

Social Networking Data Research Using Frequent Pattern Mining and Machine Learning Data



Lakshmi N, Krishnamurthy M

ABSTRACT:- The data are generated by the sources are very large in number with variety of form. These data are organized in to specific format in order to handle properly. Data mining methods are addressed various problem during data extraction process to analytical process. The relevant data are extracted by applying pattern over the huge databases. Association rule mining introduces the method to extracts the related data from the datasets using the performance metrics like support and confidence. Traditional algorithm uses this metrics which is restricted to common attribute format. This problem is addressed by using generic attribute format with frequent pattern mining. The main objective of the paper is to analyze the algorithm and performance metrics related to the frequent patter mining or relevant data. Association rule mining has analyzed with various parameters in single connectivity and multi connectivity rules. Social networking suffers various problem because of uncertain data arrived for processing which is analyzed with various efficiency related elements. The analysis and prediction are also compared with the machine algorithms like classification and clustering and so on. Various frequent pattern mining algorithm is analyzed and review has been carried out based on the performance level.

Index Terms— Data Mining, Frequent Pattern Mining, Association Rule Mining, Machine Learning, Social Networking.

I. INTRODUCTION

Frequent item set is a subset of data mining which is used to search the sequence of items with association for maximum profit in the global market. The sequence is identifying based on the customer behavior with huge amount of existing knowledge. There are two important attributes used in the analysis namely transactions and items. In this algorithm the transaction and item sets should have at least minimum support and confidence. The algorithm uses recursive procedure for finding and generating new item sets with maximum transactions.

Revised Manuscript Received on August 30, 2019.

* Correspondence Author

Lakshmi N*, Research Scholar, CSE, KCG College of Technology, Chennai, Tamilnadu, India.

(Email: livela20@gmail.com)

Krishnamurthy M, Professor, CSE, KCG College of Technology, Chennai, Tamilnadu, India.

(Email: mkrishmails@gmail.com)

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

There are several algorithms are applied for finding frequent items sets such as apriori algorithm, FP growth algorithm and so on. Frequent items are searches from different data bases like relational databases and transactional data bases etc. The item sets are selected by performing correlation analysis between various data sets based on their interest is called association mining. The item sets are classified into maximum, closed and K-item set. There are various application area which covers by the association rule based data selection namely prediction, web usage mining for extraction. It has two stage processes they are frequent item generation and rule generation. The main performance bottleneck occurs due to the main memory usage because the items sets are stored in the memory. If the data sets are very less the main memory is enough whereas the larger dataset are maintained in the disk for improving algorithm efficiency. Nowadays the data are collected from various sources with high dimensional data it will be maintained proper manner. There are multiple passes can be executed during items set identification. Initially, entire data sets are considered and then applied filter for identifying L1 items set. The operation for generation of item set is filter operation followed by construction operation in multiple passes in order to generate exact item set. Hash table mechanism can be applied over the passes then multi hash is extended for single pass. Randomized algorithm is also applied for selecting random samples for generating frequent items with high confidence level. Pattern matching is a technique to finding a frequent item set but it lacks in large number of scanning over database. This problem is overcome by implementing tree based algorithm with association property called FP tree. Frequent pattern mining suffers a scalability problem due to distributed data in large data storage with big data applications. It is applied to various application areas which solves a data mining problem such as web based application, software reliability analysis, medical and engineering application through relevant patterns. Sequential pattern mining is also applied for generating items sets by applying association rule in sequence order. Temporal pattern extraction process has been applied over the GPS and location related contents. Normally the data are categorized into certain data and uncertain data. Certain data has high quality attributes and relevant properties. Uncertain data are low quality in terms of attributes and properties.

Data privacy also handled with their boundary because the real time data can crossed multiple regions. Association rule mining process has been applied either online transactional data as well as offline transaction. Online data are classified into various types such as structured data, semis structure data, unstructured data and quasi structure data. These data are handled very difficult way so efficient method can be implemented. The data types specification plays vital role in the mining process they are streams, strings, spatial based data, graph based and so on. Structured mining extracts the data from semi structured data for generation of item sets. Frequent pattern growth method is used to extract required data items without generation of frequent sets when compared to existing method o generation. Constraint based data extraction process has been carried out for generating relevant items which satisfies condition set. The number of passes in the frequent pattern mining is reduced by applying divide and conquers strategy by handling subsets of candidate items are more frequent. Table 1 represents the advantages of various basic algorithms.

