Selective Marketing for Retailers to promote Stock using improved Ant Colony Algorithm

S.SURIYA [#1], Dr.S.P.SHANTHARAJAH [#2]

[#1] Assistant Professor, Department of Computer Science and Engineering,
Velammal College of Engineering and Technology, Madurai,
Anna University, Chennai, India
[1] suriyas84@gmail.com

[#2] Professor. Department of Master of Computer Applications,
Sona College of Technology, Salem,
Anna University, Chennai, India.
[2] spshantharaj@gmail.com

Abstract— **Data mining is a knowledge discovery process which deals with analysing large storage of data in order to identify the relevant data. It is a powerful tool to uncover relationships within the data. Association rule mining is an important data mining model to mine frequent items in huge repository of data. It frames out association rules with the help of minimum support and confidence value which in turns paves way to identify the occurrence of frequent item sets. Frequent pattern mining starts from analysis of customers buying habits. From which various associations between the different items that the customers purchase are identified. With the help of such associations retailers perform selective marketing to promote their business. Biologically inspired algorithms have their process observed in nature as their origin. The best feature of Ant colony algorithm, which is a bio inspired algorithm based on the behaviour of natural ant colonies, is its parallel search over the problem data and previously obtained results from it. Dynamic memory management is done by pheromone updating operation. During each cycle, solutions are constructed by evaluation of the transition probability through pheromone level modification. An improved pheromone updating rule is used to find out all the frequent items. The proposed approach was tested using MATLAB along with WEKA toolkit. The experimental results prove that the stigmeric communication of improved ant colony algorithm helps in mining the frequent items faster and effectively than the existing algorithms.**

Keywords - association rule mining, minimum support, confidence, selective marketing, biologically inspired algorithm, stigmergic communication, pheromone updating, transition probability.

I. Introduction

Data mining is a novel process of extracting previously unknown but potentially useful information from large repository of data [1]. It deals with mining of knowledge from huge data storages. It is also known as one of vital processes of Knowledge Discovery in Databases (KDD) [2]. The KDD process (shown in Figure1) involves three phase's namely (i) pre-processing, (ii) data mining process and (iii) post-processing [1]. The pre-processing phase is carried out before applying data mining techniques to sample datasets. It involves cleaning and integrating of databases. As data is brought together from different data sources, it is necessary to remove the noisy data, inconsistent data and duplicate data. Later integration of data from various data sources is done. The data mining process is the second and major phase of KDD [14]. It deals with selection of relevant data from the integrated sources. Later those data are transformed into a convenient format for mining. Various data mining techniques can be carried out over that format of data. Data mining technique makes use of various algorithms to extract data from multidimensional databases. Few examples are Apriori algorithm for association rules, neural network algorithm for prediction, classification and clustering, regression algorithm for classification and prediction. The last phase post-processing deals with evaluation and representation of results obtained from second phase. It completely deals with visualization of results in one of the following forms namely tables, data cubes, charts, decision trees, etc.

Data mining strategies were classified into 3 categories namely descriptive modeling, predictive modeling and market basket analysis. Predictive modeling approach makes use of input attributes to determine and predict output attributes. It is further classified into 3 sub categories like classification, estimation and prediction. Descriptive modeling approach identifies the best set of attributes and evaluates the performance of the model with the help of input attributes. Market Basket analysis determines the interesting relationship among the stocks that promotes the business. A pattern is said to be a frequent pattern or frequent item set only if has a frequency either greater than or equal to a user specified count. For example, a person who buys biscuits tends to buy cool drinks or fruit juices i.e., {biscuits, cool drinks}, {biscuits, fruit juices}. Similarly a person who buys a pencil tends to buy a sharpener later a rubber then we call it to be a sequential frequent pattern. These frequent patterns contribute towards identifying hidden interesting relationships among the data. Therefore frequent item set mining [18] has occupied a vital place in the research works of data mining.

