A Survey On: Analysis of Pattern Mining and Behavior Prediction in M-Commerce

Tejal V. Deshmukh
Student
Pune Institute of Computer Technology

Anant M. Bagade
Assistant Professor
Pune Institute of Computer Technology

ABSTRACT

Today is the world of science, mobile technologies, web applications, internet applications, business transactions these technologies can be combined to online business on mobile device which will increase the business performance. This is a real world practical application where a need arises to think about user using the system application which will run on mobile device. Thinking of user who makes actually transactions with the mobile device is a prime requirement. The main concept is to think about end user who interacts with the system taking into concern of m-commerce and mobile mobility, location area, data mining, behavior of user spending patterns which will have an enormous effect on business industry and to the society also. These methods also increase growth in various kinds of apps in mobile devices from low version phones to smart phones. So a need generates to use these in our transactions to locate shops, malls which will be situated at some distance far from our place and predict the user behavior in purchasing or making transactions into the shop with help of low version phones to smart phones. Since every user does not have smart phones so a system or an application should be made for the user using simple phones to generate automatic view of nearest shops or malls, with help of cellular phone. This will naturally make the business industry grow and give benefit to users which all counts for an m-commerce business economy industry.

Keywords

M-commerce, Mobile phones, Mobile mobility, K-nearest algorithm for capturing nearest areas(states, cities),data mining, behavior prediction, images, location area.

1. INTRODUCTION

The various methods like data mining and pattern mining being used nowadays but there are many different areas like m-commerce industry and e-commerce where each and every day millions of transactions take place. The user will surf a lot and search for related items and then move to the original page and all these traditional methods are being followed. so why not to provide the user with an application which executes on cellular devices and which can be carried to a long distance and performing business transactions irrespective of various areas through the cell phone. There are various concepts about similarity of store measuring similarity among stores and items generate the next behavior. Performing combination of these entire tasks will give benefit to the user and the economic business industry also. Today transaction on mobile phone are banking, internet banking, online shopping, mobile banking, call set up and call end, internet apps, different applications are being performed and there is no work in m-commerce area like location wise transactions(purchasing). Giving the user a view to select and not to search through the basic model of mobile phones i.e. simplest phone should perform different applications.The idea behind this view is to create a system in the cell phone where the user can do all the transactions on a phone which can find its behavior depending upon its previous made transaction with help of set of images the next behavior should come automatically seeing the selecting method of the user from images of items while making transactions in the cell phone. M-commerce is a wide area not only related to business but to other areas also. The different areas can be linked to m-commerce(Mobile-commerce) as to spatial databases, multidimensional cubes, huge complex where the user doesn't have to search but locate all the required things in all huge buildings, malls, shopping complex. For e.g. an unknown user from a different state enters into a complex arcade searching for some products where user gets confused which product is situated where and the way to reach and all these types of related queries creates a question. So such a system in the mobiles should be designed to ease the user path and the transactions made and to detect the future behavior of the next coming transaction. The system should automatically show user the path to reach upon giving the input to system then the system should give the correct path and location to reach to that area. This all working method will be executed on a cellular device. This is a very challenging area and work to be done in the m-commerce field which would give a rise to economic industry and to m-commerce economy also. This does not end here while m-commerce is wider area like for e.g. a user goes to a bank for deposit of money for more ease this work should be done on cell phone even the money with drawl should also be done on cell phone itself i.e. through a mobile device the user should interact to a bank for all transactions like opening bank account, with drawl, deposit and so on through the user handheld device. The other industry is medicine industry where day to day transaction is made these also should be performed on mobile device. In the medicine field the patient visits number of times to the doctor or gives a call what the patient stores in its home place medicines and looking at the image of patient the doctor can diagnose all these things through a mobile device. Images and symbols also play a very important part in surfing the whole data and selecting appropriate task which the user wants. The selecting task for the user is easier as compared to text since images play a clear view and idea to the end user. The K-nearest algorithm will help practically to locate the nearest locations where the user wants to visit to a shop, malls, retailers, this will provide an ease to the end user who actually work on mobile devices and concerning to the present

behavior of purchasing the next coming behavior will be notified on cellular device through an set of images of items into the shops related items from users selection. Considering the present behavior and next coming behavior through image set will be provided. The paper is divided into six different sections the first is Introduction, Literature survey, proposed concept, Implementation details, General architecture, Conclusion and Future scope.

