A Framework of an Automated Data Mining Systems Using ERP Model

First A. Abdullah S. Al-Mudimigh, Second B. Zahid Ullah, Third C. Farrukh Saleem

Abstract- **The proposal of the paper is based on the data mining effects using ERP framework. Using the ERP database we applied data mining applications to evaluate the best result for the growth and establishment of a company. In this paper we proposed a model which integrates the database, customer queries, transactions, and all other specifications used in ERP systems, then use data mining techniques to integrate decision making and forecast flows. By using ERP's characteristics and background we gathered the data from central database in cluster format which is based on the action taken against the queries generated by the customers. Furthermore, the clustered data used by Apriori Algorithm to extract new rules and patterns for the enhancement of an organization. This is a complete implementation of data mining applications on ERP framework to predict the solution of upcoming queries. This will make the best association between the customers and organization, and customer will always satisfied with company's policies.**

*Index Term-***ERP, Data Mining Applications (Clustering & Apriori)**

I. INTRODUCTION

The core of the ERP system circulates within the company as well as the management information and control needs of the entire production processincluding reducing inventory, labor, and operation costs, improving business processes to enhance operation efficiency and improving customer response [1]. However, Markus and Robey (1988) pointed out that although industry specific ERP has already focused on industry characteristics and includes the optimal business operation management model, the promotion of ERP is still significantly related to interaction with the organization [1]. Beyond that CRM often describes a strategic or philosophic approach for managing customers [2]. Hence CRM could be seen from a process oriented, technological, capability-oriented, philosophical, and/or strategic perspective [2]. CRM has different meanings, It is a business strategy to select and

manage customer to optimize long-term value, CRM also is a strategy that increases the importance of relationship marketing and integrates with other organization strategies [3,4]. Despite the efforts of CRM on efficiency and affectivity of management decisions, statistics reflect that unreasonable decisions are made by CRM systems; IT experts said errors were caused by incorrect or not enough data [5]. According to Forrester Research, 57% of business firms cannot justify CRM investments because they cannot measure customer profitability [9]. The necessary broad level of security is determined by the CRM security objectives which an organization needs to meet[10]. A CRM security strategy outlines in general terms how an organization will achieve its CRM security objectives [10].

In order to increase the use of ERP systems it is recommended to begin with the financial section, the applications invoicing, cost control, accounting and financial then it should be added many functions from Financial, Relation Management Production, Distribution, e- Business and Analyses [6]. The order of the customer is routing automatically to the next department when one department finishes their work of the customer order and each department have access to the single database that holds the customer's new order [7].

II. EXPERIMENTAL MODEL

The presented model shown in the Fig.1, described the abstract of all concerned departments exist in every organization. We generally divided the structure of an organization in three layers. Each layer has its own departments and officials having specific responsibilities and burden. This is basically the combination of CRM and ERP which are essential part of the enterprise. We enhanced this abstract by adding knowledge discovery layer,

where the rule generator plays the special and critical role to generate new rules and patterns from the large database by using data mining techniques (Clustering, Association, Classification). We implemented our model on MADAR (Customer Based Organization) data, by using clustering and Apriori algorithms for the generation of new rules and recommended these rules for the future perspective of MADAR. According to the MADAR environment whenever a customer request engender directly forward to the proper department of CRM Layer for the assessment and positive response. CRM works in front of any organization after receiving the query it will redirect to the ERP layer for the appraisal and configuration of these queries. After statistical analysis and evaluation of the request the answer back to the corresponding customer and this feedback will be saved in the database for the future requirements. Now the work start in knowledge discovery view which select

F. A. Author is an Associate Professor in the department of Information Systems, King Saud University , Riyadh, Kingdom of Saudi Arabia. Mob: 00966-504171443, mudimigh@ksu.edu.sa.

S.B. Author is a Researcher in the Department of Information Systems, King Saud University, Riyadh, Kingdom of Saudi Arabia. Mob: 00966-543807381, zahidullah45@ksu.edu.sa.