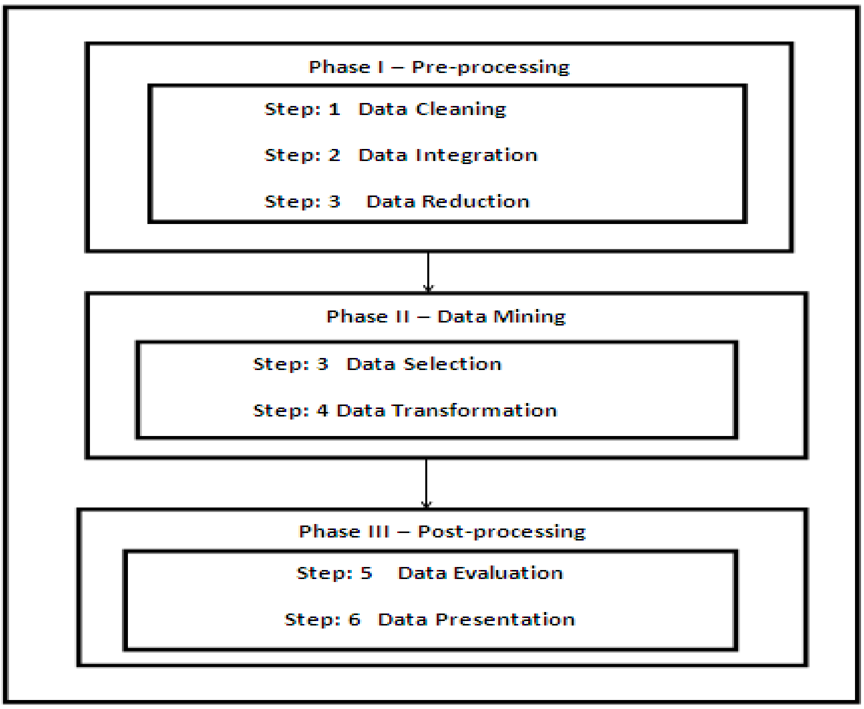

Figure: 1 Knowledge Discovery in Databases

Lemma: let T be a transaction database and S be a set of items {SI1, SI2, SI3… SIn} with 'n' number of items. The k-item set with frequency § is said to be a *frequent pattern* only if § >µ|T| where |T| is the total number of transactions in T.

In 1993, Agrawal [15] introduced association rule mining as one of most important techniques of data mining for point of scale (POS) systems in supermarkets. The main intention of association rule mining is to extract interesting pattern of data from huge data repositories [3]. A rule is defined as an implication of the form A=>B where A∩ B≠Ǿ. The left-hand side of the rule is called as antecedent. The right-hand side of the rule is called as consequent. For example [2,17] the rule { Onions, Potatoes}=>{beef} found in the sales data of a supermarket would indicate that if a customer buys Onions and potatoes together then the customer is likely to buy beef also. Such information is useful to make decisions about marketing activities. Association rules are also used in many applications including Web usage Mining, Intrusion Detection and Bio-informatics. I = {i1, i2, i3, …, im} is a collection of items. T be a collection of transactions associated with the items. Every transaction has an identifier TID [4]. Association rule A=>B is such that AЄI, BЄI. A is called as Premise and B is called as Conclusion. The support, S, is defined as the proportion of transactions in the data set which contains the item set. The confidence is defined as a conditional probability

Support(X=>Y) = Support (XUY) = P (XUY)

Confidence(X=>Y) = Support (XUY) / Support (X) = P(Y/X)

Lemma: The occurrence of a pattern1 { SI1,SI2, SI3,…, SIj } assures the occurrence of another pattern2 { SIj+1,…, SIk } i.e., P({ SI1,SI2, SI3,…, SIj } | { SIj+1,…, SIk }) can be determined as follows based on the scenario: the knowledge about { SIj+1,…, SIk } may make the occurrence of

{SI1, SI2, SI3… SIj}

(i)more likely (i.e.,*P({ SI1,SI2, SI3,…, SIj } | { SIj+1,…, SIk })>P({ SI1,SI2, SI3,…, SIj })*),

(ii)less likely (i.e., *P({ SI1,SI2, SI3,…, SIj } | { SIj+1,…, SIk })<P({ SI1,SI2, SI3,…, SIj })*), or

(iii)have no effect (i.e.,*P({ SI1,SI2, SI3,…, SIj } | { SIj+1,…, SIk }) =P({ SI1,SI2, SI3,…, SIj })*).

Transaction Identity	Item sets
1	i1,i2,i5,i8
2	i1,i3,i5,i6,i9
3	i2,i3, i7,im
….	……..
m	i1,i2,i7,i15

Table: 1 Transaction Database Table

Two popular problem solving strategies are (1) Exact method – deals with logical or mathematical manner of solving problems (2) Heuristic approach [21] – deals with complex optimization problems which cannot be solved by traditional methods with acceptable time and space complexity. Bio-Inspired algorithms are best examples of heuristic approach. Biologically inspired algorithms has routed out from the natural behavior of organisms. It acts as source of inspiration for development of problem solving techniques. Bio-inspired computing has emerged as a new era for of providing computational solutions to complex problems in data mining [22]. The motivation to study such algorithms is based on the innovative bio-inspired solutions [5] for complex problems which cannot be done by traditional methodologies. These heuristic approaches play a vital role in various applications of web mining [23]. Categories of such bio inspired computing (inspired by nature) are (i) Swarm Intelligence (ii) Artificial Immune Systems (iii) Evolutionary Computation and (iv) Neural networks. The Swarm Intelligence includes (1) Ant Colony (2) Particle Swarm and (3) Bee algorithm.