Table 1 Merits of algorithms

Name of the Algorithm	Advantages
Apriori Algorithm	Handle large database efficiently
Continuous Association Rule Mining	Real time item set with minimum number of database scan
ARM with FP Tree	Perform pruning and selection operation with efficient classification algorithm.
Weighted ARM Algorithm	Weighted parameters are considered with less database scan
Genetic Algorithm	Optimization operation is performed with support and confidence metrics
FP-Tree Algorithm	Construction of tree with two database scan

II. ISSUES IN ASSOCIATION RULE MINING

Current FUP method suffers low efficiency problem for updating new candidate set in the database. This problem is overcome by implementing new rule mining algorithm has been introduced with association relationship without any change in the support and confidence parameter in the data extraction process. This method is suitable for online data mining with association rules [10]. Association rules are categorized based on the dimensions such as single dimensional and multidimensional with single connectivity in the rules and it handling process. This process produced only minimum number of rule which gives minimum set of frequent item set, so multiple connectivity rule need to be implemented for better performance. Multi connectivity rules are complex to implement because of huge number of defined rules. This will be tested by using double connected association rule [26]. Most of the association rules are based on the support and confidence metrics, so all the problems are prepared to common set of attributes. The database attributes are same in all the time which never gives accurate output so it will be handled by using dual confidence. it uses graph with corresponding weight for rules for reducing the useless association rules. This method

always based on the user interest and their own set of attributes with maximum efficiency [7].

Tradition association rule with support and confidence metric based method produced huge number of duplicate and inaccurate rules leads the performance problem. This problem is overcome by classifying the association rule as accurate as well as inaccurate rules. Accurate rules are more relevant rules which produced correct set of item sets whereas inaccurate rules are irrelevant which produce wrong result using correction parameter. It also improved by using correlation with dual confidence for better result. The number of rules is minimized by getting high level of frequent item sets using efficient algorithms [16]. Data mining is a process of extracting the facts by performing the operation like classification, clustering, outlier analysis, regression based and association rule based mining [17]. The business decisions are taken by the organization with customer recommendation with strong association rules. The huge data are handled by dividing the data into segment based on the customer sentiments with association rules. This is more suitable for unstructured data with relevant association rules in order to take decisions by domain experts [2].

Ontology is a method of organizing the topics with their semantic relationships over various domain areas. One of the major problem areas is gene organization in biology domain. Gene ontology is applied in gene maintenance through segment based annotations with multiple phases. Association rule with weighted condition with relationship provides expected result during data extraction. Predictive association rule has been introduced over gene ontology as a group by eliminating redundant ontology. The optimization process also applied for achieving fruitful result with high efficiency [9].

Traditional association rule based mining considered data sets and transaction details in the database table. Database always provides association relationship with available data is called positive rules. The database with unavailable data is extracted without any kind of known data is called negative rules. These rules suffer conflicting issues while handling the data. These rules are applied properly without considering the redundant database scan and additional metrics [5]. Recent development of E-commerce based business is handled through information and communication with high end data center using social media and internet technologies. Transactions on internet grows enormously because of huge number of data is produced by e-commerce web application in an rapid manner.

Association rule based mining and recommendation systems are followed by retain the customer in long term business. Incremental through interactive mode of association rule mining is used to perform online business via web application [1]. Datasets used in the association rule mining is a numerical attributes so all kind of data are converted to numerical values before processing. Supervised data is simple to convert when compared to unsupervised data so it will be addressed carefully.

Two-Step clustering method is used to determine the conversion with unique data. It also considered the intervals to evaluate various algorithms [19]. Various algorithms are analyzed and identify the pros and cons of the association rule mining in social media based data.