T.C. Author is a Researcher in the Department of Information Systems, King Saud University, Riyadh, Kingdom of Saudi Arabia. Mob:00966-530706625, farrukh800@ksu.edu.sa.

some of the useful data from the large database for its work.
Description of Three Layers

A. CRM Layer

Whenever a customer contact with the company the customer support officer receive customer's request. In the company's prospectus this department has much importance because of correspondence directly with the customer. In our model we presented customer relationship management as an outer view. CRM is responsible for receiving requests and replying to the customer directly. These requests includes queries, complaints, suggestions and orders then forward these requests to the inner view enterprise resource planning (ERP) through the query generator. After taking action on the perspective request the answer will forward

through the CRM Layer. And result will also be saving in the database for knowledge discovery view.

B. ERP Layer

The important part of the model is ERP Layer. In this view each department have equal access to a single database that holds the customer's data or complaints. In this layer the customer queries rotating and evaluating by the concern department. For example a customer contact with customer support department (CRM layer) after initial review and statistics this query will throw to the ERP Layer. The ERP's department is responsible to find the solution and give proper reply to the customer and forward feedback back to customer through CRM layer as well as in the central database for future assistance.
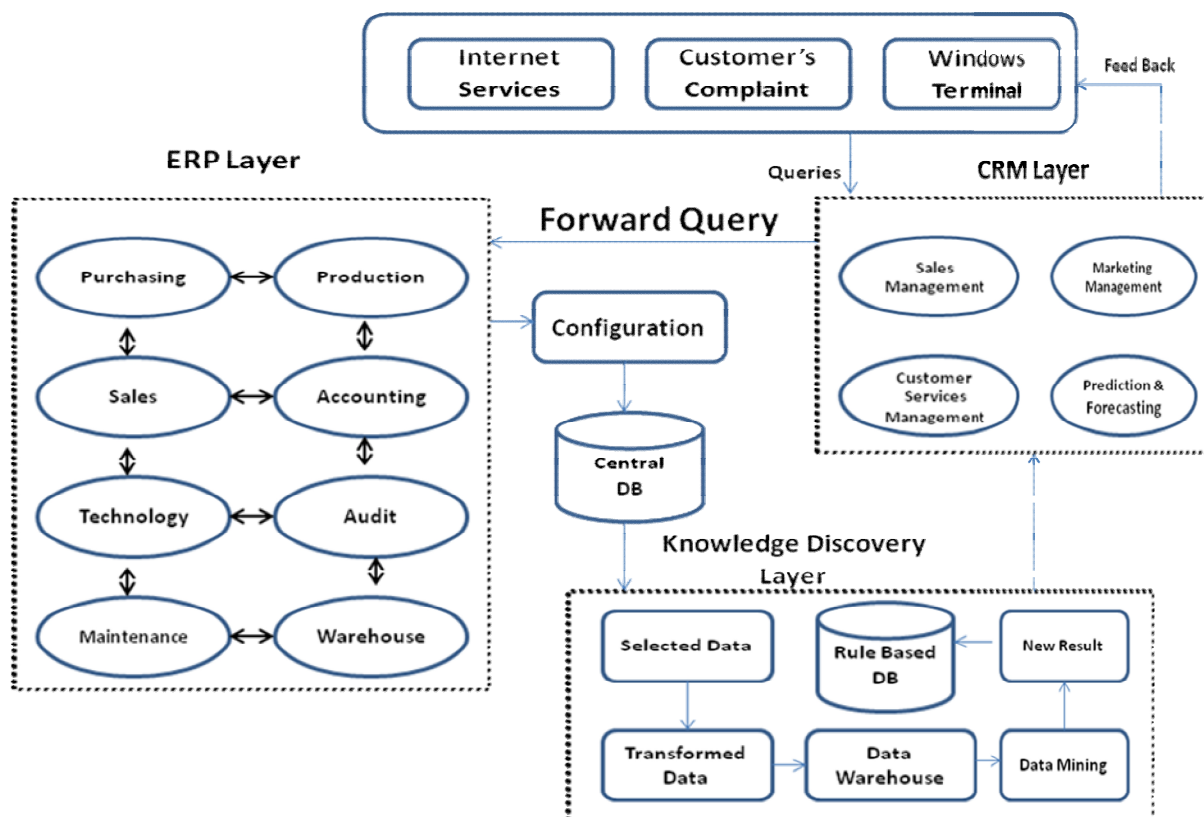

