



SURVEY ON FREQUENT PATTERN MINING

¹M. PUSHPALATHA, ²Dr. A. VIJAYA

¹ Assistant Professor, ² Assistant Professor and Head,

¹ Department of Computer Science, ² Department of Computer Application,

¹ Padmavani Arts & Science College for Women, ² Government Arts College (Autonomous)

^{1,2} Salem – 636 007.

ABSTRACT: Dynamic Pattern recognition is viewed as a significant test inside the field of data mining and knowledge discovery. Frequent pattern (itemset) mining assumes a significant function in association rule mining. The Apriori and FP-growth algorithms are the most renowned algorithms which can be utilized for Frequent Pattern mining. This paper presents the study of different Frequent Pattern Mining and Rule Mining calculation which can be applied in different fields. The examination of writing study would give the data about what has been done already in a similar zone, what is the current pattern and what are the other related regions. This paper clarifies the ideas of Frequent Pattern Mining and it additionally clarifies different frequent pattern algorithms and how it very well may be applied to various zones. This paper without a doubt causes the investigates to get away from about the use of frequent pattern mining calculation in different zones.

Keywords: [Frequent Pattern Mining, Apriori, FPgrowth, and Association Rule Mining.]

1. INTRODUCTION

As of late measure of data in the database has expanded quickly. The expanding size of the database has prompted developing enthusiasm for extraction of valuable data from the majority of data. Data mining is a method helpful for achieving valuable data from tremendous databases. Certain data inside a database can be helpful in errands, for example, marketing, financial forecast and so forth. This data must be inferred productively. Frequent pattern mining finds noteworthy connections among factors or things in a dataset. Motivation behind this paper is to get familiar with the principle significant ideas of frequent pattern mining. In data mining we may state that a pattern is a specific data

conduct, course of action or structure that may be of a business interest. Itemset is set of things, a gathering of component that speaks to all together substance. A frequent itemset is an itemset that occurs frequently. In frequent pattern mining to check whether an itemset happens frequently or not we have a boundary called backing of an itemset. An itemset is named frequent if its help tally is more prominent than the base help check set up at first.

In mining pattern stage various methods are applied to discover contender for frequent patterns and afterward frequent patterns are produced. There are two primary issues with frequent pattern mining methods. First issue is that the database is checked commonly,

second is intricate applicant age measure with an excessive number of up-and-comer itemset created. These two issues are productivity bottleneck in frequent pattern mining. Studies exhibit that a great deal of endeavors have been performed for concocting best strategies and worth referencing approaches are Apriori, RARM, ECLAT, FP Growth and ASPMS algorithms.

2. TECHNIQUES FOR FREQUENT PATTERN MINING

There are different strategies are proposed for creating frequent itemsets with the goal that association rules are mined effectively. The methodologies of creating frequent itemsets are separated into essential three procedures.

Horizontal design based data mining procedures: Apriori calculation

Vertical design based data mining procedures: Eclat calculation

Projected database based data mining procedures: FP-Growth calculation

3. LITERATURE SURVEY

1. KirilGriazev, SimonaRamanauskaitė said analysis of logical papers uncovered there is a lack of holistic web mining taxonomies or classificatory. This confounds the comprehension of this territory as by and large perspective on the region must be assembled from various sources too an issue of lopsided phrasing may happen. Every one of these realities feature the need of new and more definite, more extensive web mining taxonomy. They proposed taxonomy depends on existing arrangement standards consequently guarantees a similar phrasing, which is utilized something like date. New classificatory and potential classifications are introduced in the taxonomy as well. The additional components permit the introduction of recently risen issues or techniques in the zone of web mining. As a result of incorporation of existing and new characterization models the proposed taxonomy will be more modern and will

introduce the general perspective on web mining.

2. Sanjay K.S and Dr.AjitDanti proposed strategy utilized WEKA instrument to gather and sort conclusions surveys about an item. The proposed strategy, assessments are ordered utilizing different factual measures to give appraisals to help the sentimental analysis of enormous data. Exploratory outcomes show the effectiveness of the proposed strategy to help in analysis of nature of item, marketers assessment of accomplishment of another item dispatched, figure out which forms of an item or administration are mainstream and distinguish demographics like or dislike of item includes, and so forth.