III. ISSUES IN FREQUENT PATTERN IN SOCIAL NETWORKING

Social networking data are extracted based on heterogeneous properties over different time intervals. These behaviors are organized as a graph and sub graph whenever the topology of the data gets changes. RB miner algorithm considered regular interval for datasets which is acquired from networks. It produces large number of dynamic data which are abstracted in the sub graph itself, so it will be considered during analysis [11]. Heterogeneous data are produced by people in rapid without any time latency. Social media mining technique is used to identify the behavior of the youngsters. This will help to improve the profit in the global market. Natural Language Processing and frequent pattern mining are integrated for classifying the trending topic with an efficient manner. The sentiment analysis is carried out through textual information i.e. tweeter data. The generic data handling process is done for global trending. Behavioral pattern of recent trends are analyzed and identified based on their interests in order to improve the performance and efficiency [26].

Geotags are attached to the social media information is used to track the end users activities and movements periodically. These data are different in nature because of the various vendors application handled as mobile as well as web application. This is handled properly for accurate analysis. Existing technique of handling these social media pattern faces difficulty during data movement collection process. This is overcome by using visual based interactive system for better data exploration. heuristic model has been introduced for eliminating uncertainty which is occur over social media information [3]. Most of the eastern countries use social media for sharing the data globally. The social data are represented in a generic manner by using natural language processing technique. SFWI (Set of Frequent Word Itemset) has been introduced by representing and handling meaning of the data. Compact pattern mining with incremental manner is considered for better understanding and relationship.

The large text are handled as a small text by using tweeter social media for experiment in order to use minimum time as well as memory interaction [2]. Social data with common groups and their relationships are similar when compared to ungroup data. Common responses are analyzed through frequent pattern mining through different social media data, data sources and related rules. Individual member of the group is consistent across each other with effective behavior. If there are any changes in the group response will provides the accurate data to the peers in the corresponding group [7]. Social networking changes the entire people life such as representation of data, groups and communication between the person etc. the main challenges in the knowledge extraction process is volume of data because heterogeneous data are pulled the social networking application. The stream of the data suffers issues during

sharing of common data over individual as well as community. Integration of FP-tree with LRU organization provides better efficiency for discovering frequent and association rules [8]. Social data are not true because the data with sensitivity is never shared by any people who are part of the social networking. Selection of the data is problem during the data analysis. Integrated data are handled for discovering a pattern with more accuracy through demographics.

These data like audio, video, photographs and any form of representation. The data structure which is used for maintaining such types of data is a challenging task in frequent item set generation [13]. Social stream streaming identifies the controversial topics and objects through pattern mining. Sliding window based approach is implemented for data streams. Peak time rumor data are analyzed with various stages in order to captures the frequent pattern as a single pattern. This kind of approach also uses short texts which are extracted from long text by using conversion method [18]. Social data are classified as functional based behaviors and social based behaviors. The users and its behaviors uses two different time plane for collecting and analyzing different people data between vertical and horizontal manner.

GSP (generalized sequential pattern algorithm) uses frequent pattern of the certain people with accurate prediction [6]. Frequent pattern mining suffers explosion in pattern discovery and less descriptive pattern. This problem is overcome by using distributed data handling with high level of understanding over the pattern. Multimedia data are considered with visual characteristics via vector based cluster analysis. It handles the high dimensional data by using dimensionality reduction methods [13].

IV. ISSUES IN FREQUENT PATTERN IN MACHINE LEARNING

Machine learning with e-commerce data supports a novel as well as accurate knowledge accumulation through past information. Big data provides the services like decision making over the organization by following frequent parent rules. Initially the users are classified based on their interest as a vector of features. These features are clustered and classified according to the feasibility of the association rule [5]. Pattern mining techniques is used to find the set of frequent items from the item sets. Real time data are predicted by considering application area such as policy making in the business, recommendation systems in web application, stock behavior and other relevant application. There is a problem occur in the online sources due to data transformation form one format to another format. The solution of this problem is to maintain generic data set for different kind of domain. Machine learning addresses this problem by handling various data set in any format [16]. Machine learning uses instance based learning method for minimizing memory requirement and running time in order to improve the classification performance. Fuzzy based frequent pattern are handled by considering two level of threshold.