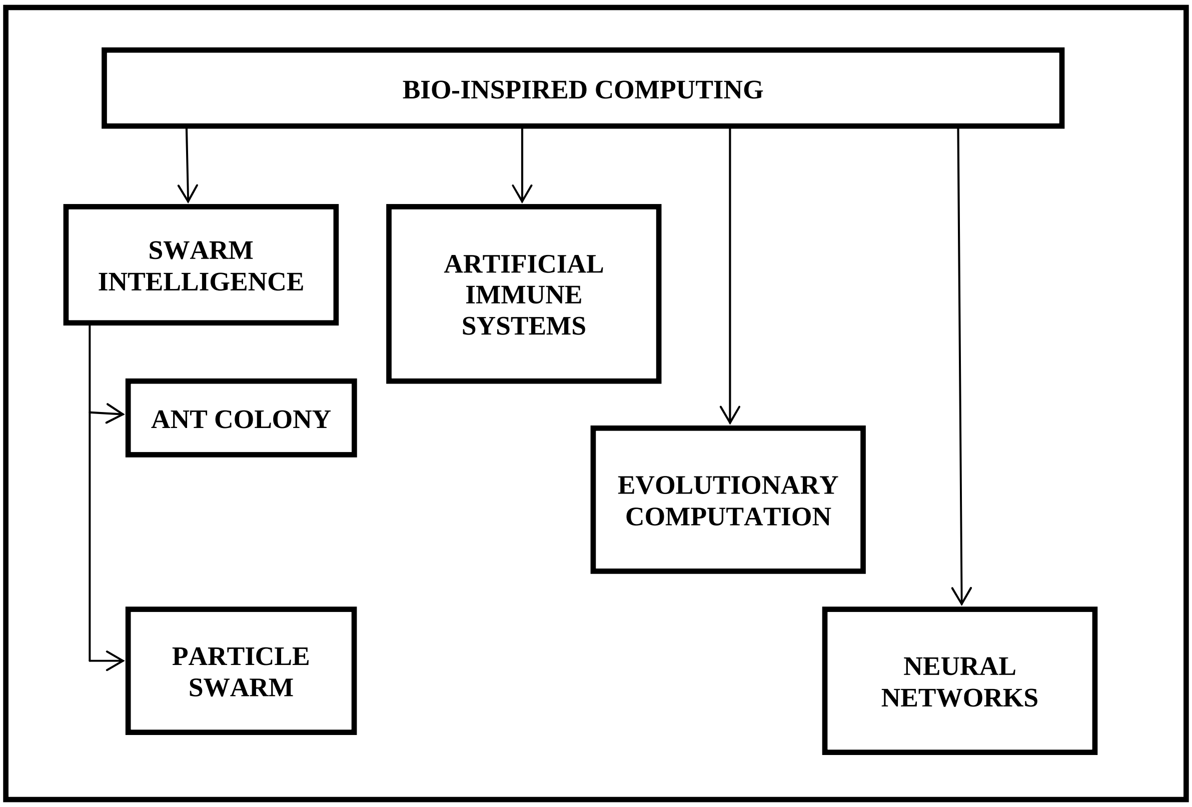

Figure: 2 Categories of Bio-Inspired Computing

This paper has major focus on Ant Colony. The source of inspiration of Ant colony algorithm is foraging behavior of natural ant colonies. It was introduced in 1990 as an optimization technique which is popularly termed as Ant Colony Optimization (ACO) [6]. The attractive feature of ACO is its parallel search over the sample data with the help of dynamic memory structure which stores the previous obtained good results. Ant-based models [7] are applied in accordance to data mining context to preform clustering, sorting, topographic mapping, etc. This algorithm is capable of adapting (at run-time) to the dynamic underlying environment [8]. The modified pheromone updating rule of ant colony algorithm proposed in my previous research paper [9] drives faster than the basic ant colony algorithm. The proposed approach will make use of this modified ant colony algorithm to generate new association rules for the dynamic changing database which consists of set of transactions along with different collection of items.

II. Association Rule Mining

Association Rule Mining is a process of framing out association rules that ensures the predefined minimum support and confidence for any given dataset [2, 19]. It solves the given task using two strategies [10, 15, 17, and 18]. First strategy deals with identifying frequent items in the given task. It is done in two steps namely: (i) Candidate Generation process and (ii) Frequent item generation process. Second strategy involves generation of association rules for the identified frequent items with the help of confidence value. The two basic parameters [15] (Agrawal et al., 1993) related to association rules are support and confidence. A set of items is popularly known as an item set. Support of an association rule is defined as the percentage of records that contain combination of certain item sets to the total number of transactions in the database. Confidence is often referred as measure of strength of the association rules. Confidence of an association rule is defined as the percentage of number of transactions that contain the combination of certain item sets to the total number of records that contain one subset of combination of those item sets. Consider the following sample dataset.

TID	Item sets
T1	i1,i2,i5
T2	i2,i4
T3	i2,i3,i5
T4	i1,i2,i3
T5	i1,i3,i4
T6	i2,i3,i4
T7	i1,i3,i5
T8	i1,i2,i3,i5
T9	i1,i2,i5
T10	i1,i2

Table: 2 Sample Dataset

The association rule mining [24] provides two powerful measures for evaluating the association between items which is expressed by a rule. The confidence of a rule measures the degree of the correlation among items, while the support of a rule measures the significance of the correlation among items.