3. Ziyin Zhang, Wei Hao was proposed another cloudlet content caching algorithm dependent on web mining. The new algorithm misuses the k-means clustering and association rule to find most frequently got to web objects to be reserved to the cloudlet. Also, this algorithm takes the size of the web object and cloudlet store size into thought. A reproduction and analysis joined test was performed to check this algorithm. The outcome demonstrated that contrasted with the customary caching algorithm, the new algorithm will expand the store hit proportion and diminish the reaction time. Nonetheless, it must be noticed that in our examination, they utilized a VM to reproduce the cloudlet and a few components are disregarded. For instance, the web object search time in the cloudlet isn't taken into thought. In this way, later on they will develop the total cloudlet design utilizing OpenStack++ to actualize and assess the proposed cloudlet content caching algorithm.

4. Pratik Saraf, R. R Sedamkar, and SheetalRathi, proposedPrefixspan Algorithm is created dependent on projection is utilized in PrefixSpan algorithm [5] for mining sequential patterns. The fundamental thought behind this strategy is, as opposed to

projecting sequence databases by assessing the frequent events of sub-sequences, the projection is made on frequent prefix. This assists with diminishing the preparing time which eventually expands the algorithm proficiency. Jian Pei et al. proposed a novel algorithm called PrefixSpan (Prefix-extended Sequential Pattern Mining) algorithm [5] which works on projection of database and sequential pattern growth. PrefixSpan algorithm is assessed by running the algorithm on various datasets (C16D248k, C16D150k and C21D36k). The two boundaries least help and greatest prefix length are given at beginning of the execution of algorithm. The sequences having esteem more noteworthy than least help are separated from sequential datasets. Least help is the quantity of sequences which are determined by partitioning the pattern happens with the complete number of sequences in the database. The most extreme prefix pattern esteem used to indicate the length of the sequence to be there in yield sequential patterns which is advantageous while executing the algorithm on enormous datasets. For getting the sequential yield dependent on least help and most extreme prefix length, the two boundaries time unpredictability and memory usage are set as the benchmark for execution assessment of algorithm on various datasets. Both the boundaries differ from one dataset to other. These outcomes are plotted in reference diagram and are valuable so as to break down the presentation of existing algorithm.

5. Niek Tax, Natalia Sidorova, Wil M.P. van der Aalst, proposed Prefixspan Algorithm is created dependent on Process mining intends to remove novel experiences from occasion data (van der Aalst, 2016). Cycle discovery assumes a conspicuous part in measure mining. The objective is to find a cycle model that is delegate for the arrangement of occasion sequences regarding begin to-end conduct, for example from the beginning of a case till its end. Many cycle

discovery algorithms have been proposed and applied to an assortment of genuine cases. A more traditional point of view on finding bits of knowledge from occasion sequences can be found in the territories of sequential pattern mining (Agrawal&Srikant, 1995) and scene mining (Mannila et al., 1997), which center around finding frequent patterns, not focusing on portrayals of the full occasion sequences from begin to end.

6. Prof. AlpaReshamwala, Ms.Neha Mishra, proposed Prefixspan Algorithm is developed based on pattern mining is the mining of frequently occurring ordered events or subsequences as patterns literature [5,6]. Compared with projected databases and subsequence connections, PrefixSpan [11] was more efficient than SPADE and SPAM. PrefixSpan does not require candidate generation, also it can reduce the scale of projected databases substantially relative to the original sequence database, and the major cost of PrefixSpan is the construction of projected databases. In addition, scanning projected databases repeatedly also reduce the efficiency of the algorithm. Generally speaking, reducing both the scale of projected databases and the time of scanning projected databases are the main ways of improving PrefixSpan [7,8,9]. However, when mining long frequent concatenated sequences, this method is inefficient. Therefore, it is impractical to apply PrefixSpan to mine long contiguous sub-sequences from sequential database. The comparison study of SPADE, SPAM and the Prefixspan algorithm is done on the results collected. As per the results, SPADE performs better in both the dense as well as sparse dataset such as sign and kosarak10k respectively. Performance of SPAM is worst when executed on sparse dataset. Prefixspan is approximately approaching the performance of SPADE in sparse dataset whereas the performance of SPAM and Prefixspan is same in dense dataset. The number of sequences generated is same in both the dataset. For dense dataset

prefixspan is uses less memory whereas in sparse dataset it utilizes the most. In Dense dataset SPAM and SPADE are utilizing approximately constant memory. In sparse dataset minimum utilization of memory is by SPADE.