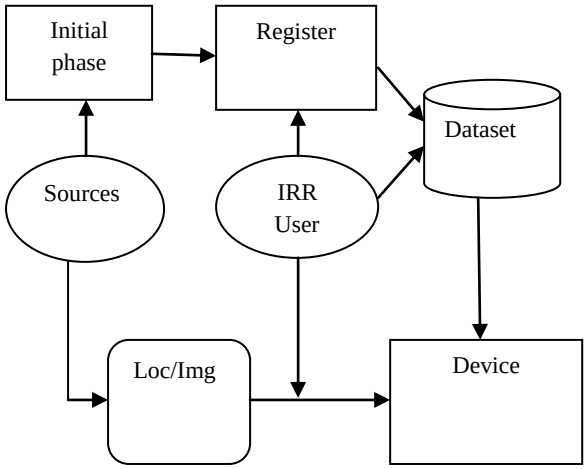

Fig 1: Working phases in M-commerce

The fig.1 gives an general idea of the layout in the system which is being developed. There are mainly four methods and its related components. The most important block is the device where the actual system will be executed. These blocks are as follows. First method is initial phase which consists of various sources in the dataset these sources may from different kinds of subjects. The second method is register which consists of IRR (Irregular user) and RU (Regular user) which has to generate first authentication for security point of view. The third method is the dataset which consists of images, locations=shops/malls, paths, transactions, shopkeeper items/products information. The last one is most important is the device method where the application will have an icon and will be executed on an actual mobile device.

2. LITERATURE SURVEY

The related work is divided into various methods[1] which mainly focuses on framework for pattern mining and prediction in M-commerce this gives an framework called MCE(Mobile commerce Explorer) for mobile users movement and purchases done by them. This is followed by three different components named SIM (Similarity Inference Model), PMCP (Personal mobile Commerce Pattern Mine) and the third is MBP (Mobile behavior Predictor). The first component is similarity inference model which measures the similarities among stores and items where set theory is not applicable. The idea behind this method is any two objects are similar if they are related to similar objects. The next component is mobile behavior prediction which is divided into two categories the first is vector based prediction and the other is pattern based prediction. The idea behind vector–based prediction its predicts the very next location of the object according to its direction and velocity. Here the mobile behavior of a user is assumed that it can be represented in mathematical form based on user recent movement in geographic information. The second is pattern-based prediction in this patterns are taken which match with user's recent mobile behavior. The SIM model works on two methods 1) two stores are similar if the items they sell are similar 2) two items are dissimilar if the stores that sell them are dissimilar. An algorithm is also proposed called PMCPs algorithm which is transformed into three sub parts which are 1) frequent transaction mining 2) mobile transaction database transformation and 3) PMCP mining.

SimRank [2] also specifies nearby similarity to improve the efficiency of Sim-Rank then Sim-tree came into existence.

The other component is mobile pattern mine which focuses on pattern mining from transaction data and the data related to mobile. An apriori algorithm [3] for mining data for association rules. The next is WAP-Mine [4] to discover web access converted into tree data structure. SMAP-Mine [5] also proposes for mining efficiently patterns based on FP-Tree. The sequential pattern is taken into account [6] moving paths from left and right side of the mobile sequential pattern. Hybrid prediction models are the combination of both vector based and pattern based methods. The other concept is behavior patterns in mobile web a system [7] which mainly focuses on services given by mobile phone anywhere and anytime. The ease of mobile behavior increases system performance and quality of services offered by mobile devices. This concept is that user's do not have much time to surf so response should be in very short period. This also proposes sequential mobile access pattern which comprises of sequential movement with requested service. Here both parts have been considered like movement patterns and the services. There is no attention paid to particular user behavior which works on images. The other proposed method is Markov Model[8] which effectively gives services, location and location with appropriate services, moreover three rules are being generated next location, next requested service, and next location and associated service. Here the N-gram model has been extended. In this main focus is on purely mobile behavior user patterns. The other concept represents mining frequent patterns without candidate generation[9] here the main concept is any long patterns are being generated the novel Frequent Pattern (FP-Tree) which is an tree structure storing compressed information about frequent occurring patterns and generate an FP-tree. Here this FP-Tree works in three ways 1) large databases are compressed into smaller data structures to avoid same scan number of times 2) the second is frequent mining to avoid generation of large sets and 3) a partitioning based with divide and conquer method is used to mine smaller task to reduce space.