Fig. 1- ERP – Data Mining Model

C. Knowledge Discovery Layer

Knowledge discovery can be defined as the extraction of contained, hidden and useful information from the large database. As per the definition, in the presented model this layer is also concern with the central database having all kind of data posted from any department of ERP and CRM layers. The data has several attributes and characteristics according to the queried customer. As the specialty of this layer we used Clustering and Apriori algorithms for advantageous and high-quality results.

III. FINDINGS AND RESULTS

For the practical implementation of presented model we applied data mining technique on MADAR data. MADAR is an ERP based organization working in King Saud University dealing with all administrative software of the

university and also do work for outer projects. The efficient implementation of presented ERP-CRM model using data mining techniques is applied on MADAR data. First we applied clustering on the data thereafter For this we used Association mining- Apriori Algorithm for finding new rules and patterns from the experienced data. The description of all phases on our case study are as follows:

A. Programming Tool Used

For the implementation of Apriori Algorithm to get new rules on the MADAR data we used VB technology. Input and output data files can be in any format e.g databases, Notepad, Excel worksheets, etc. But in this case we used Notepad and Excel worksheet as input and output files.

B. Data Clustering

Data clustering is a method in which we make cluster of objects that are somehow similar in characteristics. The

criterion for checking the similarity is implementation dependent. Clustering is often confused with classification, but there is some difference between the two. In classification the objects are assigned to pre defined classes, whereas in clustering the classes are also to be defined [12]. The goal of data clustering, also known as cluster analysis, is to discover the natural grouping(s) of a set of patterns, points, or objects [13].

B.1. Clustering Implementation

Clustering implementation on the data is a process in which we try to make some groups of objects together having similar characteristics according to the pre-defined criteria. The criteria to merge similar objects in one group are depending on the implementation. In this case study we collected data shown in Fig. 2. In which we have four (04) kinds of attributes includes; customer's query, queried department, forwarded to concerned department and action of the department. We clustered the data on the basis of actions of the customer's query. Moreover, action's column has again four (04) types of action have taken against the particular query. The similarity criterion which we set here in this case is dependent on the action. We described these actions as numbers 1,2,3,4 for gathering them in separated groups shown in Fig.3. As the result we got four different clusters, in which each cluster is fulfilling the selection criteria shown in Fig.4. Cluster's evaluation graph is shown in Fig.5