7. Jian Pei, Jiawei Han, Wei Wang, proposed Constraint Based Sequential Pattern Mining Algorithm is developed based on A constraint C for sequential pattern mining is a boolean function $C(\alpha)$ on the set of all sequences. The problem of constraint-based sequential pattern mining is to find the complete set of sequential patterns satisfying a given constraint C . Constraints can be examined and characterized from different points of views. We examine them first from the application point of view in this section and then from the constraint-pushing point of view in the next section, and build up linkages between the two by a thorough study of constraint-based sequence mining. From the application point of view, we present the following seven categories of constraints based on the semantics and the forms of the constraints. The problem of pushing various constraints deep into sequential pattern mining. We characterize constraints for sequential pattern mining from both the application and constraint-pushing points of views. A general property of constraints for sequential pattern mining, prefixmonotoneproperty, is identified.

8. Yu Hirate, Hayato Yamana, proposed Constraint Based Sequential Pattern Mining Algorithm is developed based on constraint-based mining and extended sequence-based mining, in this paper, we generalize sequential pattern mining with item interval. The generalization includes three points; (a) a capability to handle two kinds of item-interval measurement, item gap and time interval, (b) a capability to handle extended sequences which are defined by inserting pseudo items based on the interval itemization function, and (c) adopting four item-interval

constraints. Generalized sequential pattern mining is able to substitute all types of conventional sequential pattern mining algorithms with item intervals. Evaluations using a Japanese earthquake dataset confirmed that generalized sequential pattern mining with time interval is able to extract interval extended sequences that include time interval with variable segmentation size. In addition, by adapting our types of constraint related to time intervals, it also excludes extraction of interval extended sequences with time intervals in which the user is not interested.

9. Bhawna Mallick , Deepak Garg , and Preetam Singh Grover, proposed CFM-PrefixSpan Algorithm (CFM-Compactness, Frequency, Monetary) is created based on PrefixSpan [25] is the most favorable pattern-growth approach, which is based on developing the patterns recursively. Based on apriori (e.g., GSP algorithm) and pattern growth (e.g., PrefixSpan algorithm) approaches, many algorithms have been proposed for fruitful sequential pattern mining. A vigorous CFML-PrefixSpan algorithm for mining all CFML sequential patterns from the client exchange information base. The CFMLPrefixSpan algorithm has used a pattern-growth approach that finds sequential patterns through a partition and-overcome technique.

10. Niti Ashish Kumar Desail and Amit Ganatra, proposed CFM-PrefixSpan Algorithm (CFM-Compactness, Frequency, Monetary) is created based on Most of the current SPM techniques work simply on frequency parameter, officially known as help edge. Backing is critical to recognize if patterns show up over and again or not. Then again, the proposed Constraint-based Prefix Span algorithm would be worried about chief's observation. Proposed algorithm through utilization of Recency constraint decides current purchasing behaviors and through utilization of benefit constraint decides the more gainful purchasing

behaviors. Further compactness constraint can be utilized to recognize purchasing conduct of client during explicit time span including seasonal patterns. Proposed Constraint-based Prefix Span algorithm isn't limited to traditional Sequential Pattern Mining (SPM) parameter frequency yet joins six more significant parameters like Gap, Recency, Compactness/Duration, Profitability, Item and Length. Joining of these constraints in FP-growth based—Prefix Span prompts more productive and compelling outcomes by decrease of patterns. Compact patterns present pertinent and exact outcomes regarding clients' advantage. Seven distinct examinations are performed on IBM created six engineered datasets. Examination made for run times and pattern age of three algorithms: proposed constraint-based Prefix Span with RFM and Prefix Span. Proposed constraint-based Prefix Span algorithm is more proficient and viable regarding decrease of patterns age of fascinating patterns for client.