KNN (K-Nearest Neighbor) classifier uses for achieving maximum accuracy[10]. NetFeatureMap is a representation is used to describe extracted features from various data sources with different format. Manual selection of the features leads the inaccuracy problem which is handled by introducing automatic way of extraction process i.e. via intelligent agent. The agent extracts the features based on the frequency of the occurred data items. Genetic Algorithms uses the optimization function through heuristic functions. The efficiency of the patterns are handled by using frequent pattern based intelligent software agent [24]. Frequent pattern mining handles real time data such as online streaming data, network traffic data, web data, record stream data and other multimedia data. The non-frequent data are removed from the set faces a problem due to continuous arrival data from different sources. Efficient data structure is used to maintain any kind of data using CPS-tree with complete data stream. Depth its based scanning is done by addressed the imbalance conditions over parallel processing and its algorithm [25].

V. ALGORITHM ANALYSIS OF FREQUENT PATTERN MINING

Support Framework has been implemented for identifying pattern by using raw frequency and also compared with threshold value. This is used a property called phase wise property for defining frequent patterns with item sets. Frequency pattern mining with conditions are imposed in order to generate item set is called constraint based frequent pattern mining. The mining pattern is so large in volume which leads the performance problem which is addressed by using compressed version of frequent patterns. Join based method generates new candidate set from already available candidate set. Apriori algorithm is a level based algorithm which generates new level from previous level. Direct Hashing and Pruning (DHP) algorithm uses two optimization method called pruning and trimming with efficient support count metrics.

TreeProjection algorithm uses recursive projection in hierarchical manner i.e. level by level. It uses depth first and breath first traversal for identifying better candidate generation migrated to one level to another level. It reduces the level counting which improve the efficiency. VIPER method is used to generate frequent pattern using bit vector in compressed version is called snakes. This is the way to count the frequent pattern from various data sources. FP-growth algorithm explores the pattern from the corresponding database with accurate item set count. It uses a tree with constraints based transactions for frequent pattern with labeling scheme. Association rule mining for planning a travel by the customer using customized K-day item sets [4].

Quine-McCluskey techniques are introduced by evaluating the customer's behaviors with clusters of frequent items [2]. Mining operation is carried out over XML data by traversing the node of an XML tree to find the frequent patterns [17]. Normally the data are handled in the mining methods are binary which leads the accuracy problem, so decimal based mining has been introduced [3].

Utility bases data mining is introduced for achieving high quality items set with association rules with better accuracy [19]. Figure 1 shows that the comparison of various frequent pattern mining algorithms.

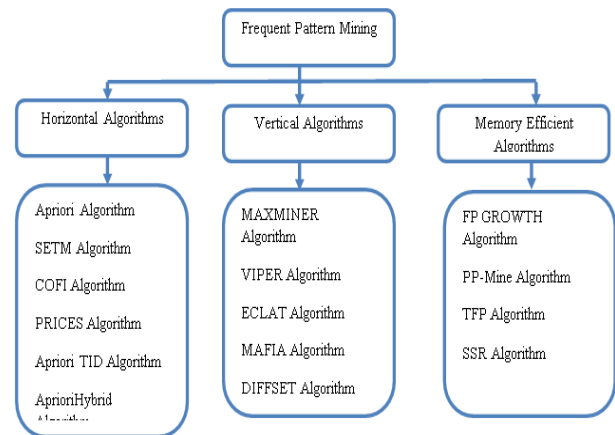


Figure 1. Comparison of frequent pattern mining algorithms

VI. RESULTS AND DISCUSSION

Apriori algorithm extracts the frequent item sets from large collection of data using association rule for better analysis. FP-Growth (Frequent Pattern Growth) algorithm performs mining for generating candidates using divide and conquer methodology. It also retains the related information of frequent item sets through association process. ECLaT (Equivalence Class Transformation) algorithm extracts the suitable items from the collection of data for all transaction. It uses vertical format which maintains transaction into single record and database scan. TreeProjection (TP) algorithm extracts the frequent item sets by constructing lexicographic tree with various searching method such as BFS, DFS and Combination of these methods for better performance. COFI (Co-Occurrence Frequent Item set) algorithm also mines the item sets by pruning concept in order to reduce the memory and usage with proper threshold level.

TM (Transaction Mapping) algorithm maps the different locations transactions with specified time interval through DFS search method over tree. LP-Growth (Linear Prefix Growth) algorithm reduces the data which is required for mining process by using linear prefix tree with arrays. It mainly used for memory in more efficient manner [9].

