



INTEGRATING E-SERVICE WITH A OIL REFINERY E-COMMERCE USING DATA MINING

T.Saranya¹, J.K. Anu Shakthi Priya², K.Poornima³

^{1,2,3}School of Computing Science and Engineering,
VIT University, Vellore, Tamil Nadu, INDIA.

Abstract: E-business is the relevance of information and communication technology (ICT) to carry out all the business activities. When any business go online, it has to decide which e-business model best suits their organization's goals. A business model should describe on what grounds the organizations generates, distributes and confines values in profitable, societal, intellectual and in other circumstances. To select best appropriate business strategy we made use of association rule mining and also we integrated both e-business models and cloud computing features to overcome the general problems faced by online vending. Common trading involves consumer demands such as shipping, mobile access, customization, information and convenience. All these mentioned strategies are explained and illustrated with case study: Oil Refinery process.
Keywords: E-commerce, cloud computing, data mining, association rule, oil refinery.

I. Introduction

In this day and age 'e' is gaining drive and if not the whole thing is made digitally possible. There are types of e-commerce and it will be chosen according to the type of bond involving diverse faces of commerce. Some of the types which being made use in this paper is as follows, B2B- Business to Business, B2C- Business to Consumer, C2B- Consumer to Business, B2I- Business to Internal.

B2B: The commerce transactions between businesses i.e., between manufactures and wholesaler or retailer are illustrated by this type. The transaction degree will be comparatively higher than others and prime cause behind this is that the sequence of process involved in the production and distribution of a commodity have many B2B transactions of secondary elements or unprocessed stuffs. It also involves the circumstances of exchanging data and partnership.

C2B: It is a type which is vice versa of B2C i.e., user proposes services or products to third party therefore company will compensate them. For example let us put forth the elance, it is one of the website to promote this type of business. In this it permitted vendor to publicize their proficiency and eventually bargain hunter to advertise their ventures.

B2I: It is an emerging segment of electronic commerce types. Here it distinguishes the people's online acquiring routines are alike whether they're retailing for individual or for their industry use.

B2C: As name implies this type of businesses supplies the end users with their goods or as service. Any business or organization which wishes to sell its products can make use of this business to consumer type of model for its best adoption. It not only deals with the online transactions but also the serves the e-banking, public sale in online, health related sites, travel services etc.,

Data mining is one kind of technique which is used to extract interesting patterns or knowledge from large databases. There are lot of techniques that have been used to find such kind of knowledge, most of them resulting from statistics and machine learning. The main aim is to discover accurate knowledge from large datasets. The task performed by data mining depends on what kind of knowledge someone needs to mine. Data mining techniques are the consequence of a lengthy process of research. The kinds of tasks performed by data mining techniques are Regression, clustering, classification, association and prediction.

II. Association Rule Mining

In data mining, the association rule is introduced to find hidden patterns from large datasets and conclude useful result on how a subset of items influences the presence of another subset. Let $T = \{T_1, T_2, T_3, \dots, T_n\}$ be a universe of items and $A = \{A_1, A_2, A_3, \dots, A_n\}$ is a set of transactions. Then expression $X \rightarrow Y$ is an association rule where X and Y are item sets and $X \cap Y = \Phi$. Here X and Y are called antecedent and consequent of the rule respectively. The main feature of this rule is support and confidence.

Support is a set of transactions in set A that contain both X and Y .

$\text{Support}(X \Rightarrow Y) = \text{Support}(XUY) = P(XUY)$

Confidence is percentage of transactions in A containing X that also contain Y .

$\text{Confidence}(X \Rightarrow Y) = \text{Support}(XUY) / \text{Support}(X) = P(Y/X)$.

If support $\geq$ minsup and confidence $\geq$ minconf then the association rule is termed as strong association rule. There are two types of association rules named as positive association rules and negative association rules.

Using Cloud Computing in E-commerce Application: In E-Commerce a new practice evolved as Cloud computing. The ecommerce and cloud computing has not yet stretched a advanced stage and still needs the test of practice. At present, there are still many glitches that requisite to be solved. They are policy security, supervisory issues of services, data refuge and privation of relevant laws and conventions. As procedure propagates, energy depletion and the atmosphere become more dissimilar, due to a countless variety of computing resources being offered, so grows managing complexity.

IV. Case Study (oil refinery)

Apriori Algorithm:

Apriori, the lain word means “from what comes before”. It is a classical algorithm used to mine frequent patterns from huge amount of data. It follows bottom up strategy. This algorithm works upon categorical data. It follows breadth first search technique.

Terms Used in Apriori Algorithm

➤ Min_supp:

Stands minimum support. It is used for searching frequent patterns that satisfy this threshold.