11. BhawnaMallick, P. S. Grover and Deepak Garg, proposed CFM-PrefixSpan Algorithm(CFM-Compactness, Frequency, Monetary) is created based on CFM patterns from the static information base are productively mined utilizing the CFM algorithm proposed in our past work . The applicable aspect of the CFM algorithm that is based on PrefixSpan is introduced here for the fulfillment of this article. We have utilized two ideas in particular, financial and compactness that are gotten from the total and duration constraints which are introduced in the accessible writing. an effective reformist CFM-miner algorithm to deal with the support issue of CFMsequential patterns. We have assembled a refreshed CFM-tree utilizing the CFM-sequential patterns acquired from the static information base to control the dynamic idea of information refreshing cycle and cancellation measure into the sequential pattern mining issue. Therefore, the information base gets refreshed from the conveyed data set that might be static,

embedded, or erased. At whatever point the information base is refreshed from the numerous sources, CFM tree is likewise refreshed by including the refreshed arrangement. At that point, the refreshed CFM-tree is utilized to mine the reformist CFM-patterns utilizing the proposed tree pattern mining algorithm. In the end, the experimentation is completed utilizing the engineered and genuine datasets that are given to the reformist CFM miner utilizing string condition.

12. Probst, Oliver, proposed COPRE framework(Constraint Based Prefix) is created based on constraints which are characterized in a Platform-Independent Model (PIM) into running code that mirrors the predefined constraints. One reason for the overview is that this semantic change is viewed as an open issue which must be unraveled to have the option to utilize Model-Driven Development (MDD) approaches when all is said in done for building data frameworks. The creators notice four preferences of their approach: • The monitoring code which handles the information approval and constraint assessment is independent of the functional code • Easy XML based constraint detail which can be naturally meant monitoring code • The functional code stays unaffected • Cleaner monitoring and functional code in addition to simpler support We classify this idea as a cross-tier approval technique in light of the fact that the interceptor based approach could hypothetically be applied to the solicitation/reaction messages between the introduction tier and the rationale tier or the rationale tier and the information base tier, for example not coupled to a particular tier which means cross-tier.

13. JoernWuebker, Spence Green, John DeNero, SašaHasan, Minh-ThangLuong, proposed COPRE framework(Constraint Based Prefix) is created based on three measurements underneath that score interpretations by the attributes that are

generally applicable in an intuitive setting: the exactness of the primary expressions of the postfix and the general nature of the addition. Every measurement takes model triples (f, ep, e^*) delivered during an intelligent MT session wherein ep was created during the time spent building e^* . A reproduced corpus of models can be created from an equal corpus of (f, e^*) matches by choosing prefixes of every e^* . A comprehensive reproduction chooses all conceivable prefixes, while a tested recreation chooses just k prefixes consistently at arbitrary for every e^* . Registering measurements for thorough reproductions is costly in light of the fact that it requires performing addition forecast derivation for each prefix: $|e^*|$ times for each reference. state based and neural interpretation approaches can be utilized to finish halfway interpretations. The repetitive neural framework gives higher word forecast exactness, yet requires long derivation on a GPU. The expression based framework is quick, produces various n-best records, and gives sensible prefix-Bleu execution. The reciprocal qualities of the two frameworks recommend future work in joining these procedures.

14. Alessandro Dal Pal'u, Agostino Dovier proposed COPRE system (Constraint Based Prefix) is created based on quick and modest techniques, it is conceivable to recover exact data about a DNA grouping, its methylation (utilized for epigenetic examines), histones changes, and quality and protein articulation. The cycle can be more than once applied to similar example over years, for example, when a lot of pharmacological treatments. The development of a living being and additionally a particular example of cells at genomic scale can be followed when watching such natural properties. The malignant growth cells incorporate highlights, for example, quick changing genome and cross blend of various offsprings of tumoral cells. Work-in-progress research, we quickly examined the underlying modeling of the

developmental haplotype derivation issue; the issue is attached to examination of genome advancement in malignancy (e.g., as consequence of pharmacological intercessions). The issue is combinatorial in nature, and reasonable for modeling and investigation utilizing rationale programming strategies.

