 2009. p. 803–8.
7. Abdullah N, Xu Y, Geva S, Chen J. Infrequent purchased product recommendation making based on user behavior and opinions in e-commerce sites. 2010 IEEE International Conference on Data Mining Workshops, ICDMW, Sydney, NSW; 2010. p. 1084–91.
8. Liu LZ, Wang WT, Wang HS. Summarizing customer reviews based on product features. 2012 IEEE 5th International Congress on Image and Signal Processing, (CISP), Chongqing; 2012. p. 1615–19.
9. Sun Z, Finnie G. A unified logical model for CBR-based e-commerce systems. International Journal of Intelligent Systems. 2005; 20(1):29–46.
10. Srinivasan S, Singh J, Kumar V. Multi-agent based decision support system using data mining and case based reasoning. International Journal of Computer Science. 2011 Jul; 8(4):340–9.
11. Al-Shrouf F, Turani A, Al-Shqeerat K. Software agents for e-Commerce data workflow management. Software Engineering and Computer Systems, Springer-Verlag Berlin Heidelberg; 2011. p. 96–106.

12. Vigneshwari S, Aramudhan M. Social information retrieval based on semantic annotation and hashing upon the multiple ontologies. Indian Journal of Science and Technology. 2015; 8(2): 103–7. DOI: 10.17485/ijst/2015/v8i2/57771.

13. Anandaraj A, Kalaivani P, Kumar VR. Development of ontology based intelligent system for software testing. International Journal of Communication and Innovation Journal. 2012; 2(2):157–61.

14. Cimiano P, Lucker J, Nagel D. Unger C. Exploiting ontology lexica for generating natural language texts from RDF data. Proceedings of the 14th European Workshop on Natural Language Generation; 2013. p. 10–19.

15. Hur J, Ozgur A, Xiang Z, He Y. Identification of fever and vaccine-associated gene interaction networks using ontology-based literature mining. Journal of Biomedical Semantics. 2012; 3(18):1–18.

16. Kalaivani P, Anandaraj A, Raja K. Ontology construction approach for the domain of poultry science using protege. International Journal of Information Technology and Management Sciences. 2011; 1(2):134–40.

17. Kalaivani P, Anandaraj A, Raja K. Retrieval of knowledge based information using ontological classification in poultry management. International Conference on Competency Building Strategies in Business and Technology for Sustainable Development. Sri Sai Ram Institute of Management Studies, Chennai in association with Asian Management Science Association; 2011.

18. Choi JH, Lee S-H, Kim M-K. Integrated user authentication method using BAC (Brokerage Authentication Center) in Multi-clouds. Indian Journal of Science and Technology. 2015; 8(25):1–7. DOI: 10.17485/ijst/2015/v8i25/80241.

19. Mannava V, Ramesh T. A composite design for pattern for service injection and composition of web services for peer-to-peer computing with service oriented architecture. International Journal on Web Service Computing. 2012 Nov; 3(3):49–63.

20. Shah MA, Azam F. QOS based service search and composition algorithm. 2012 IACSIT Hong Kong Conferences, IPCSIT12, IACSIT Press: Singapore; 2012. p. 121–6.

21. Yu J, Cai H, Bu F, Liu A. OWL-S based service composition of three-dimensional geometry modeling. Journal of Multimedia. 2012; 7(1):98–105.

22. Rohallah B, Ramdane M, Zaidi S. Semantic web service discovery based on agents and ontologies. International Journal of Innovation, Management and Technology. 2012 Aug; 3(4):467–72.

23. Liu C, Liu D, Han N. A novel web service composition algorithm for multiple QoS constraints. International Journal of Software. 2012 Aug; 7(8):1867–72.

24. Allauddin M, Azam F. Dynamic web service composition and parameters matchmaking. International Journal of Computer Applications. 2011 Dec; 36(9):21–6.