# ARAA: A Fast Advanced Reverse Apriori Algorithm for Mining Association Rules in Web Data

Bina Bhandari [*1], Bhaskar Pant[2], R H Goudar[3]
[1] CS/IT Department, Graphic Era Hill University, 510, Society Area, Clement Town Dehradun, India.
[1] kotiyalbina@gmail.com
[2]CS/IT Department, Graphic Era University, Dehradun, India.
[2] pantbhaskar2@gmail.com
[3]Department of CNE, Visvesvaraya Technological University, Belagavi-590018, India.
[3] rhgoudar@gmail.com

***Abstract*— This paper proposed an effective algorithm for mining frequent sequence patterns from the web data by applying association rules based on Apriori, known as Advanced Reverse Apriori Algorithm (ARAA). It also shows the limitation of existing Apriori and Reverse Apriori Algorithm. Our approach is based on the reverse scans. An experimental work is performed that shows that proposed algorithm works better than the existing two algorithms. The advantages of ARAA are that it can deeply reduce the multiple scans for frequent sequence pattern generation which results in less processing overhead. A comparative study performed on all three approaches shows that our algorithm improve the mining process significantly as compared to Apriori and Reverse Apriori based mining algorithms especially for the all database. The advantages of ARAA are reduced execution time and increase throughput.**

**Keyword -** Association Rule, Apriori Algorithm, Reverse Apriori, Web Usage, Frequent Sequence Patterns

## I. INTRODUCTION

Data mining is the course of extracting useful information from large dataset by merging statistical and artificial intelligence methods. It aims at finding interesting correlations, frequent patterns, associations among sets of items [1] in the data sources. Association rule mining has been a topic of research in data mining. In order to find the associations among large set of items in a transaction database, mining algorithm are introduced. The mining process is divided into two phases; first phase discovers the frequent large set of items which are based on counts by scanning the transaction data whereas in next phase association rules are made on the basis of the large sets of item originated in the first phase.

The problem associated with existing Apriori Algorithm and Reverse Apriori Algorithm for association rule mining are discussed in this paper and we have also proposed an algorithm which is based on existing Reverse Apriori Algorithm known as Advance Reverse Apriori Algorithm (ARAA) for finding the sequence patterns in the filtered data set of web log. Sequential pattern mining is 0the procedure of employing data mining techniques to a sequential database for the purposes of uncovering the association that exist among an ordered list of events.

This paper focuses on discovering the frequent sequence patterns from the web data through Apriori Algorithm (AA), Reverse Apriori (RAA) and an Advance Reverse Apriori Algorithm (ARAA) and also performs a comparative study on all the three algorithms in terms of the number of scans required to achieve the frequent sequence patterns. It also justifies why ARAA is better than AA, RAA.

### *A. Problem Statement*

Agrawal and Srikant formerly proposed the problem of Sequence pattern mining [3, 14]. Sequential pattern mining is a problem associated with finding the combinatorial explosive number of intermediate sequences from the large datasets. This paper discusses the problem of sequence pattern mining with respect to web log data.

The problem associated is divided into two sub problems. First problem is to find the item sets whose existences exceed a predefined threshold value in the large database. These item sets are termed frequent or large item sets. This first problem is further divided into two sub-problems.

- Candidate Large item sets generation
- Frequent item sets generation

Out of these we considered the item sets whose support exceeds the user defined threshold value as frequent or large item sets. The second problem of sequence pattern mining is to engender association rules from those large item sets with minimal confidence as constraint. Let I = { i1, i2……. in } be a set of n discrete literals called items, T be a set of transactions (variable length) over I, where each transaction contains an item set of (i1,

i2,…….. ik) $\subseteq$ I. All transactions are associated with an identifier (called TID). An association rule is an allegation of the form A $\Rightarrow$ B, where A, B $\Rightarrow$ I and A $\cap$ B = 0. Here A is known as antecedent and B is known as consequent of the rule. The selection of the rule is based on two important properties that are support and confidence.