15. Wensheng Gan, Jerry Chun-Wei Lin, Philippe Fournier-Viger, Han-Chieh Chao, Senior Member, IEEE and Philip S. Yu, Fellow, IEEE proposed Fuzzy Sequential Pattern Summarization is created based on different sorts of patterns from different kinds of databases. Simultaneously, in late many years, data mining has been studied broadly and applied generally [1], [3], [4], [9], [13], [14], [15]. These procedures perform well on little datasets, notwithstanding, because of the restricted memory limit and calculation ability of a solitary node, these data mining techniques become wasteful over big data. The memory prerequisites for dealing with the total arrangement of wanted outcomes increment rapidly, and the computational cost can be costly on a solitary machine. All previously mentioned techniques are serialized. When taking care of large-scale data, these strategies are in a general sense inappropriate because of numerous reasons, including the enormous measures of data, infeasibility of transmission capacity constraint, just as the way that larger sources of info requests parallel processing, and security concerns. The principle commitments are that we investigate late advances in parallel sequential pattern mining and give the status of the field in detail, including sequential pattern mining (SPM), parallel frequent itemset mining (PFIM), and parallel sequential pattern mining (PSMP). Both essential algorithms and progressed algorithms for parallel sequential pattern mining are assessed in a few classifications, the key thoughts, preferences and disservices of each approach are likewise called attention to in subtleties. We further give some related

open-source software of PSPM, that may lessen obstructions from examination and algorithm usage. At long last, we quickly call attention to certain difficulties and chances of parallel sequential pattern mining for future exploration.

16. Vinay Kumar Khare, Vedant Rastogi, proposed Fuzzy Sequential Pattern Summarization is created based on Mining Positive and Negative Sequential pattern from databases is helpful for information revelation. The patterns were mined distinctly from the Existing transaction database. New forthcoming transactions databases can't be converged into existing transaction database. So every time new transactions database is mined independently. In this approach we can without much of a stretch update existing transaction database with the appended transaction database. The Merged transaction database (refreshed database) will be mined to get the Positive and Negative Sequential patterns. Converging of Existing and Appended database is performed by utilizing the refreshed compact pattern tree approach. Proposed model is Mining Positive and Negative Sequential patterns in incremental transaction Databases. To mine Positive and Negative Sequential patterns in incremental transaction database in this Approach we can refresh, existing transaction database with appended transaction database by the utilization of Updated Compact pattern tree approach at that point as indicated by their help the new refreshed transaction database table is kept up and we can mine positive and negative sequential patterns with the assistance of CPNFSP algorithms proposed by Weimin Quyang and Qinhua Huang.

17. Gowtham Atluri , Anuj Karpatne , Vipin Kumar, proposed Fuzzy Sequential Pattern Summarization is created based on a data mining algorithm works upon is known as a data instance. In traditional data mining settings, a data instance is unambiguously spoken to as a lot of watched highlights with

discretionary directed names. Notwithstanding, with regards to ST data, there are numerous methods of characterizing instances for a given data type, each subsequent in an alternate STDM formulation. In this part, we survey five regular classes of ST instances that one experience in STDM issues, to be specific, focuses, directions, time arrangement, spatial guides, and ST rasters. An essential approach for speaking to ST rasters is utilizing N-way clusters additionally called as tensors. In a tensor portrayal of a ST raster data, a few measurements are utilized to speak to the arrangement of areas while the rest of the measurement is utilized to speak to the arrangement of time stamps accessible in the ST matrix. For instance, precipitation data is spoken to as a 3-dimensional cluster where the first two measurements catch 2D space and the third measurement catches time. Essentially, fMRI data is spoken to as a 4-dimensional exhibit where the first three measurements catch 3D space and the third measurement catches time. A tensor portrayal of a ST raster data would then be able to be summed up utilizing space-time subspaces that have comparative qualities, which are what might be compared to picture division in ST domains.