AIS Algorithm

AIS (Agrawal, Imielinski, Swami) algorithm [11] was first proposed algorithm for mining association rules in 1993. The major drawbacks of this algorithm are high memory space requirements; more number of passes over the sample database, too many candidate item sets that finally turned out to be small was generated. For the above dataset in Table 2,

AIS Algorithm:

Item set	Count
i1	7
i2	8
i3	6
i4	3
i5	5

Item set	Count
i1,i2	5
i1,i3	3
i1,i4	1
i1,i5	4
i2,i3	4
i2,i4	2
i2,i5	4
i3,i4	2
i3,i5	3

Item set	Count
i1,i2,i3	2
i1,i2,i5	3
i1i3,,i4	1
i1i3,,i5	2
i2,i3,i4	1
i2,i3,i5	2

Large 1 item set
i1
i2
i3
i5

Large 2 item set
i1,i2
i1,i5
i2,i3
i2,i5
i1,i3
i3,i5

Large 3 item set
i1,i2,i5
i1,i2,i3
i1,i3,i5
i2,i3,i5

APRIORI Algorithm

Apriori algorithm was most efficient association rule mining algorithm of 1990's. It was proposed by Agrawal in 1994 [1]. It was able to overcome the drawbacks of the previous AIS algorithm. The best features of apriori algorithm are efficient candidate generation process and usage of pruning techniques to avoid measuring non frequent item sets. But still, the two major bottlenecks [10] of this algorithm are usage of most of the time, space and memory for candidate generation process and multiple scans over the sample database. For the above dataset in Table 2,

APRIORI Algorithm: (Minimum Support:3)

Item set	Count
i1	7
i2	8
i3	6
i4	3
i5	5

Large 1 item set
i1
i2
i3
i4
i5

Item set	Count
i1,i2	5
i1,i3	3
i1,i5	4
i2,i3	4
i2,i5	4
i3,i4	2
i3,i5	3

Large 2 item set
i1,i2
i1,i5
i2,i3
i2,i5
i1,i3
i3,i5

Item set	Count
i1,i2,i3	2
i1,i2,i5	3
i1i3,,i5	2
i2,i3,i5	2

Large 3 item set
i1,i2,i5

FP-Tree Algorithm

FP-Tree (Frequent Pattern) Mining algorithm was introduced by Han in 2000. It works comparatively faster than apriori algorithm. FP-Tree is a prefix tree structure algorithm which stores information related to frequent item sets. It involves two steps namely construction of FP-Tree and generation of frequent patterns from the tree. The major drawbacks of this algorithm are lack of interactive mining system (where users don't have the freedom of changing the support values) and not suitable for incremental mining environment. For the above dataset in Table 2,

FP-Tree Algorithm: (Minimum Support:3)

Item set	Count
i1	7
i2	8
i3	6
i4	3
i5	5

Large 1 item set
i2
i1
i3
i5
i4

Item set	Count
i2,i1	5
i1,i3	3
i1,i5	4
i2,i3	4
i2,i5	4
i3,i5	3

Large 2 item set
i2,i1
i1,i5
i2,i3
i2,i5
i1,i3
i3,i5

Large 3 item set
i2,i1,i5

III. Improved Ant Colony Algorithm

Basic Ant Colony Algorithm: Ant Colony algorithm [12] was first proposed in 1990's, by Marco Dorigo for his research activities to solve optimization problems effectively. It is the most successful swarm based algorithm with attractive swarm principles [13] namely simple space and time requirements, fast response to the environment and adaptability towards the environment. This algorithm emerged out with its inspiration from activities carried out by the natural ant colonies. It is universally observed that whenever ant moves in colonies to search food source, they lay a chemical substance on the ground. It is termed as pheromone. This substance has the characteristic of getting evaporated within certain period of time. It acts as a source of indirect communication for the other fellow ants following it. This type of indirect communication is called as stigmeric communication. During a random selection of paths at initial stage from source to destination, ants lay pheromone and proceed in their own way. But the fellow ants tend to follow the path which has pheromone (which must obviously be the shortest route).

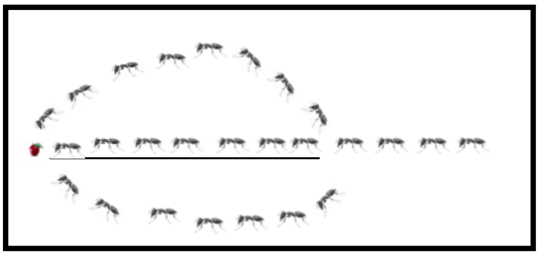

Figure: 3 Initial Stage of search for food source

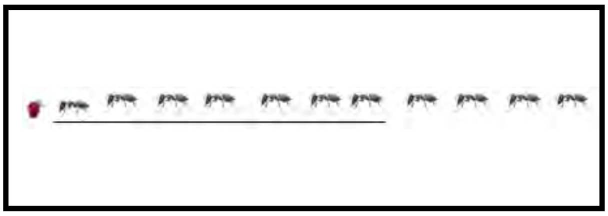

Figure: 4 Final Stage of search for food source

Consider the above example (in figure 3), where a group of ant colonies are observed to move in search for food source. At Initial stage of search, they randomly move in possible directions. As they move, pheromone is laid on the ground on all the paths wherever they travel. After some time, it is observed that the pheromone on the other paths except the shortest path gets evaporated as shown in figure 4. Thus the fellow ants follow the shortest path with the help of pheromone laid on that path. This algorithm has been adopted towards various applications like task scheduling, vehicle routing, grid computing, travelling salesman problem, etc. The main aim of this paper is to investigate the ant colony algorithm technique for mining frequent item sets in the given sample database. The proposed technique is equipped with a colony of ants with initial minimum support value. Using this knowledge as guidance, they proceed towards mining of frequent items. The sample database consists of number of transactions. Each transaction has a collection of items. Any existing algorithm like AIS, Apriori, FP-Tree can identify frequent item sets and the interesting rules. But the execution time differs when the transactions become dynamic in nature. The pheromone trail is updated in order to allow ants to share information related to good solutions. At the end of every iteration, only the best ant is allowed to leave the pheromone. It is assumed to be the best ant solution explored so far. The pheromone level is decremented to make the ants forget about the poor solutions. It is indicated using the parameter ρ. The value of ρ lies between 0 and 1. If the value of ρ is 1 then there is no idea of decrement. If the value of ρ is 0 then the pheromone trail will be switched off. Pheromone updating rule is given by:

$\tau ij(t)new = [\ \tau ij(t)old\ ] + [\ \{\rho\} * \Delta\ \tau ij(t)\ ]$ ---(1)

Where $\tau ij(t)$ denotes trail intensity of the edge(i,j), ρ denotes evaporation rate and $\Delta\tau ij(t)$ denotes additional pheromone when ant moves from one transaction to another.Using information stored in the pheromone trail and the heuristic information (related to the transactions), the ant starts to mine the database. Each ant starts with the minimum support value and the transaction list (TID), which consists of collection of transactions along with their own collection of item sets as shown in table2. Each transaction T_i is then probabilistically chosen to mine next based on the pheromone value. the probability of selecting an item to be included for frequently occurring item is generated using the equation (2). Equation (2) consists of two parameters namely α and β. The parameter α indicates the relative importance given to pheromone and the parameter β indicates the relative importance given to heuristic information. The probability selection is given as

$P_{ij}(t)^k = [\tau_{ij}(t)]^{\alpha}*[\eta_{ij}(t)]^{\beta}/\Sigma_{u\in allowed(k)} [\tau_{iu}(t)]^{\alpha}*[\eta_{iu}(t)]^{\beta}$ ………….. (2)

Where $P_{ij}(t)$ denotes the probability to move along the path (i to j), $\tau_{ij}(t)$ denotes the trail intensity of the edge(i,j) and $\eta_{ij}(t)$ denotes the visibility (1 / $distance_{ij}$).At the end of each iteration, the identified item satisfying the minimum support value is allocated to the best selected ant. This process is repeated until all the transactions in the list are traversed and a complete solution is built. Every ant in the system follows the same manner to build the solution. The pheromone trail is updated after all the ants build a solution. Ants make use of the pheromone value (minimum support value) in identifying the item which is frequently occurring in many transactions. According to the dynamic change in the transactions in the database and also to change in minimum support value, ants generate the frequently occurring items. The pseudo code of basic ant colony algorithm for mining frequent item set is as follows:

```
procedure BasicACO_Mining
Begin
   Initialize the pheromone (minumum support value)
   While lists in the database  is not empty do
       For each transaction in the list do
           Start  each ant to mine for frequent_check
       end for
     Choose next item with the help of state transition rate
     until every ant has build a solution
     Update the pheromone (if required)
   end while
end
```

Figure: 5 Basic Ant colony algorithm for mining frequent item sets

The proposed improved ant colony algorithm has a modified pheromone updating rule. It is shown in equation (3). The complexity of the proposed algorithm can be represented using Θ notation as it helps in identifying the interesting rules very quickly than that of basic algorithm whose complexity can be represented using Ω notation. $\tau_{ij}(t)new=[\{(1-\rho)/(1+\rho)\}*\tau_{ij}(t)old]+[\{\rho/(1+\rho)\}*\Delta\tau_{ij}(t)]$…(3) Where $\tau_{ij}(t)$ denotes the trail intensity of the edge(i,j), ρ denotes evaporation rate and $\Delta\tau_{ij}(t)$ denotes additional pheromone when next transaction is traversed. The pseudo code of improved ant colony algorithm [9] is as follows:

```
procedure ImprovedACO_Mining
Begin
   Initialize the pheromone (minumum support value)
   While lists in the database  is not empty do
       For each transaction in the list do
           Start  each ant to mine for frequent_check
       end for
     Choose next item with the help of state transition rate
       Pij(t)^k = [τ ij(t)]^α*[ηij(t)]^β/Σu∈allowed(k) [τ iu(t)]^α*[ηiu(t)]^β
     until every ant has build a solution
     Update the pheromone
     τij(t)new=[{(1- ρ)/(1+ρ)}*τij(t)old]+[{ρ/(1+ρ)}*Δτij(t)]
   end while
end
```

Figure: 6 Improved Ant colony algorithms for mining frequent item sets

IV. Experimental Results

MATLAB collaborates itself with WEKA for gene selection [25] which is now utilized for frequent itemset selection over Market Basket Analysis. WEKA is well for tool for Machine Learning [26]. The proposed approach is implemented in MATLAB environment along with WEKA tool. Convert weka data, stored in a java weka Instances object to a matlab. It is done with the help of 4 basic input parameters namely Input java instances (from WEKA), Transaction list with their collection of items, Minimum support value and Frequent item. Preprocessing is carried out using matlab code. The item selection is done by selector using trainWekaAttrSelector function. Later training, validating and testing are done using useFilter function. Finally non frequent items are filtered out. The table 3 shows the number of transactions considered for every cycle of execution. The results are tabulated for intervals of every 50 transactions starting from 50 transactions to 950 transactions respectively. Table 3 shows the experimental results of Basic Ant colony algorithm and Improved Ant colony algorithm for mining frequent items.