The next proposed is Transaction Management for m-Commerce at mobile terminal[10] which concentrates on distributed transaction properties fascinating an m-commerce view an wireless terminal based transaction manager architecture with requirement analysis. The main idea is architecture is based on a concept that, a certain application that has some business purchased transacts which uses a transaction manager to store state information upon a link application or an terminal down. There is a way to represent a TM called ontological transaction monitor used between application and servers when any m-commerce transaction are being made.

The next concept [11] is mining mobile sequential patterns in a mobile commerce environment which states a mining model called sequential patterns in mobile commerce which takes moving and purchase patterns into consideration this also consists of three algorithms to find frequent sequential patterns from mobile transactions. If a user moves in an m-commerce environment transactions are being done and those

are being put in a chart pattern and underlined in a schematic diagram. For e.g. {(A,{i1}),(B),(C,{i2,i3})} this simple equation states that A is a store and i1 item has been purchased B is null and C is a store where two items are being purchased they are i2 and i3.

The next method states[12] an algorithm to work on frequent pattern mining based on apriori which proposes apriori is the first to be used in frequent pattern mining for large data using knowledge discovery uses three different techniques called record filter, intersection and a algorithm.

3. PROPOSED CONCEPT

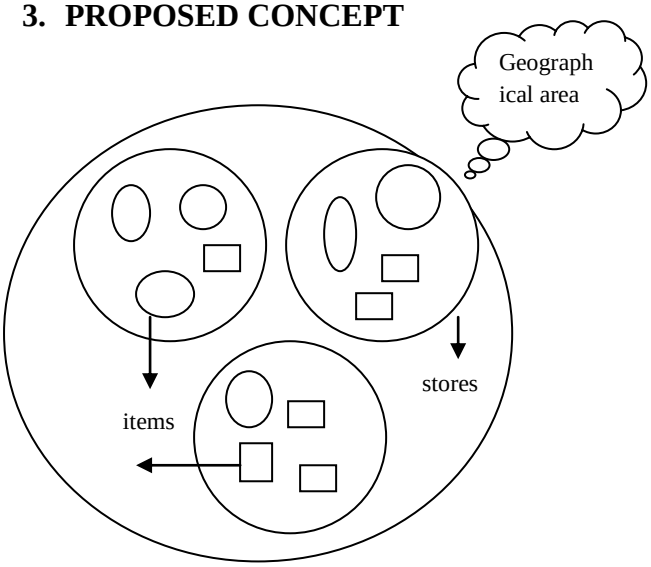

Fig 2: Schematic stores in respective areas

Fig 2.gives a schematic view of the geographical area wise stores and the different items present in the stores. These products or the items vary from each other and range from different brands to other brands. The given figures also vary from brand to brand. There are "N" number of stores here it shows three for ease of the user so these stores are shown in a circular fashion in which all is surrounded by an geographic area this area constitutes for a single region/state. There are different provinces where different geographic area are situated. For a different geographic area different store required items are also generated. The above method can be shown by set theory for example which is an easy method to understand.