➤ Min_conf:

Stands Minimum confidence. It is used for finding the strong association rule that satisfy this constraint.

➤ Frequent Itemset (Li):

It is denoted by Li, where I means ith item. The item sets satisfies the minimum support (min_supp) threshold .

➤ Apriori Property:

This property is very useful for trimming irrelevant data. It states that any subset of frequent item set must be frequent.

➤ Prune step:

It is used to find frequent itemsets, for any (n-1)-itemsets that is not frequent cannot become subset of a frequent n-itemset .

Oil Refinery System:

The current paper shows the e business components used in the oil industry based on the largest companies worldwide. The top twenty companies in the worldwide oil industry based on their revenues are Permex,Ecopetrol,Suncor,Chesapeake,Surgutneftegas,Pertamina,Libya,Socar,Adnoc,Lukoil,Rosneft,Sinopec,Ecopetrol,Inpex corp,Woodside, Xcel energy,Holyfrontier,Nobel energy , Caltex, Osaka.

TABLE-1

B2B									
Items	P1	P2	P3	P4	P5	P6	P7	P8	P9
X1	1	1	1	1	1	1	1	1	1
X2	1	1	1	1	1	1	1	1	1
X3	1	1	0	1	1	1	1	1	1
X4	1	1	0	1	1	1	1	1	1
X5	1	1	1	1	0	1	1	1	1
X6	1	0	0	1	0	1	1	1	1
X7	1	1	0	1	1	1	1	1	1
X8	1	0	1	1	0	1	1	1	1
X9	1	1	0	1	0	1	1	1	1
X10	1	0	0	1	0	1	1	1	1
X11	1	1	1	1	1	0	1	1	1
X12	1	1	1	1	1	1	1	1	1
X13	1	1	1	1	1	1	1	1	1
X14	1	0	1	1	1	1	1	1	1
X15	1	1	1	0	1	1	1	0	1
X16	1	0	1	0	0	0	1	0	1
X17	0	0	1	0	0	0	1	0	1
X18	1	0	1	0	0	0	1	0	1
X19	1	0	1	0	0	1	1	0	1
X20	1	1	1	0	1	1	1	0	1

B2C																
Items	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16
X1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1
X2	0	1	1	1	1	1	0	0	1	1	0	1	1	1	0	1
X3	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1
X4	0	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1
X5	0	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1
X6	0	0	1	1	0	1	0	0	0	1	0	0	1	0	0	1
X7	0	0	0	1	0	1	0	0	0	1	0	0	1	1	0	0

X8	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0
X9	0	0	1	1	0	0	0	0	1	1	0	0	0	1	1	1
X10	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	1
X11	0	0	0	1	0	0	0	0	1	1	0	0	1	1	0	1
X12	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0	1
X13	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0
X14	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0
X15	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
X16	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
X17	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
X18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
X19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
X20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

C2B			
Items	R1	R2	R3
X1	0	0	1
X2	0	0	1
X3	0	0	0
X4	0	0	0
X5	0	0	0
X6	0	0	0
X7	0	0	0
X8	0	0	0
X9	1	1	0
X10	0	0	0
X11	0	0	0
X12	0	0	0
X13	0	0	0
X14	0	0	0
X15	0	0	0
X16	0	0	0
X17	0	0	0
X18	0	0	0
X19	0	0	0
X20	0	0	0

B2I								
Items	S1	S2	S3	S4	S5	S6	S7	S8
X1	1	1	1	0	1	1	1	1
X2	1	1	1	0	1	1	1	1
X3	1	1	1	0	1	1	1	1
X4	1	1	1	0	1	1	1	1
X5	1	1	1	0	1	1	1	1
X6	0	1	1	0	1	1	1	1
X7	0	1	0	0	1	1	1	1
X8	0	1	1	0	1	1	1	1
X9	1	1	1	0	1	1	1	1
X10	0	1	0	0	1	1	1	1
X11	1	1	1	0	1	1	1	1
X12	1	1	0	0	1	1	1	1
X13	1	1	0	0	1	1	1	1
X14	0	1	0	0	1	1	1	1
X15	0	1	0	0	1	1	1	1
X16	0	1	0	0	1	1	1	1
X17	0	1	0	0	1	1	1	1
X18	0	1	0	0	1	1	1	1
X19	0	1	0	0	1	1	1	1
X20	0	1	0	0	1	1	1	1

Note: X1-X20 are oil companies, including Pemex (X1), Ecopetrol (X2), Suncor (X3), Chesapeake (X4), Surgutneftegas (X5), Pertamina (X6), Libya (X7), Socar (X8), Adnoc (X9), Lukoil (X10), Rosneft (X11), Sinopec (X12), Ecopetrol (X13), Inpex Corp (X14), Woodside (X15), Xcel Energy (X16), HollyFrontier (X17), Nobel Energy (X18), Caltex (X19), and Osaka (X20).