The sample dataset for 5 transactions are:

T1.	i1,i2,i3,i4,o,o,i7,o,o,i10
T2.	i1,i2,i3,i4,o,o,o,i8,o,i10
T3.	i1,i2,i3,4,o,o,o,o,i9,i10
T4.	i1,i2,i3,4,o,o,o,i8,i9,i10
T5.	i1,i2,i3,4,o,o,i7,o,i9,i10

	Percentage of Execution Speed for Mining Frequent Item sets	
No: of Transactions	**Basic Ant colony algorithm**	**Improved Ant Colony algorithm**
50	0.25	0.4
100	0.28	0.42
150	0.29	0.44
200	0.3	0.45
250	0.32	0.47
300	0.34	0.49
350	0.35	0.54
400	0.42	0.57
450	0.44	0.63
500	0.46	0.68
550	0.49	0.72
600	0.52	0.74
650	0.56	0.77
700	0.57	0.79
750	0.62	0.8
800	0.67	0.84
850	0.69	0.87
900	0.7	0.9
950	0.71	0.92

Table: 3 Basic Ant colony algorithm VS Improved Ant colony algorithm for mining frequent items

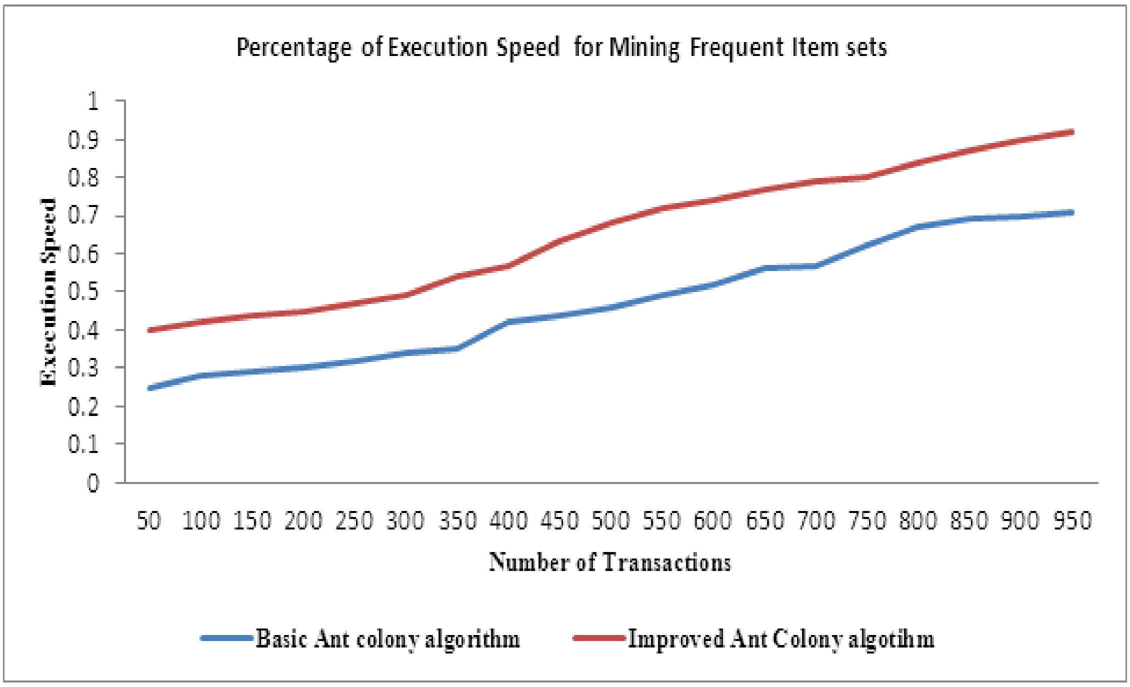

Figure: 7 Performance of Basic Ant colony algorithm VS Improved Ant colony algorithm for mining frequent item sets

The Figure 7 visualizes that the improved ant colony algorithm mines faster than the basic ant colony algorithm. The speed of mining every set of transactions (starting from 50 transactions,100 transactions, 150 transactions, 200 transactions, 250 transactions, 300 transactions, 350 transactions, 400 transactions, 450 transactions, 500 transactions, 550 transactions, 600 transactions, 650 transactions, 700 transactions, 750 transactions, 800 transactions, 850 transactions, 900 transactions and finally 950 transactions) is comparatively effective than the basic ant colony algorithm. The table 4 shows the comparative performance of AIS algorithm, Apriori algorithm, FP-Tree algorithm, basic ant colony algorithm and improved ant colony algorithm. The results are tabulated for intervals of every 50 transactions as in case of table 3. Figure 8 make us realize the tremendous performance of proposed algorithm when compared to some existing algorithms for mining frequent item sets.