The vertical and horizontal organization of algorithm gives better comparisons and analysis with reduces memory management techniques. The execution time of various horizontal and vertical algorithms is show in figure 2 and figure 3 respectively. Figure 4 represents the efficient memory usage of various algorithms.

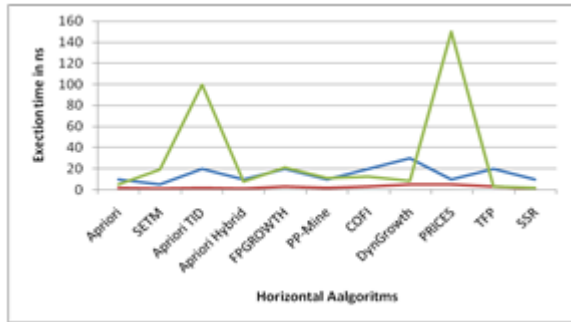


Figure 2. Comparison of Horizontal Algorithms

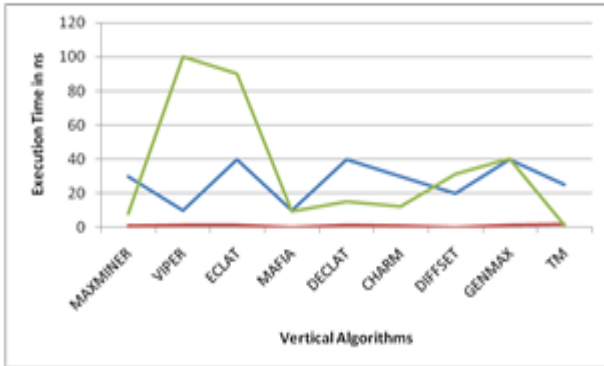


Figure 3. Comparison of Vertical Algorithms

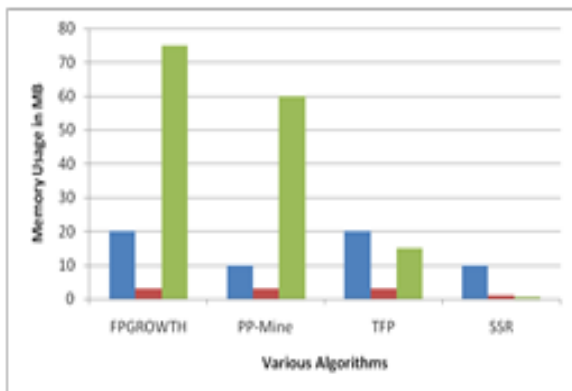


Figure 4. Comparison of memory consumption of various Algorithms

VII. CONCLUSION

Data mining process performs various tasks for extracting the knowledge from the data warehouse. It collects data from different sources then represented in the schema which will help to identify the data format. Association rule mining is analyzed and performed comparison over existing algorithms. The support and confidence metrics have been used and also evaluated with dual and multi-connectivity rule. Frequent pattern mining is analyzed and identified performance-related metrics. Sequential and random pattern mining also compared with various algorithms and pros and cons are assessed. Frequent pattern mining with social networking is applied and reviewed in different perspectives because different data formats exist in the social data. The data accuracy is achieved with the help of machine learning and its operations like regression, classification, clustering and deep neural networking. The prediction-based analysis is carried out for better understanding of knowledge. Various algorithms over frequent pattern mining with corresponding rules have been assessed and analyzed. The overall objective

of this paper is to perform a review of literature over frequent pattern mining for item sets.