Ti= Total items in stores

Ti= {(Ti-Si), (Tj-Sj), (Tk-Sk)}

Then Tijk= Dijk(Discounts/Free vouchers)/Tijk%

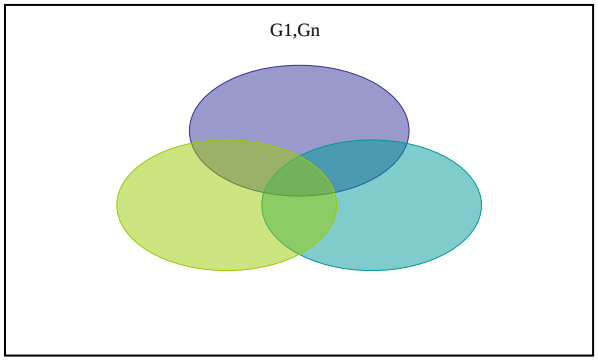

Fig 3 Distributed geographical area

Fig 3. which states if there are different areas at far distance in kilometers the effect will be on to the stores also. Since these areas may have an effect from nature, atmosphere, environment, humidity, temperature which will have an impact on the present items in the stores with respect to the stores underlying its respective area. The items in the stores vary due to the geographic area or not. The transactions made by the user the purchasing strategies of the items calculate the user behavior and the patterns can be generated from it this will count the spending patterns corresponding to user behavior. Fig 3. Shows G= {g1...gn} which is the area gives locations and locations gives information about stores. Fig 3 has three different areas where the shops may have similar, common items or not. The related work up till now does not have this concept some apriori are used for association rules then FP are frequent pattern generation in large data but taking geographic area with images and giving it a support of various stores situated in that particular area is an novel approach. Whole system will work on a mobile device and will make an easy transaction to the user. Applying k-nearest algorithm to the system will gives nearest possible known locations with its corresponding stores, the items, products will have an impact on specified area. An algorithm is been proposed which will generate the user behavior respect to the underlined purchased transactions. The next transaction will be automatically uploaded to the user for the ease of surfing through the system. The user doesn't have to survey or search or around to look for the products/items but the user will get appropriate choice depending upon user's transaction methods.

4. IMPLEMENTATION DETAILS

The system will be implemented in Microsoft visual.net framework. The programming is in c# language with dataset. Simulation of whole system will be done on android simulator and finally on mobile device. The dataset will match to present database for its efficiency, performance and scalability. The framework design will be shifted to an actual mobile device having an android operating system. Operating system will be windows 7 / windows XP. The whole application will be converted to a file and an icon will be created so that system will execute on an android mobile device. This whole system will have a dataset into it and will operate on low version phone to higher version phone.

5. GENERAL ARCHITECTURE

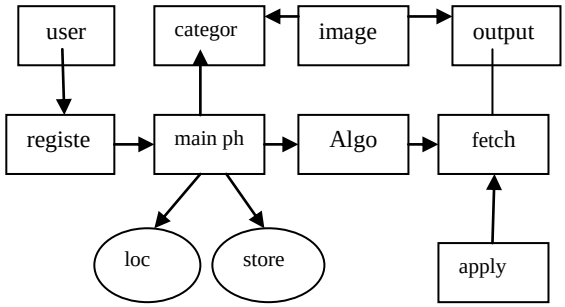

Fig 4 System Architecture

Fig 4. States a general system architecture where user is subdivided into two types regular and nonregular. Regular user is one who visits the shops number of times purchases items and makes transactions in the shops. The non regular user is the one who may be a window shopper go through user survey user the one who visits the shop or the system just for sake. These both users must be registered while entering into the system. If at all the user is a regular considering user behavior patterns the items will be displayed to the user if the user wants to select another product/items user is free to do so. The irregular user has to first clarify details and then move forward to enter the system.

Once user is registered system creates a file and enters in main phase where user is categorized by two ways and is separate by its location. As the location is being decided various related images will be displayed to user of the nearest areas with help of an algorithm which will find the nearest shops/ malls/ arcade close to user location. The personal behavior will be judged depending on user personalize behavior algorithm. This algorithm will decide very next coming spending behaviors of particular user so that the user may not need to surf and go to abounded web data. The user will automatically get a loaded image which will correspond to its behavior from selecting/ spending patterns in the shops. This is the main phase where two algorithms will work out and measure the behavior of the user. The user will fetch the whole data with the enclosure will be displayed to user to fill certain criteria. Finally the user will go to final stage and receive some offering from the shops/arcade.