## II. WEB USAGE MINING

Web Usage Mining an application of data mining techniques is used to discover interesting usage patterns from web data so as to understand the need of the users. The web data contains web log data, web structure data and user profiles data [2]. Log or Usage data stores the information related to the identity or origin of Web users along with their surfing behavior at a Web site. Web personalization, web prefetching, site reorganization and link prediction are the application areas of web usage mining [4, 5, 9, 15]. The most important phases of web usage mining are data cleaning, user identification, session identification and formatting of the data. Heuristic technique is used for the reconstruction of user sessions [6] and discovering interesting patterns from these sessions in the pattern discovery [7] phase using the techniques such as association rule mining, Apriori [8]. With respect to web usage mining, association rule mining helps to understand the behavior of the users on web, aims to attract their visitors by personalizing the web site as per the need of the users.

## III. ASSOCIATION RULE MINING

In web usage mining, an association rule discovers the correlations between web pages which are visited together during a server session [10]. Association rules shows the potential relationship between pages that are often visited together although they are not directly associated. It can depict the associations between groups of users with precise interests. Association rule mining are used for business applications, web recommendation, personalization and improving the system's performance via predicting and pre-fetching the web data. These rules assist in developing effective marketing strategies for those organizations which are engaged in electronic commerce. According to many researchers Apriori is the best well-known, easy to implement and understand algorithm under association rule for mining the frequent sequence patterns from the large databases. However, later on researchers started working on the limitations of Apriori algorithm and either developed new algorithms or improved the working of existing Apriori algorithm.

## IV. MATERIAL AND METHODOLOGIES

For conducting the experiment we took the filtered data from the access log file. The access log file is generated as an interaction between the client and server. Therefore it is essential to perform the preprocessing of the log data so that the quality data can be obtained.

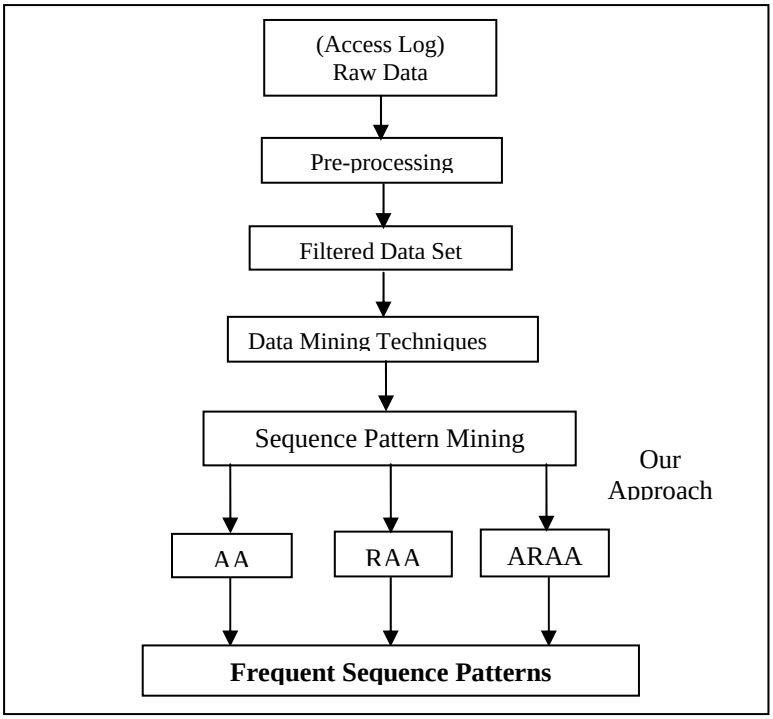


Fig. 1 Proposed System Architecture

This process of extracting the interesting data from the log is known as feature extraction i.e. removing the unwanted data such as redundant data, video, images etc. This feature extracted data is also known as filtered data. Now once the data is filtered then we apply sequence pattern mining algorithm with an extension of three algorithms for discovering the frequent sequence patterns. Fig. 1 shows the proposed system architecture. Fig. 2 shows the data from access file in the tabular format. This dataset is prepared manually into number of sessions. Out of that we have taken one session that contains the eight transactions along with the information of web pages accessed by the users in those transactions.