S.No.	Cutomer's Query	Queried Department	Forwarded to: Concerned Deparment	Actions of the Department
1	Bugs in Software	Software Maintenance	Debuggers	Declined
2	Salary Calculation mistakes	Salary Department	Programmer-Salary Dept.	Need some time
3	Changes in contracts for new employee	Employees Department	Programmer-Employee Dept.	Call a Meeting
4	Salary structure for each designation	Salary Department	Programmer-Salary Dept.	Replied on the spot
5	Budget approval for coming year	Finance Department	Programmer-Finance Dept.	Declined
6	Changes in budget in current year	Finance Department	Programmer-Finance Dept.	Need some time
7	Last 6 months purchased items	Purchased Department	Programmer-Purchased Dept.	Call a Meeting
8	Project expenses	Research Department	Programmer-Research Dept.	Call a Meeting
9	Webmail problem	Website Controller	Web Developers	Replied on the spot
10	User Id problem	Website Controller	Web Developers	Declined
11	Salaries payroll problem	Salary Department	Programmer-Salary Dept.	Call a Meeting
12	Keeping record of residential permit - (Aqaama)	Passport Department	Programmer-Passport Dept.	Need some time
13	Passport keeping record	Passport Department	Programmer-Passport Dept.	Replied on the spot
14	Login problem in the KSU website	Website Controller	Web Developers	Call a Meeting
15	Online course registration form	Admission Department	Programmer-Admission Dept.	Replied on the spot
16	File uploading and download problem	University Database Controller	Database Administrator	Need some time
17	Mail server problem	Website Controller	Web Developers	Replied on the spot
18	Faculty members Website maintenance problem	Website Controller	Web Developers	Replied on the spot
19	University magazine	Publishing Department	Programmer-Publishing Dept.	Need some time
20	University calendar	Publishing Department	Programmer-Publishing Dept.	Call a Meeting
21	Colleges portals	College's Web Developer	Web Developers	Replied on the spot
22	Information for international students	Research Department	Programmer-Research Dept.	Replied on the spot
23	Information sharing problems	Database Administrator	Database Administrator	Call a Meeting
24	Online transferring of salaries to bank accounts	Salary Department	Programmer-Salary Dept.	Declined
25	Percentage increment control in salaries	Salary Department	Programmer-Salary Dept.	Need some time
26	New employee entry problem	Employees Department	Programmer-Employee Dept.	Call a Meeting
27	Employee's ID exceed from previous limit	Employees Department	Programmer-Employee Dept.	Need some time

Fig. 2- Program Interface   Fig. 2- Data Gathering

| S.No.	Cutomer's Query	Queried Department	Forwarded to: Concerned Deparment	Actions of the Department	Value
1	Bugs in Software	Software Maintenance	Debuggers	Declined	4
2	Salary Calculation mistakes	Salary Department	Programmer-Salary Dept.	Need some time	2
3	Changes in contracts for new employee	Employees Department	Programmer-Employee Dept.	Call a Meeting	3
4	Salary structure for each designation	Salary Department	Programmer-Salary Dept.	Replied on the spot	1
5	Budget approval for coming year	Finance Department	Programmer-Finance Dept.	Declined	4
6	Changes in budget in current year	Finance Department	Programmer-Finance Dept.	Need some time	2
7	Last 6 months purchased items	Purchased Department	Programmer-Purchased Dept.	Call a Meeting	3
8	Project expenses	Research Department	Programmer-Research Dept.	Call a Meeting	3
9	Webmail problem	Website Controller	Web Developers	Replied on the spot	1
10	User Id problem	Website Controller	Web Developers	Declined	4
11	Salaries payroll problem	Salary Department	Programmer-Salary Dept.	Call a Meeting	3
12	Keeping record of residential permit - (Aqaama)	Passport Department	Programmer-Passport Dept.	Need some time	2
13	Passport keeping record	Passport Department	Programmer-Passport Dept.	Replied on the spot	1
14	Login problem in the KSU website	Website Controller	Web Developers	Call a Meeting	3
15	Online course registration form	Admission Department	Programmer-Admission Dept.	Replied on the spot	1
16	File uploading and download problem	University Database Controller	Database Administrator	Need some time	2
17	Mail server problem	Website Controller	Web Developers	Replied on the spot	1
18	Faculty members Website maintenance problem	Website Controller	Web Developers	Replied on the spot	1
19	University magazine	Publishing Department	Programmer-Publishing Dept.	Need some time	2
20	University calendar	Publishing Department	Programmer-Publishing Dept.	Call a Meeting	3
21	Colleges portals	College's Web Developer	Web Developers	Replied on the spot	1
22	Information for international students	Research Department	Programmer-Research Dept.	Replied on the spot	1
23	Information sharing problems	Database Administrator	Database Administrator	Call a Meeting	3
24	Online transferring of salaries to bank accounts	Salary Department	Programmer-Salary Dept.	Declined	4
25	Percentage increment control in salaries	Salary Department	Programmer-Salary Dept.	Need some time	2
26	New employee entry problem	Employees Department	Programmer-Employee Dept.	Call a Meeting	3
27	Employee's ID exceed from previous limit	Employees Department	Programmer-Employee Dept.	Need some time	2