REFERENCES

1. A. Sabri Alvar, M. Saniee Abadeh, "Efficient instance selection algorithm for classification based on fuzzy frequent patterns", DOI: 10.1109/CINTI.2016.7846426, 2017, IEEE, Pp 12-18.
2. Ardra, Blessy Merin Varughese, Merline Susan Joseph, Preethi Elsa Thomas, K. K. Sherly, "Analyzing the behavior of youth to sociality using social media mining", International Conference on Intelligent Computing and Control Systems (ICICCS), 2017, DOI: 10.1109/ICCONS.2017.8250665, Pp 1-6.
3. Casey Doyle, Alex Meandzija, Gyorgy Korniss, Boleslaw Szymanski, Derrick Asher, Elizabeth Bowman, "Mining Personal Media Thresholds for Opinion Dynamics and Social Influence", IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), DOI: 10.1109/ASONAM.2018.8508316, IEEE, 2018, Pp 1-6.
4. Chowdhury Farhan Ahmed, Md. Samiullah, Nicolas Lachiche, Meelis Kull, Peter Flach, "Reframing in Frequent Pattern Mining", 27th International Conference on Tools with Artificial Intelligence (ICTAD), DOI: 10.1109/ICTAI.2015.118, IEEE, 2017, Pp 1-6.
5. D Sathyanarayanan, M. Krishnamurthy, "An Efficient Association Rule Based Dynamic Support Count Adaptation Model for XML Databases Using XQuery", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 6 (2015) pp. 16129-16147.
6. Dian Sa'adillah Maylawati, Muhammad Ali Ramdhani, Ali Rahman, Wahyudin Darmalaksana, "Incremental technique with set of frequent word item sets for mining large Indonesian text data", 5th International Conference on Cyber and IT Service Management (CITSM), DOI: 10.1109/CITSM.2017.8089224, 2017, Pp 1-8.
7. Hemant Kumar Soni, Sanjiv Sharma, Manisha Jain, "Plausible characteristics of association rule mining algorithms for e-commerce", Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), DOI: 10.1109/AEEICB.2017.7972379, IEEE, 2017, Pp 1-10.
8. Hui-ke Rao, Zhi Zeng, Ai-ping Liu, "Research on personalized referral service and big data mining for e-commerce with machine learning", 4th International Conference on Computer and Technology Applications (ICCTA), DOI: 10.1109/CATA.2018.8398652, IEEE, 2018, Pp 1-8.
9. Jia Yubo, Duan Yuntao, Wang Yongli, "An Incremental Updating Algorithm for Online Mining Association Rules", International Conference on Web Information Systems and Mining, 2009, ISBN: 978-0-7695-3817-4, DOI: 10.1109/WISM.2009.37, IEEE, 2009, Pp 1-6.
10. Kannan N. Jaishankar, M. Krishnamurthy, S. Kousalya and A. Kannan, "Frequent Item Set Generation using Efficient Utility Mining Algorithm", International Review on Computers and Software, Volume 9, Issue 6, 2014, Pp 1-10.
11. Leiming Yan, Jinwei Wang et al, "Extracting Regular Behaviors from Social Media Networks", Third International Conference on Multimedia Information Networking and Security, DOI: 10.1109/MINES.2011.137, IEEE, 2011, Pp 1-5.
12. M. Krishnamurthy, D. Sudha, Y. Narasimha Rao, R. Baskaran, A. Kannan, "An Effective and Customized Itinerary Planning System Using Association Rule Mining Technique with Personalized Points of Interest", Circuits and Systems, 2016, 7, Pp.1120-1131.
13. M. Krishnamurthy, K. Manivannan, A. Chilambuchelvan, E. Rajalakshmi and A. Kannan, "Enhanced Candidate Generation for Frequent Item Set Generation", Indian Journal of Science and Technology, Vol 8(13), DOI: 10.17485/ijst/2015/v8i13/60756, 2015, Pp. 1-7.
14. M. Krishnamurthy, K. Manivannan, A. Chilambuchelvan, E. Rajalakshmi and A. Kannan, "Frequent Itemsets Generation from Transaction Database and Rule Generation using Quine-McCluskey Method", Middle-East Journal of Scientific Research 23 (Sensing, Signal Processing and Security):2015, Pp. 222-231.