| Time | User Name | User Group | Domain | URL | Category | IP Address |
|---|---|---|---|---|---|---|
| 2012-11-26 14:37:13 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:36:54 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:36:50 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:36:26 | csdt57 | Facuty_Tech_Staff | urlfilter.vmn.net | urlfilter.vmn.net/vmnsbf/stamp.txt | SearchEngines | 192.168.6.54 |
| 2012-11-26 14:36:19 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:35:59 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:35:58 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:35:57 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:35:52 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:35:51 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:35:23 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:34:54 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | DownloadFreewareAndSh | 192.168.6.54 |
| 2012-11-26 14:34:51 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:34:27 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:33:58 | csdt57 | Facuty_Tech_Staff | mail.google.com | mail.google.com/mail/u/0/?ui=2&ik= | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:33:55 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:33:53 | csdt57 | Facuty_Tech_Staff | tools.google.com | tools.google.com/service/update2?v | InformationTechnology | 192.168.6.54 |
| 2012-11-26 14:33:48 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:33:31 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:32:59 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:32:54 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:32:53 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | Government | 192.168.6.54 |
| 2012-11-26 14:32:33 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:31:56 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:31:49 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:31:39 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:31:33 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |
| 2012-11-26 14:31:02 | csdt57 | Facuty_Tech_Staff | *.mail.google.com | *.mail.google.com | WebBasedEmail | 192.168.6.54 |

Fig. 2 Data from Log File

The first algorithm is an existing Apriori Algorithm, second is Reverse Apriori and the third algorithm is an improved algorithm of Reverse Apriori, we called it Advanced Reverse Apriori Algorithm.

All the algorithms are based on two main steps: Candidate Generation and Pruning.

### *A. Apriori Algorithm*

The Apriori algorithm is one of the well-known association based algorithm for mining the frequent sequence patterns from the web log data. The frequent sequence patterns are the patterns whose value exceeds the minimum defined support and later on it is used for generating the rules. The general outline of the AprioriAlgorithm is as follows:

- Define the support threshold of the one division item sets and discard the rare items.
- Form candidate item sets with an increasing order i.e. of two items then three items and so on (pair items must be frequent), determine their support threshold, and discard the rare occurring item sets. The pseudo code for the AA is given in [11, 13, 14].

### *B. Reverse Apriori Algorithm*

The Reverse Apriori Algorithm is proposed by Kamrul Abedin Tarafteret. al. [12] to find the frequent item sets from the large database. The working of RAA is opposite to that of AA. RAA scans the k- itemset first then if the frequent itemset is not found in the largest scan then it selects the next k to k-i frequent itemset and the process goes on until frequent itemsets are found. RAA works well when the frequent itemsets are found in the beginning but if the itemsets are found in the later phases then the working of Reverse Apriori does not scale well. The pseudo-code for the RAA is given in Fig. 3.

**Reverse Apriori Algorithm**
**Input:**
Web Database wd
Minimum supportminsup
**Output:**
Large Frequent itemsets
**Method**
1. k = Select maximum number of elements
2. i=0
3. For each combination of (k-i) number of elements
4. Do
5. Generate Candidate itemsets of k-i
6. Generate frequent itemsets from candidate itemsets of k-i
7. Where, support in generated itemsets>=minsup
8. If true, then goto step 10
9. Else increment i and goto step 3
10. Return set of large itemsets
11. End

Fig. 3. Pseudo code for Reverse Apriori Algorithm

*C. Advance Reverse Apriori Algorithm*

Advance Reverse Apriori Algorithm is based on association rule mining. It works just opposite to the Apriori algorithm and therefore scans k^thitemset first and then move to the lower level sets. At a particular kth level it only scans k-length attribute only. The scans in ARAA are constant and at each level the number of scans is equal to the number of transactions. If we get frequent set at starting level we can predicts most of the datasets of all its lower level sets. The number of scans in ARAA is almost 50%-60% reduced as that of AA. The pseudo code for ARAA is given in Fig. 4.