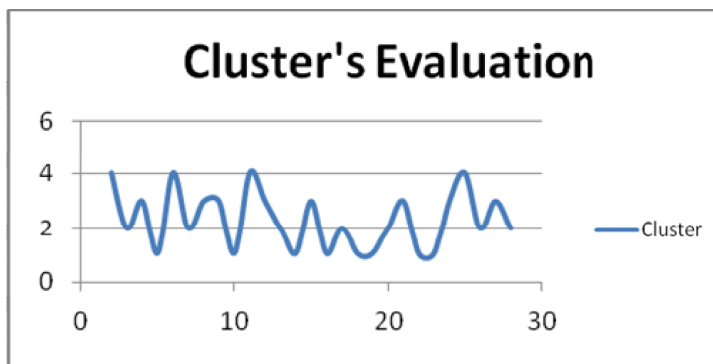Fig. 3- Data Enhancement
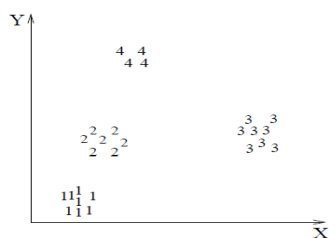
Fig. 4- Data Clusters

Fig. 5- Cluster's Evaluation Graph

C. Apriori Algorithm

Apriori is a classic algorithm for learning association rules. Apriori is designed to operate on databases containing transactions (for example, collections of items bought by customers, or details of a website frequentation) [7, 8]. As is common in association rule mining, given a set of itemsets (for instance, sets of retail transactions, each listing individual items purchased), the algorithm attempts to find subsets which are common to at least a minimum number of the itemsets. Apriori uses a "bottom up" approach, where frequent subsets are extended one item at a time (a step known as candidate generation), and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found [8, 9, 11]. One way to construct a simpler model computed from data, easier to understand and with more predictive power is to create a set of simplified rules [11]. Apriori Algorithm is suitable to compute the rules and patterns and predict for any organization to improve the customer satisfaction. We implement Apriori algorithm on MADAR data and generated some rules and patterns for MADAR.