15. P. Dhanalakshmi, R. Porkodi et al, "Extraction of Association Rules from Tobacco Smoke Effect on the Placenta Microarray Dataset using Gene Ontology Based Optimized Association Rule Mining", International Conference on Current Trends towards Converging Technologies (ICCTCT), DOI: 10.1109/ICCTCT.2018.8550982, 2018, Pp 1-12.
16. Peddi Kishor ; Sammual Porika et al, "An efficient approach for mining positive and negative association rules from large transactional databases", International Conference on Inventive Computation Technologies (ICICT), DOI: 10.1109/INVENTIVE.2016.7823240, IEEE, Pp 1-6.
17. Saki Kawanobe, Tomonobu Ozaki, "Extraction of Characteristic Frequent Visual Patterns by Distributed Representation", 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA), DOI: 10.1109/WAINA.2017.71, IEEE, 2017, Pp 1-8.
18. Shihan Wang, Izabela Moise, Dirk Helbing, Takao Terano, "Early Signals of Trending Rumor Event in Streaming Social Media", IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), DOI: 10.1109/COMPSAC.2017.115, 2017, IEEE, Pp 15.
19. Shubha S. Jain, B. B. Meshram, Munendra Singh, "Voice of customer analysis using parallel association rule mining", IEEE Students' Conference on Electrical, Electronics and Computer Science, 2012, DOI: 10.1109/SCEECS.2012.6184770, IEEE, Pp 1-8.
20. Siming Chen, Xiaoru Yuan, Zhenhuang Wang, Cong Guo, Jie Liang, Zuchao Wang, "Interactive Visual Discovering of Movement Patterns from Sparsely Sampled Geo-tagged Social Media Data", IEEE Transactions on Visualization and Computer Graphics, Volume: 22 , Issue: 1 , 2016, Pp. 270 – 279
21. Surbhi K. Solanki, Jalpa T. Patel et al, "A Survey on Association Rule Mining", Fifth International Conference on Advanced Computing & Communication Technologies, DOI: 10.1109/ACCT.2015.69, IEEE, 2015, Pp 1-6.
22. Suwook Ha, Yong Mi Lee, Kwang Woo Nam, Keun Ho Ryu, "An algorithm for finding frequent patterns in social media stream", International Conference on ICT Convergence (ICTC), DOI: 10.1109/ICTC.2013.6675438, 2013, IEEE, Pp 1-8.
23. Swee Chuan Tan et al, "Improving Association Rule Mining Using Clustering-based Discretization of Numerical Data", International Conference on Intelligent and Innovative Computing Applications (ICONIC), DOI: 10.1109/ICONIC.2018.8601291, IEEE, 2018, Pp 1-9.
24. Valquíria Aparecida Rosa Duarte, Rita Maria Silva Julia, "Improving NetFeatureMap-Based Representation through Frequent Pattern Mining in a Specialized Database", 28th International Conference on Tools with Artificial Intelligence (ICTAI), DOI: 10.1109/ICTAI.2016.0147, 2016, IEEE, Pp 34-42.
25. Xi Fu, Lei Shi, Jing Li, "Balanced Parallel Frequent Pattern Mining over Massive Data Stream", IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService), DOI: 10.1109/BigDataService.2017.15, IEEE, 2017, Pp 1-8.
26. Xiufeng Piao, Zhanlong Wang et al, "Research on Mining Positive and Negative Association Rules Based on Dual Confidence", Fifth International Conference on Internet Computing for Science and Engineering, 2010, DOI: 10.1109/ICICSE.2010.28, IEEE, 2010, Pp 1-8.
27. Xunwei Zhou, Hong Bao et al, "Mining Double-Connective Association Rules from Multiple Tables of Relational Databases", International Conference on Computer Science and Software Engineering, Print ISBN: 978-0-7695-3336-0, INSPEC Accession Number: 10437339, DOI: 10.1109/CSSE.2008.559, IEEE, 2008, Pp 1-8.
28. Yiheng Zhou, Numair Sani, Jiebo Luo, "Fine-grained mining of illicit drug use patterns using social multimedia data from instagram", IEEE International Conference on Big Data (Big Data), DOI: 10.1109/BigData.2016.7840812, IEEE, 2016, Pp 1-8.
29. Yihua Zhong, Yuxin Liao et al, "Research of Mining Effective and Weighted Association Rules Based on Dual Confidence", Fourth International Conference on Computational and Information Sciences, DOI: 10.1109/ICCIS.2012.232, IEEE, 2012, Pp 1-6.
30. Yunan Pan, Zhongjie Wang, "Data Collection and Modeling of Personal Social Behaviors", International Conference on Service Science (ICSS), DOI: 10.1109/ICSS.2015.32, IEEE, 2015, Pp 1-8.