**Advance Reverse Apriori Algorithm**
1. i=0, count =0
2. Generate $C_{k-i}$ and $F_{k-i}$ together and make set of TID for each set (Scan only k-i set)
3. $C_{k-i}$>= mini_support
    a. Put in $F_{k-i}$ by performing Union operation according to their TID and delete $C_{k-i}$
4. i=i+1
5. ∀ set, make set combination of $C_{k-i+1}$& $F_{k-i+1}$ of size k-I, for item having count of ($F_k$– 1) >mini_support make combination of that itemset in $C_{k-i}$ and do not put that in $F_{k-i}$combination.
6. Perform union operation of $C_{k-i}$&$F_{k-i}$ of the TID's and increment their respective count.
7. Perform union of item in $F_{k-i}$which is present in both $C_{k-i}$ and $F_{k-i}$ and delete that from $C_{k-i}$
8. Repeat step 2 till $C_1$

Fig. 4. Pseudo code for Advance Reverse Apriori Algorithm

## V. EXPERIMENT AND RESULTS

In the experimental work we have taken a transaction and each transaction contains some sequences. The sequences are represented with the alias for the ease of writing and understanding it in the datasets. The table uses the following alias names as follows:

- Web : Normal Navigation
- IT : Information Technology Related Websites
- SE : Various Search Engines Navigated
- EDU : Educational Site

- Down : Downloaded Sites
- Govt : Government Organization Related Sites

Table 1 shows the sequences that are generated from the filtered dataset along with the eight transactions. In this paper we are going to apply the algorithms on this dataset and based on that we will perform the comparison between the existing algorithms and our proposed algorithm. In table 2 the sequences are represented with the items for the ease of writing. Finally table 3 shows the transactions that contain the sequences in the itemsets form.

TABLE I. Sequences in a Transaction

| Transactions | Category |
|---|---|
| T1 | {Web, IT, SE} {IT} {IT, Web, Edu} |
| T2 | {Down, Govt, IT} {IT, Govt} {Web, IT, SE} |
| T3 | {Edu, Down} {Edu, IT, Web} {Web, IT SE, Down} |
| T4 | {IT, Web, Edu} {IT, Govt, Down} |
| T5 | {Web, IT, SE} {IT, Govt} {Edu, IT, Web} {IT}{Down, Govt, IT} |
| T6 | {Down, Govt, IT} {IT, Web, Edu} |
| T7 | {Web, IT, Down, SE} |
| T8 | {Edu, Down} {IT, Govt} {IT} |

TABLE 2. Shows Item set sequence

| Category | Items |
|---|---|
| {Web, IT, SE} | A1 |
| {IT, Web, Edu} | A2 |
| {IT} | A3 |
| {Down, Govt, IT} | A4 |
| {IT, Govt} | A5 |
| {Edu, Down} | A6 |
| {Web, IT, SE, Down} | A7 |

TABLE 3 Transaction Table

| Transactions | |
|---|---|
| T1 | A1, A3, A2 |
| T2 | A4, A5, A1 |
| T3 | A6, A2, A7 |
| T4 | A2, A4 |
| T5 | A1, A5, A2, A3, A4 |
| T6 | A7 |
| T7 | A6, A5, A3 |

### *A. Working Steps of Advance Reverse Apriori Algorithm*

From the table 3 we have taken the transactions that contain the maximum number of itemsets. In ARAA the transaction that contains the largest itemsets is taken that forms the C1 table. The candidate itemsets and the frequent itemsets are generated together in the proposed algorithm. The table contains the information related to the support or counts as well the transaction which contains that itemsets. In this algorithm we are using support 1 until we get the support of itemsets more than one i.e. extracting all frequent sequence itemsets in C1, C2 and then support of 2 in rest of the tables. In C2 the candidate sets are generated based on the frequent sequence itemsets of C1. Based on C2 the candidate and the frequent sequence itemsets are generated in C3 along with TID. Here the itemsets whose support is greater than one is placed under the frequent sets with TID and support; however support less than one are placed under the candidate sets. In C4 the candidate and frequent itemsets are generated based on the C3 frequent itemsets. At last in C5 frequent itemsets are generated from the C4 frequent itemsets. The advantage of proposed algorithm is that the number of scans is equal to the number of transactions which drastically reduces the execution time of the system. The disadvantage of proposed algorithm is that it contains the TID and support together therefore increases the complexity. The working of the algorithm is shown in Fig. 5.