C.1. Apriori Implementation

According to the Fig. 4, four clusters have been generated on the basis of selection criteria. These clusters we took as an input for the Apriori implementation to generate new rules. This is called data mining on mined data because we used two best data mining algorithm for acquiring the best solution and rules. First clustering and then on clustered data we employed Apriori algorithm on it. For the Apriori implementation we compared all the clustered with each other to find out the subset of all clusters shown in Fig.6.
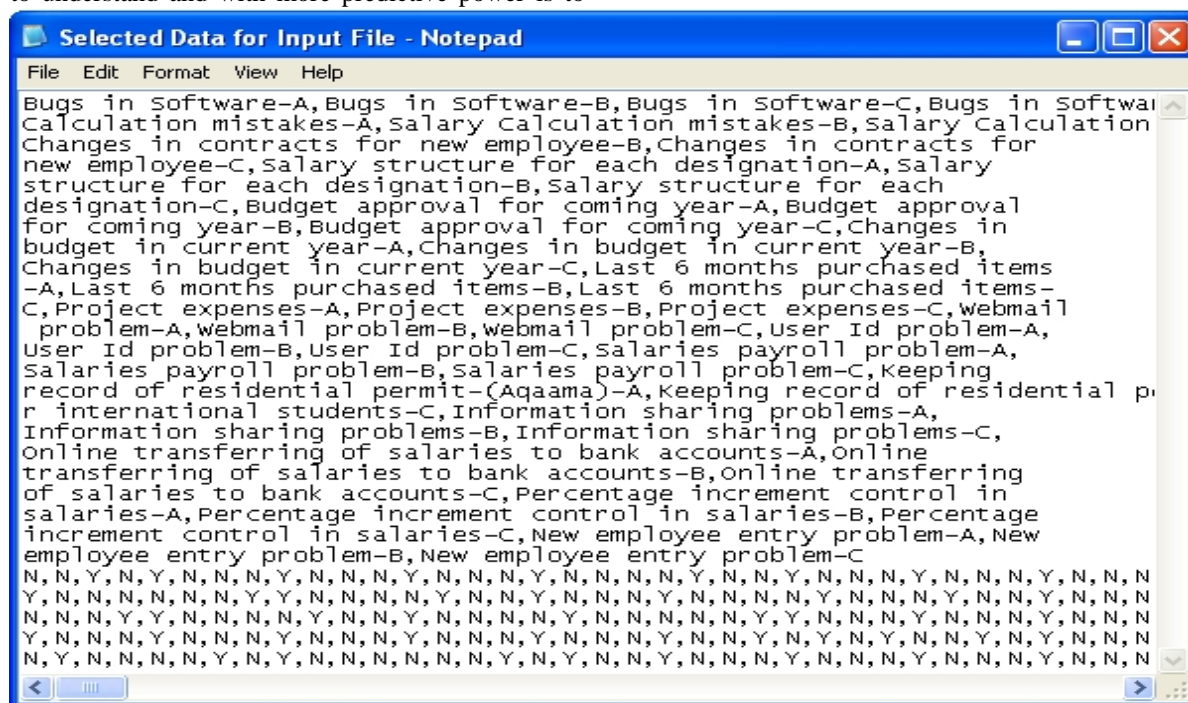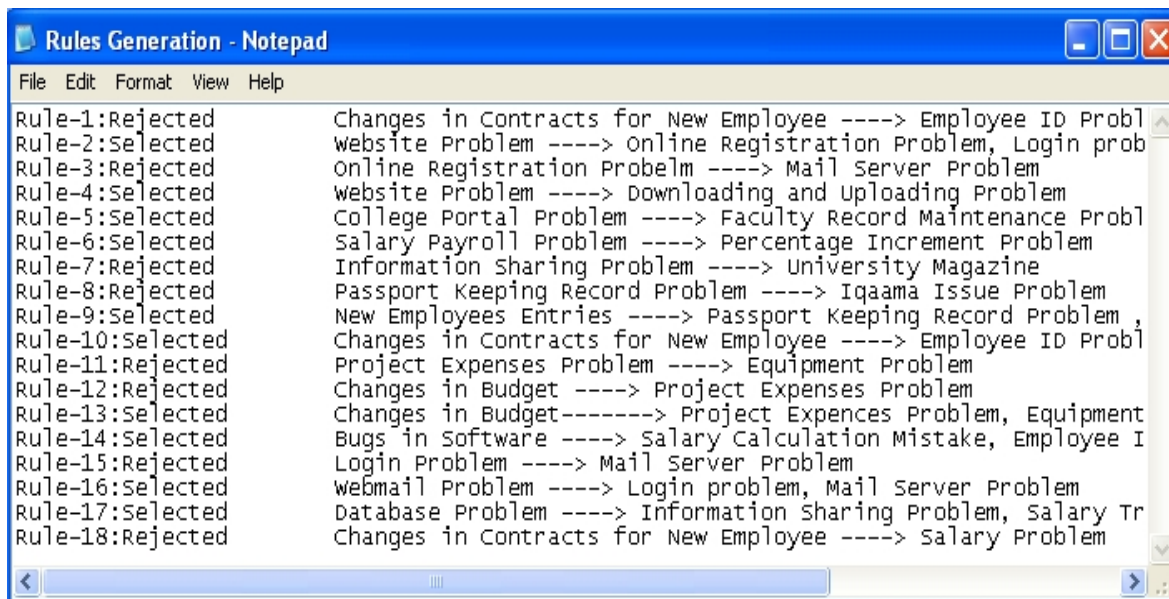