TABLE-2

Analysis of e-Business components using association rule.

Company	B2B Count	B2B Support	B2C Count	B2C Support	C2B Count	C2B Support	B2I Count	B2I Support
X1	9	1	13	0.81	1	0.33	7	0.87
X2	9	1	11	0.68	1	0.33	7	0.87

X3	8	0.88	13	0.81	0	0.0	7	0.87
X4	8	0.88	12	0.75	0	0.0	7	0.87
X5	8	0.88	12	0.75	0	0.0	7	0.87
X6	7	0.78	8	0.5	2	0.66	7	0.87
X7	8	0.88	6	0.37	0	0.0	7	0.87
X8	9	1	5	0.31	0	0.0	6	0.75
X9	6	0.66	7	0.43	0	0.0	6	0.75
X10	9	1	4	0.25	0	0.0	6	0.75
X11	8	0.88	5	0.31	0	0.0	5	0.62
X12	7	0.78	3	0.18	0	0.0	6	0.75
X13	8	0.88	3	0.18	0	0.0	5	0.62
X14	6	0.66	4	0.25	0	0.0	5	0.62
X15	7	0.78	2	0.12	0	0.0	5	0.62
X16	7	0.78	0	0.0	0	0.0	5	0.62
X17	4	0.44	2	0.12	0	0.0	5	0.62
X18	5	0.55	0	0.0	0	0.0	5	0.62
X19	3	0.33	1	0.06	0	0.0	5	0.62
X20	4	0.44	0	0.0	0	0.0	5	0.62

Analysis of e-Business components using association rule:

In Table-1, the companies and e-business components are listed. 1 in table refers that a particular company implements that particular components and 0 refers that component does not implement.

In table-2, the support of particular components implemented by certain companies are calculated.

V. Conclusion and Future Work

Now a days several organizations were in the necessity to obtain benefits of electronic business and thus IT has exaggerated oil industry in abundant means. As far as oil industry is concerned the greatest benefit that technology offers to make process safer and the major focus will be on rate minimization and prevention of shortages at gas location so this can be made by energy security which engrosses the running of operations and proficient utilization of oil. This energy security will reduce the disequilibrium between the supply and demand. Almost all the applications like e-commerce, medicine, clustering, etc uses the Data mining concepts and techniques. In this paper association rule algorithm has been used which profits our mentioned area in many ways and some of the notable advantages are the improved efficiency of process and reduced I/O costs. Through the discovery it is understood that, the assimilation of e-commerce application model based on cloud computing guarantees huge volume of data storage, outstanding speed rate of computing, and cloud's own characteristic as its allocation and resource sharing. So it will pave a new way for even growth of e-commerce. Even though this is in the near beginning point of discovery, few problems are not well carried out yet. Whichever way, e-commerce application model based on cloud computing will not prevent its rapidity to progress. As the cloud computing increasingly used, e-commerce will undoubtedly marshal in a new-fangled era of cloud computing.

VI. References

- [1] Jelena Zascierinska, Andreas Ahrens, "E-business applications in Engineering education" in international Conference .
- [2] Aakif Nazeer Khan and M. Hussain , "Application of Semantic Web in e-Business and Telecommunication" in International Conference on Advanced Computer Control.
- [3] Myint Myint Khaing, Nilar Thein, "An Efficient Association Rule Mining For XML Data" in SICE-ICASE International Joint Conference 2006.
- [4] Z. Yang, W. H. Tang, A. Shintemirov, and Q. H. Wu, "Association Rule Mining-Based Dissolved Gas Analysis for Fault Diagnosis of Power Transformers" .
- [5] Idheba Mohamad Ali O. Swesi, Azuraliza Abu Bakar, Anis Suhailis Abdul Kadir, "Mining Positive and Negative Association Rules from Interesting Frequent and Infrequent Itemsets " in proceeding of 9th International Conference on Fuzzy Systems and Knowledge Discovery, 2012.
- [6] Security Attack Mitigation Framework for the CloudEsha Datta ,Neeraj Goyal , Indian Institute of Technology
- [7] On Evaluating and Publishing Data Concerns for Data as a Service Hong-Linh Truong and Schahram Dustdar Distributed Systems Group, Vienna University of Technology
- [8] Software Development for Cloud: An Experiential Study Marimuthu Cand K. Chandra Sekaran,Department of Computer Science and Engineerin,National Institute of Technology, Karnataka, India.