### *B. Reverse Apriori Algorithm*

The working of the RAA is opposite to the Apriori algorithm. It scans the $k^{th}$– itemsets first. This algorithm can perform good for higher datasets and poor for lower dataset. It generates the candidate set first and then the frequent itemsets.Kamrul Abedin Tarafteret. al. [12] has given the working of Reverse Apriori Algorithm in their paper. Table 4 shows the frequent itemsets.

TABLE 4. FREQUENT SEQUENCE PATTERNS

| Candidate Set(5 Items) | Frequent Set(5 items) | Scans |
|---|---|---|
| | A1,A2,A3,A4,A5 – 1 {T5} | 7 |

### *C. Apriori Algorithm*

The AA is an iterative level-wise algorithm. It finds all the 1- frequent itemsets then 2- frequent itemsets, 3-frequent itemsets and so on. The working of AA is discussed by many researchers in their papers. Table 5 shows the number of scans taken by Apriori algorithm for generating the frequent sequence patterns.

TABLE 5. Number of Scans in Apriori Algorithm

| Itemsets | Scans |
|---|---|
| Table 1 (C1) | 49 |
| Table 2 (C2) | 105 |
| Table 3 (C3) | 84 |
| Table 4 (C4) | 35 |
| Table 5 (C5) | 7 |