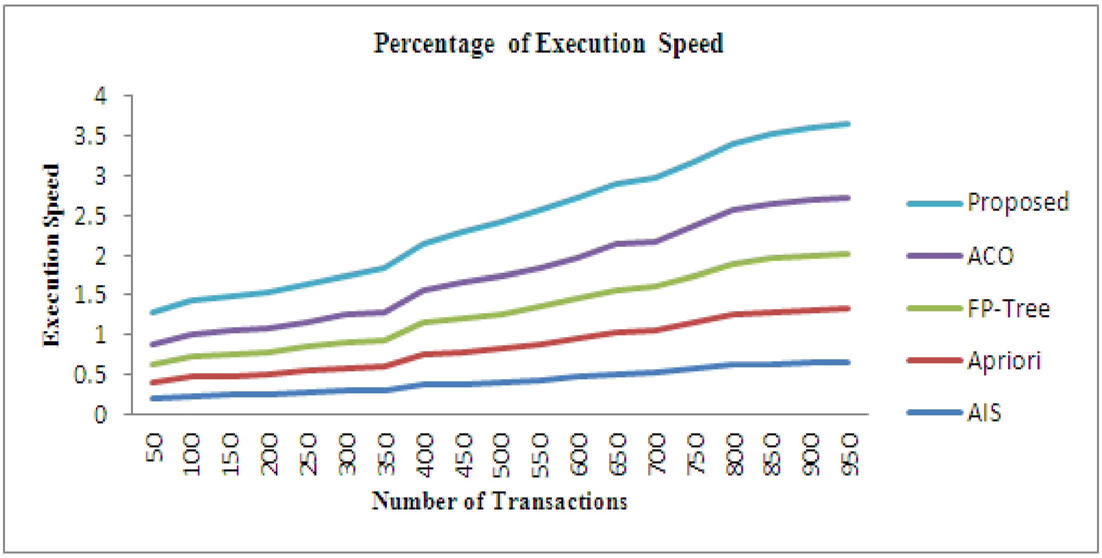

Figure: 8 Performance Rain of Proposed approach

No: of Transactions	Percentage of Execution Speed for Mining Frequent Item sets	
	Basic Ant colony algorithm	Improved Ant Colony algorithm
50	0.25	0.4
100	0.28	0.42
150	0.29	0.44
200	0.3	0.45
250	0.32	0.47
300	0.34	0.49
350	0.35	0.54
400	0.42	0.57
450	0.44	0.63
500	0.46	0.68
550	0.49	0.72
600	0.52	0.74
650	0.56	0.77
700	0.57	0.79
750	0.62	0.8
800	0.67	0.84
850	0.69	0.87
900	0.7	0.9
950	0.71	0.92

Table: 4 Comparison of AIS, Apriori, FP-Tree, ACO and Proposed approach

V. CONCLUSION AND FUTURE WORK

This research paper has presented a fast and effective algorithm for mining frequent item sets over large repository of data. The proposed algorithm yields good results when compared to existing algorithms. It helps out in framing qualitatively interesting association rules. This approach helps the retailers in perform the best selective marketing by identifying the interesting rules i.e the relationship between the items that every customer buys. Thus there is a remarkable improvement in promoting the business by identifying the demands of the customers. Future work deals with merging of this proposed work as hybridization with any one of swarm intelligence algorithms.

REFERENCES

[1] Jaiwei Han and Micheline Kamber, “Data Mining Concepts and Techniques”, Second Edition , Morgan Kaufmann Publishers.
[2] Qiankun Zhao and Sourav S. Bhowmick, “Association Rule Mining: A Survey”, Technical Report, CAIS, Nanyang Technological University, Singapore, No. 2003116, 2003.
[3] Feng Yucai, “Association Rules Incremental Updating Algorithm”, Journal of Software, Sept., 1998.
[4] Lei Guoping, Dai Minlu, Tan Zefu and Wang Yan, “ The Research of CMMB Wireless Network Analysis Based on Data Mining Association Rules”, IEEE conference paper – project supported by the Science and Technology Research Project of Chongqing municipal education commision under contract no KJ101114 and KJ 111103, 2011
[5] Ana Karina F. Prior, Leandro Nunes de Castro, Leandro R. de Freitas, Alexandre Szabo, “The Proposal Of Two Bio-Inspired Algorithms For Text Clustering “,Learning and Nonlinear Models - Revista da Sociedade Brasileira de Redes Neurais (SBRN), Vol. 6, No. 1, pp. 29-43, 2008.
[6] Marco Dorigo, Mauro Birattari, and Thomas Stutzle, “Ant Colony Optimization - Artificial Ants as a Computational Intelligence Technique”, IRIDIA – Technical Report Series, Technical Report No. TR/IRIDIA/2006-023, September 2006.
[7] Handl J, Knowles J, Dorigo M. , “Ant-based clustering and topographic mapping”, Artificial Life, MIT Press Journals, volume 12 , Issue 1 – winter, page no35-62, 2006.
[8] Marco Dorigo, Gianni Di Caro and Luca M. Gambardella, “Ant Algorithms for Discrete Optimization”, Artificial Life, MIT Press Journals, Spring 1999, Vol. 5, No. 2, Pages 137-172, Posted Online March 11, 2006.
[9] P. Mathiyalagan, S. Suriya and S.N. Sivanandam, “Hybrid enhanced ant colony algorithm and enhanced bee colony algorithm for grid scheduling”, International Journal of grid and Utility Computing, volume 2, Issue 1, page no: 45-58, 2011.
[10] Sotiris Kotsiantis, Dimitris Kanellopoulos, “Association Rules Mining: A Recent Overview”, GESTS International Transactions on Computer Science and Engineering, Vol.32 (1), , pp. 71-82, 2006.
[11] Agrawal, R., Imielinski, T., and Swami, A. N., “Mining association rules between sets of items in large databases”,Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, 207-216, 1993.