18. Fredrik Robertsen, proposed ProgresLatticeMiner (PLM) is created based on the lattice Boltzmann strategy is a technique for computationally recreating fluid dynamics. It has been utilized for recreations of fluid marvels at numerous scales, from infinitesimal permeable media flows [5], to blood flows in vascular systems [55] and large even aerodynamics dimulations [56]. It very well may be utilized for both single and multicomponent fluid reenactments, and permits much more perplexing wonders with molecule suspensions and liquid crystals to be recreated [57]. The technique is appropriate for parallel calculation and has been demonstrated to function admirably on distributed clusters using regular commodity CPUs [58], just as on more particular

equipment with GPUs [59]. The lattice Boltzmann strategy works by discretizing the whole recreation space with a regular lattice, with the kinetic model equation for the fluid approximated uniquely at the lattice sites. Inside every lattice site, molecule speed space is further discretized into a limited arrangement of speeds. The 2D model appeared in figure 3.1 utilizes 9 discrete speed vectors for each lattice site. This is alluded to as a D2Q9 discrete speed set [60], all aside from the middle one of these highlight a neighboring lattice site. In the lattice Boltzmann technique and how it carries on current quickened supercomputer systems. We have indicated the advantage of utilizing computational quickening agents and that they offer incredible exhibition for running LB solvers. While the solver isn't depending on the computational exhibition of GPUs and manycore quickening agents, they do give fundamentally higher memory transfer speeds than regular CPU systems. With the exhibition of the LB strategy being exceptionally dependent on the memory transmission capacity of a system, quickening agents give an amazing platform to fast LB solvers.

19. Alain Casali, Rosine Cicchetti, Lotfi Lakhal, proposed ProgresLattice Miner (PLM) is created based on the shape lattice characterizes an evaluated scan space for different multidimensional data mining issues. In this part, we study the structure of the 3D square lattice in presence of constraint conjunctions. Given such a structure, we propose consolidated portrayals (or outskirts) of the constrained block lattice with a twofold goal: characterizing in a compact way the arrangement space and choosing whether a tuple t has a place with the arrangement space or not. At last, following from standards of levelwise approaches, we give an algorithm for processing consolidated portrayals of constrained 3D square lattices. We consider droning and antimonotone constraints frequently utilized in paired data mining. The 3D square lattice as a reviewed search space.

We likewise inferred consolidated portrayals of the 3D shape lattice confronted with droning as well as antimonotone constraints. Such an outcome is based on the specific structure of constrained block lattice which is a raised space. A work in progress addresses the meaning of a conclusion administrator on the block lattice. Point is to show that the arrangement of shut tuples furnished with the speculation request is a coatomic lattice which is by all accounts isomorphic to the Galois (idea) lattice of the double connection speaking to the first database connection.

CONCLUSION

In this paper, we have overviewed a ton of old and ongoing algorithms for frequent patterns mining. Notwithstanding that, we have examined the sorts of algorithms to give thoughts for scientists to build up the specific or approximate algorithms. We have given a review of past investigations done in this field and recognize some significant gaps accessible in the current information.

REFERENCES

- [1] Kiril Griazev, Simona Ramanauskaitė, “Web Mining Taxonomy”, Published in: [2018 Open Conference of Electrical, Electronic and Information Sciences \(eStream\)](#), Publisher: IEEE, Electronic ISBN: 978-1-5386-6737-8, Print on Demand (PoD) ISBN: 978-1-5386-6738-5
- [2] Sanjay K.S and Dr. Ajit Danti, “Sentimental Analysis on Web Mining using Statistical Measures”, Published in: [2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering \(ICPCSI\)](#), Publisher: IEEE, Electronic ISBN: 978-1-5386-0814-2, Print ISBN: 978-1-5386-0813-5
- [3] Ziyin Zhang, Wei Hao, “Development of a New Cloudlet Content Caching Algorithm Based on Web Mining”, Published in: [2018 IEEE 8th Annual Computing and Communication Workshop and Conference \(CCWC\)](#), Publisher: IEEE, Electronic

ISBN: 978-1-5386-4649-6, Print on Demand(PoD) ISBN: 978-1-5386-4650-2

[4]. Pratik Saraf, R. R Sedamkar, and SheetalRathi, “PrefixSpan Algorithm for Finding Sequential Pattern with Various Constraints”, International Journal of Applied Information Systems (IJ AIS) – ISSN : 2249-0868 Foundation of Computer Science FCS, New York, USA Volume 9 – No.3, June 2015 – www.ijais.org.

[5]. Niek Tax, Natalia Sidorova, Wil M.P. van der Aalst, “Local Process Models: Pattern Mining with Process Models”Published in: Proceedings of the Twenty-Sixth Benelux Conference on Machine Learning (BENELEARN) Published: 10/06/2017.