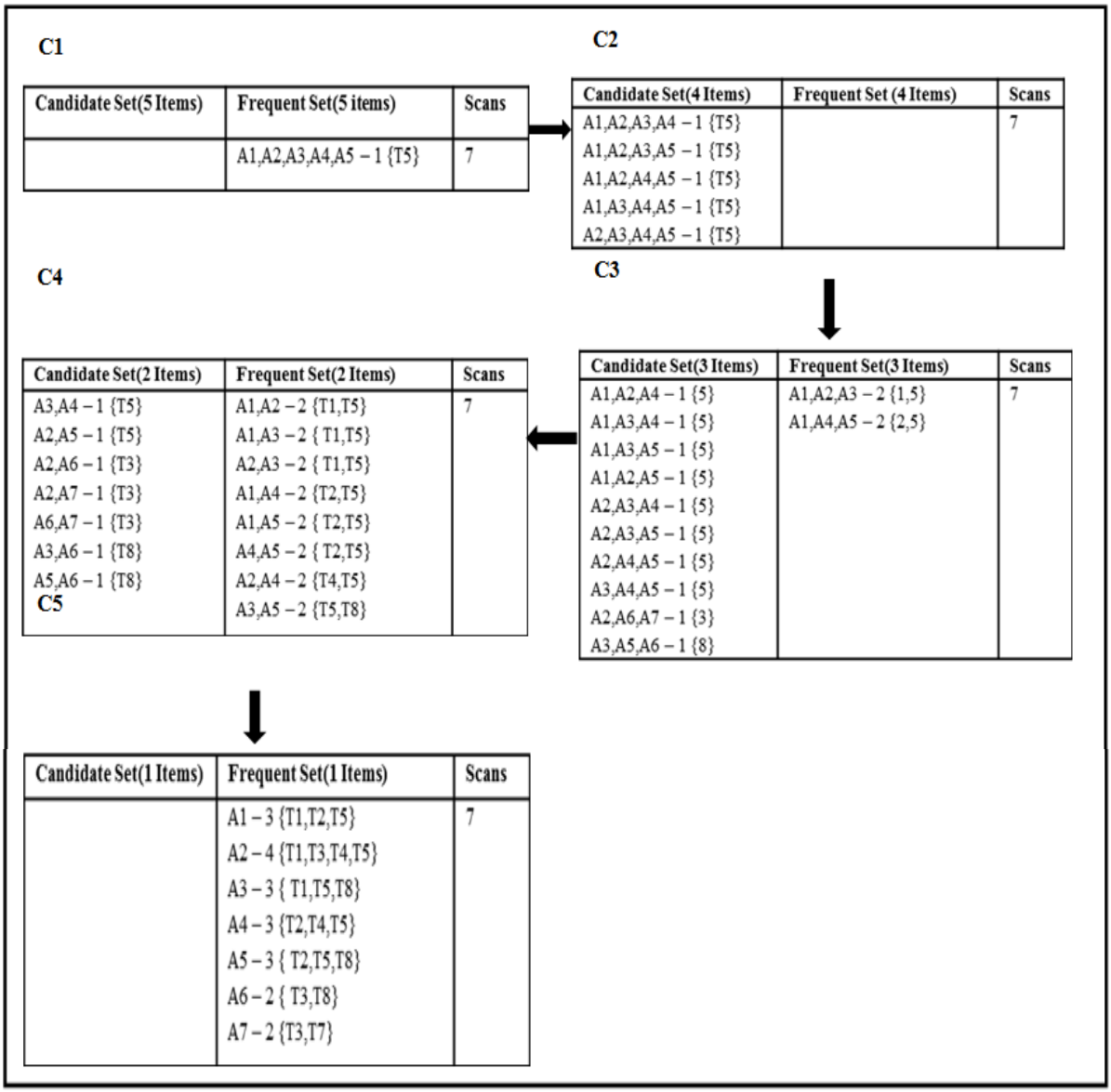


Fig. 5. Working of Advance Reverse Apriori Algorithm

## VI. COMPARISON AND JUSTIFICATION OF AA, RAA, ARAA

The working of the AA shows that AA is simple to implement, it generates the candidate set first and then the frequent itemsets. AA is poor for all rules, the output of Apriori is all types of rules, and the number of scans taken by it is 280 which are comparatively higher than the RAA and ARAA. The disadvantages of AA are that it takes number of scans that increase the time complexity of the system very much. The RAA is good for higher data set rule but it is poor for the lower data set rule. RAA also generates first candidate sets then the frequent itemsets. The output of RAA is 1- itemsets rule; it takes total 7 scans to generate the frequent sequence itemsets. The main disadvantage of RAA is that it is good for higher dataset and therefore we can get only a particular higher level frequent sequence item. Comparing the ARAA with the existing algorithms, it shows that the performance of ARAA is very good for all datasets; output of ARAA is all types of rule. It generates the candidate and frequent sequence itemsets simultaneously. The number of scans taken by ARAA is 35 which are equals to the number of transactions. The advantages of ARAA is that it takes constant and less scans for all frequent set, the lower itemsets rules can be predicted after analyzing the higher itemsets, for $i^{th}$ level ruleset we only need to scan items that contains i-items. The disadvantages of ARAA are that it is complex as it contains the TID with the datasets.

### *A. Graph*

The number of scan done by all three algorithms is shown in Table 6. Here the minimum numbers of scans are taken by the RAA; however it might not be the condition for all type of transactions.

TABLE 6. Number of Scans in All Algorithms

| Algorithm | Scans |
|---|---|
| AA | 280 |
| RAA | 7 |
| ARAA | 35 |

## VII. CONCLUSION AND FUTURE WORK

In this paper, we have improved an existing Apriori and Reverse Apriori Algorithm for mining the web log data. We have also performed a comparative study on the existing two algorithms and our proposed algorithm. The experimental result shows that the approached algorithm drastically reduces the multiple scans and results in overhead processing as compared to the Apriori Algorithm whereas Reverse Apriori Algorithm works well but has limitations based on the type of data. The future work is to reduce the internal complexity of the ARAA algorithm.