[12] M. Dorigo. Optimization, Learning and Natural Algorithms (in Italian). PhD thesis, Dipartimento di Elettronica e Informazione, Politecnico di Milano, IT, 1992.
[13] Binitha S, S Siva Sathya, "A Survey of Bio inspired Optimization Algorithms", International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307, Volume-2, Issue-2, May 2012.
[14] Heikki Mannila, Hannu Toivonen, "Levelwise Search and Borders of Theories in knowledge Discovery", International Journal of Data Mining and Knowledge Discovery (Springer), Volume 1, Issue 1, pg: 241 – 258, 1997.
[15] Rakesh Agrawal, Tomasz Imielinski, Arum Swami, "Mining Association Rules between sets of items in large databases", Proceedings of the ACM SIGMOD international conference on Management of data, Volume 22 Issue 2, Pages 207-216, 1993.
[16] Rakesh Agrawal, Ramakrishnan Srikant, "Fast Algorithms for Mining Association Rules" , Proceedings of 20th International Conference on Very Large Data Bases, VLDB, 1994.
[17] "Privacy-Preserving Data Mining: Why, How, and When", IEEE Security & Privacy, published by the IEEE Computer Society, 2004.
[18] Gkoulalas-Divanis," A Hybrid Approach to Frequent Itemset Hiding", 19th IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2007, 29-31 Oct. 2007, Volume 1, Page(s): 297 – 304, 2007.
[19] Chengqi Zhang, Shichao Zhang, "Association rule mining: models and algorithms", Book from ACM DL, Springer-Verlag Berlin, Heidelberg ©2002, ISBN:3-540-43533-6.
[20] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten, "The WEKA data mining software: an update", Published in Newsletter ACM SIGKDD Explorations Newsletter archive, Pages 10-18, Volume 11 Issue 1, June 2009.
[21] D.T.Pham and D. Karaboga, "Intelligent Optimization Techniques – Genetic Algorithms, Tabu Search, Simulated Annealing and Neural Networks", Springer, 2000.
[22] Crina Grosan, Ajith Abraham, Monica Chis, "Swarm Intelligence in Data Mining", Studies in Computational Intelligence Volume 34, page no 1-20,2006.
[23] Andrew Secker, Alex A. Freitas, Jon Timmis, "A Danger Theory Inspired Approach to Web Mining", Artificial Immune Systems, Lecture Notes in Computer Science Volume 2787, page no 156-167, 2003.
[24] Weiyang Lin, Sergio A. Alvarez, Carolina Ruiz, "Efficient Adaptive-Support Association Rule Mining for Recommender Systems", Data Mining and Knowledge Discovery, Volume 6, Issue 1, page no 83-105,January 2002.
[25] Yi Zhang, Dingding Wang, Tao Li, "LIBGS: A MATLAB software package for gene selection", International Journal of Data Mining and Bioinformatics, Inderscience Publishers, Volume 4, Number 3, Pages 348-355, 2010.
[26] Remco R. Bouckaert , Eibe Frank, Mark A. Hall, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten, "WEKA---Experiences with a Java Open-Source Project", The Journal of Machine Learning Research, Volume 11, Pages 2533-2541, 2010.

Biographical notes:

S.Suriya received her Bachelor of Engineering degree in the field of Computer Science from Raja College of Engineering and Technology, TamilNadu, India in 2008 (Anna University Rank Holder – 31st Rank in overall University). She got her M.E degree in the field of Computer Science from PSG College of Technology, TamilNadu, India 2010 (Gold Medalist). She got the Best Project Award for her PG Project at PSG College of technology, Coimbatore. She is working as an Assistant Professor in the Department of Computer Science and Engineering, Velammal College of Engineering and Technology, TamilNadu, India. She is currently carrying out her Doctoral research in frequent item set mining in Data Mining. Her area of interest includes data mining, grid computing and semantic web. She has published seven technical papers in national, international journals and conferences.

Dr.S.P.SHANTHARAJAH Professor and R&D Incharge , working in the deparment of Computer Applications, Sona College of Technology, Salem. His research area is Network Security. He is specialized in Network Security and Database Systems. His publications include 9 Journals and 21 Conferences. He got the Best paper award in "IEEE International conference' held at Srilanka on August 2006. He has acted as Resource person in several roles in different parts of country at various Conferences, Workshops, Seminars and so on, sponsored by UGC, AICTE, CSIR, ISRO, CDAC, ISTE, DIT, NASSCOM, CSI and other organisations. He is also a member in various professional bodies like The Indian Society for Technical Education (ISTE), Computer Society of India (CSI) and Cryptology Research Society of India (CRSI)