Fig. 6- Input File for Apriori

D. Rules Generation

We applied Apriori algorithm for the generation of frequent itemsets. These Itemset generation is based on the customer query, corresponding binary value, minimum support (5%) and minimum confidence (70%). Fig.7 is shown the number of rules generated by running the Apriori algorithm. This was a sample implementation of Apriori algorithm on the presented model. We can find different kind of frequent itemsets generation in the same manner by increasing or decreasing the minimum support and confidence. The selected rules then proposed and suggested to the MADAR.

IV. FUTURE WORK

The model will be enhancing with more information and more data mining technique will be applying and new rules will be generated in future for the enhancement of an organization. New ERP tools will be used to modify the existing work and make ease for the customers to access the organization's facilities with out any hesitation.

V. CONCLUSION

In today's technologies the customers have a lot of difficulties to access the organization's facilities. The customer have problem in contacting the organization

The model presented in this paper will solve these problems all the customers complaints will be recording in the central database and will be process according to the customer need. The customer can easily contact the organization and can purchase the organization products very easily. The CRM-Layer will collect the information about the products and the queries will be forwarded to the ERP-Layer to act upon these queries. The knowledge discovery layer generates new rules and patterns for the betterment of an organization for future correspondence to improve the growth of the customers for an organization.

REFERENCES

[1] Wen-Hsiung Wu, Chin-Fu Ho, Hsin-Pin Fu, Tien-Hsiang Chang, "SMES IMPLEMENTING AN INDUSTRY SPECIFIC ERP MODEL USING A CASE STUDY APPROACH", Journal of the Chinese Institute of Industrial Engineers, Vol. 23, No. 5, 2006, pp. 423-434.
[2] Sonja Grabner-Kraeuter, Gernot Moedritscher, Martin Waigunyc, Werner Mussnigb, "Performance Monitoring of CRM Initiatives", in Proceedings: IEEE conference on System Sciences, 2007.
[3] Ryals, L. and Knox, S. (2001) "Cross-Functional Issues in the Implementation of Relationship Marketing Through Customer Relationship Management", Cranfield University School of Management.
[4] Tinsley, Dillard B. (2002) "Relationshipmarketing s strategic array", Stephen F. Austin State University, Nacogdoches, Texas.
[5] Stanley B. (2000), Customer RelationshipManagement , John Wiley & Sons Canada Ltd.
[6] Virgil Chichernea, Romanian, "THE USE OF THE ERP-CRM-CIM SYSTEMS WITHIN THE MASTER'S DEGREE PROGRAMMES"
[7] Executive guidance, "http://www.netessence. com.cy/downloads/erp.pdf", Accessed date: March 21, 2009.
[8] Apriori Algorithm, http://en.wikipedia.Org /wiki/Apriori algorithm #Algorithm,Accessed Date: 12th April, 2009.
[9] Agrawal R, Imielinski T, Swami AN. "Mining Association Rules between Sets of Items in Large Databases." SIGMOD. June 1993, 22(2):207-16.
[10] Mahdi Seify, "New Method for Risk Management in CRM Security Management", in proceedings: IEEE conference on Information Technology, 2006.
[11] S. Kotsiantis, D. Kanellopoulos, Association Rules Mining: A Recent Overview, GESTS International Transactions on Computer Science and Engineering, Vol.32 (1), 2006, pp. 71-82.
[12] Raza Ali, Usman Ghani, Aasim Saeed, "Data Clustering and Its Applications", http:// members.tripod.com/asimsaeed/paper.htm, Accessed Date, 15th April, 2009.
[13] Anil K. Jain , "Data Clustering: 50 Years Beyond K-Means", 19th International Conference Pattern Recognition (ICPR), Tampa, FL, December 8, 2008.