[6]. Prof. AlpaReshamwala, Ms.Neha Mishra, “Analysis of Sequential Pattern Mining Algorithms”, International Journal of Scientific & Engineering Research, Volume 5, Issue 2, February-2014 1034 ISSN 2229-5518.

[7]. Jian Pei· Jiawei Han ·Wei Wang, “Constraint-based sequential pattern mining: the pattern-growth methods”, J IntellInfSyst DOI 10.1007/s10844-006-0006-z Received: 31 January 2003 / Revised: 23 March 2005 / Accepted: 28 June 2005 © Springer Science + Business Media, LLC 2006.

[8]. Yu Hirate, HayatoYamana, “Generalized Sequential Pattern Mining with Item Intervals”, JOURNAL OF COMPUTERS, VOL. 1, NO. 3, JUNE 2006.

[9]. ManikaVerma, Dr.Devarshi Mehta, “Sequential Pattern Mining: A Comparison between GSP, SPADE and Prefix SPAN”, International Journal of Engineering Development and Research (www.ijedr.org) © 2014 IJEDR | Volume 2, Issue 3 | ISSN: 2321-9939.

[10]. BhawnaMallick , Deepak Garg , and Preetam Singh Grover, “Constraint-Based Sequential Pattern Mining: A Pattern Growth Algorithm Incorporating Compactness, Length and Monetary [11]. Intrusion Detection” The International Arab Journal of Information Technology, Vol. 11, No. 1, January 2014. Received July 15, 2011;

accepted May 22, 2012; published online January 29, 2013.

[12]. NitiAshish Kumar Desail and AmitGanatra, “Efficient constraint-based Sequential Pattern Mining (SPM) algorithm to understand customers’ buying behaviour from time stamp-based sequence dataset” Received: 14 April 2015 Accepted: 03 July 2015 Published: 14 September 2015.

[13]. BhawnaMallick, P. S. Grover and Deepak Garg, “Progressive CFM-Miner: An Algorithm to Mine CFM – Sequential Patterns from a Progressive Database” International Journal of Computational Intelligence Systems, Vol. 6, No. 2 (March, 2013), 209-222.

[14]. Probst, Oliver, “Investigating a Constraint-Based Approach to Data Quality in Information Systems” Global Information Systems Group Institute of Information Systems Department of Computer Science 8th October 2013.

[15]. JoernWuebker, Spence Green, John DeNero, SašaHasan, Minh-ThangLuong, “Models and Inference for Prefix-Constrained Machine Translation” Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 66–75, Berlin, Germany, August 7-12, 2016. c 2016 Association for Computational Linguistics.

[16]. Alessandro Dal Pal`u, AgostinoDovier, “Proceedings of the 12th International Workshop on Constraint-Based Methods for Bioinformatics (WCB’16)” Held in conjunction with the 22th International Conference on Principles and Practice of Constraint Programming (CP-2016) September 5, 2016 Toulouse Business School, Toulouse, France.

[17]. WenshengGan, Jerry Chun-Wei Lin, Philippe Fournier-Viger , Han-Chieh Chao, Senior Member, IEEE and Philip S. Yu, Fellow, IEEEGarg, “A Survey of Parallel Sequential Pattern Mining” Journal Of Latex Class Files, Vol. 6, NO. 1, January 2018.

[18]. Vinay Kumar Khare,VedantRastogi, “Mining Positive and Negative Sequential Pattern in Incremental Transaction Databases”

International Journal of Computer Applications (0975 – 8887) Volume 71– No.1, June 2013.

[19]. Gowtham Atluri, Anuj Karpatne, Vipin Kumar, “Spatio-Temporal Data Mining: A Survey of Problems and Methods Incremental” ACM Computing Surveys, Vol. 1, No. 1, Article . Publication date: November 2017.

[20]. Fredrik Roberts'en, “The Lattice Boltzmann Method, a Petaflop and Beyond” To be presented, with the permission of the Division of Natural Sciences and Technology of the University of Åbo Akademi for public criticism in Auditorium XXI on April 20, 2018, at 13.

[21]. Alain Casali, Rosine Cicchetti, Lotfi Lakhal, “Cube Lattices: a Framework for Multidimensional Data Mining” for this publication at: <https://www.researchgate.net/publication/220906817> Conference Paper May 2003.