The above procedure is for a single geographical area which can be shown mathematically as follows,

For G1={(Rr =>Lr=>Sr={set of images}=>(i1,i2,i3,i4)}=u1

For G2={(Rr=>Ld=>Sd={images}=>(j1,j2,j3)}=u2

For G3={Irr=Lirr=>S={images}=>(k1,k2,k3,k4)}=u3

For above equations geographical area(G1) corresponds to regular user(u1) with location(Lr) provided with image set purchasing items (i1,i2,i3,i4). The second area(G2) which has an regular user(u2) but different location(Ld) with set of images and has items (j1,j2,j3). The third is for irregular user(u3) with different location nearby store to user first(u1) has images with respect to items (k1,k2,k3,k4) for the third user. Fig 4.shows system architecture is divided into three phases they are initial phase, main phase and lastly processing phase. The first phase is the initial phase which consists of two user which will create its own log in the system to confirm itself. The second phase is initial phase which is the focused phase consists of different categories of images of various products and the related items. The third phase is the processing phase where the whole data is taken and fetched to user behavior algorithm where the analysis is done for the user's behavior and which automatically predicts the user very next behavior in the form of images.

6. CONCLUSION AND FUTURE SCOPE

The work up till now is done on various area's like similarities between the stores, frequent patterns generation, frequent tree data structure generation some work also is been done on pattern mining using apriori and other is business transactions using a transaction manager. This system application will provide an novel approach where a k-nearest algorithm is used for location selection and personalize behavior of the user will be given by personalize user behavior algorithm. Moreover the whole system is being operated on android cellular device for the ease of the user to avoid number of clicks and scroll and select. Today a need arises to think practically of the areas where the sales ratio is very less and to boost the performance of shops and to give benefits to the user is the prime importance. The user will select only images based on the patterns related to user transaction surfing methods. The future scope of the system is to design an algorithm which will work for dissimilar stores in a region and to present in a Dissimilar Location Stores Model (DLSM).

7. ACKNOWLEDGEMENTS

This research was partially supported by Pune Institute Of Computer Technology, Pune and would like to thank all who have contributed directly or indirectly.

8. REFERENCES

[1] Eric Hsueh-Chan Lu, Wang- Chien Lee, Vincent S. Tseng "A Framework for Personal Mobile Commerce Pattern Mining and Prediction ",IEEE Transactions On Knowledge And Data Engineering, Vol. 24, No. 5, pp 769-782 MAY 2012.

[2] Glen Jeh, Jennifer Widom, "SimRank: A Measure of Structural-Context Similarity", Stanford University, pp 1-11.

[3] R. Agrawal and R. Srikant, "Fast Algorithm for Mining Association Rules," Proc. Int'l Conf. Very Large Databases, pp. 478-499, Sept. 1994.

[4] J. Pei, J. Han, B. Mortazavi-Asl, and H. Zhu, "Mining Access Patterns Efficiently from Web Logs," Proc. Pacific Asia Conf. Knowledge Discovery and Data Mining, pp. 396-407, Apr. 2000.

[5] V.S. Tseng and W.C. Lin, "Mining Sequential Mobile Access Patterns Efficiently in Mobile Web Systems," Proc. Int'l Conf. Advanced Information Networking and Applications, pp. 867-871, Mar. 2005.

[6] C.H. Yun and M.S. Chen, "Mining Mobile Sequential Patterns in a Mobile Commerce Environment," IEEE Trans. Systems, Man, and Cybernetics, Part C, vol. 37, no. 2, pp. 278-295, Mar. 2007.

[7] Vincent S. Tseng , Kawuu W. Lin, "Efficient mining and prediction of user behavior patterns in mobile web systems" , Science Direct Institute of Computer Science and Information Engineering, National Cheng Kung University, No. 1, Ta-Hsueh Road, Tainan 701, 2006.

[8 J. Pitkow, P. Pirolli, Mining longest repeating subsequences to predict World Wide Web surfing, in: Proceedings of the USENIX Symposium Internet Technologies and Systems, 1999.

[9] Jiawei Han, Jian Pei, Yiwen Yin, "Mining Frequent Patterns Without Candidate Generation", School of Computer Science, Simon Fraser University, pp 1-12.

[10] Jari Veijalainen, Vagan Terziyan, Henry Tirri," Transaction Management for M-Commerce at a Mobile Terminal", proceedings of the 36th Hawali International Conference on system Science-2003.

[11]Goswami D. N, Chaturvedi Anshu, Raghuvanshi C.S, "An Algorithm for Frequent Pattern Mining Based On Apriori" , International Journal on Computer Science and Engineering Vol. 02, No. 04, 942-947, 2010.

[12] ShopKick, http://www.shopkick.com/index.html, 2010.