## REFERENCES

[1] Zailani Abdullah, TututHerawan, "Mining Significant Association Rules From Educational Data Using Critical Relative Support Approach", Procedia - Social and Behavioral Sciences 28 (2011) 97 – 101, Elsevier
[2] Agrawal R, Imielinski T, Swami A. (1993), "Mining Association Rules Between Sets of Items in Large Databases", Proceeding of ACM SIGMOD International Confidence. Management of Data, Washington: 207- 216.
[3] R. Agrawal and R. Srikant, "Mining Sequential Patterns", In Proc. of the 11th Int'l Conference on Data Engineering, Taipei, Taiwan, March 1995.
[4] Junjie Chen and Wei Liu, "Research for Web Usage Mining Model", International Conference on Computational Intelligence for Modelling Control and Automation,and International Conference on Intelligent Agents,Web Technologies and Internet Commerce (CIMCA-IAWTIC'06) 0-7695-2731-0/06 © 2006 IEEE
[5] SutheeraPuntheeranurak ,Hidekazu Tsuji, "Mining Web Logs for a Personalized Recommender System", 0-7803-8932-S/05/ 2005 IEEE
[6] OlfaNasraoui, MahaSolimanEsin Saka, Antonio Badia, Memberand Richard Germain "A Web Usage Mining Framework for Mining Evolving User Profiles in Dynamic Web Sites", IEEE Transactions On Knowledge And Data Engineering, Vol. 20, No. 2, February 2008
[7] V.Chitraa, Antony SelvdossDavamani, "A Survey on Preprocessing Methods forWeb Usage Data", International Journal of Computer Science and Information Security, Vol. 7, No. 3, 2010
[8] F. Masseglia, P. Poncelet, and M. Teisseire, "Using Data Mining Techniques On Web Access Logs To Dynamically Improve Hypertext Structure", In ACM SigWeb Letters, 8(3): 13-19, 1999.
[9] DiamantoOikonomopoulou, Maria Rigou,"Full-Coverage Web Prediction based on Web Usage Mining and Site Topology", Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI'04) 0-7695-2100-2/04, April 20, 2009.
[10] OzgurCakir, Murat Efe Aras, "A Recommendation Engine By Using Association Rules", Procedia - Social and Behavioral Sciences 62 ( 2012 ) 452 – 456, Elsevier
[11] Goswami D.N., ChaturvediAnshu, "An Algorithm for Frequent Pattern Mining Based On Apriori", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 04, 2010, 942-947
[12] Kamrul Abedin Tarafter, Shah Mosafa Khaled, "Reverse Apriori Algorithm for Frequent Pattern Mining", Asian Journal of Information Technology 7 (12) ISSN 1682-3915, Medwell Journals, 2008
[13] R. Agrawal and J.C. Shafer, "Parallel Mining of Association Rules: Design, Implementation, and Experience", IEEE Trans. Knowledge and Data Eng., vol. 8, pp. 962±969, 1996.
[14] Bina Kotiyal, Ankit Kumar, Bhaskar Pant, R H Goudar, "User Behavior Analysis in Web Log through Comparative Study Of Eclat and Apriori", Proceedings of 7th International Conference on Intelligent Systems and Control (ISCO 2013) 421- 426
[15] Kosala R., Blockeel H., "Web Mining Research: A Survey", SIGKDD explorations: newsletter of the special interest group (SIG) on

## AUTHOR PROFILE

Ms Bina Bhandari working with Graphic Era Hill University as Assistant Professor. Her research area includes Data Mining, Machine Learning and Big Data.

Dr Bhaskar Pant working with Graphic Era deemed university as Associate Professor. He has done Ph.D. from Maulana Azad National Institute of Technology, Bhopal. His research areas include Data mining, Soft Computing, Machine Learning & Bio-Informatics.

Dr R H Goudar working with Visvesvaraya Technological University, Belagavi. He has published several papers in International Journals and Conferences. His research areas are Information Retrieval, Cloud, Networks and Security.