2020 IEEE INTERNATIONAL CONFERENCE ON
ADVANCES AND DEVELOPMENTS IN ELECTRICAL AND ELECTRONICS ENGINEERING (ICADEE 2020)

Issues and Research Challenges in Sequential Pattern Mining

Swati Nagori

Department of Computer Science and Engineering
Amity University Madhya Pradesh
Gwalior, India

Dr. Hemant Kumar Soni*
Department of Computer Science and Engineering
Amity University Madhya Pradesh
Gwalior, India
*hemantsoni.gec@gmail.com

2020 IEEE International Conference on Advances and Developments in Electrical and Electronics Engineering (ICADEE) | 978-1-7281-9251-2/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICADEE51157.2020.9368943

Abstract— **The Mining of data is the way to recognize the substantial data from huge databases. Sequential Pattern Mining (SPM), is one of the important aspect of data mining. The SPM is utilized in discovering the consecutive patterns which takes place in huge databases. It likewise recognizes the incessant subsequences as examples from a database. In today's era, an enormous data is required to be accumulated and get saved in the databases. Various industries are focusing towards mining sequential patterns from these databases. SPM is a very latest approach, in which the researchers uncover the sequential patterns. This following paper gives a methodical study on techniques of SPM. Also, this paper provides an organized study on SPM along with the analysis of their techniques. It also deals with the sequential pattern mining problems, research challenges and future trends of SPM.**

Keywords—***Use Sequential Pattern Mining, Substantial Data, Apriori Methods, Pattern Growth Methods, Data Mining.***

I. INTRODUCTION

All Mining of data is the process of extricating (implied, unperceived and potentially productive) data or patterns from huge information sources such as: relative database, depots of data, XML sources, etc [1]. Mining of data is also known as one of the core processes of Knowledge Discovery in Database (KDD). Before actual mining of data, data preprocessing is required. Data preprocessing involves data cleaning, amalgamation, selection and alteration. In data mining, distinct algorithms are used to generate concealed knowledge. After that post processing will take place, which estimates the result of mining according to end-users. Concerning the estimation results, the data can be introduced, if the outcome of the result is adequate, or else we have to continue running a few or all of those processes in loop until we get the adequate outcome. Several techniques of data mining are applied to the data origin; diverse data is extracted out as the outcome of mining. Those data are estimated by definite protocols, like domain data. Once we get the desired outcome, the last step is to envisage the outcome. The result can be presented as data, tables, decision trees, charts, data cubs or 3D graphics. Data mining techniques also useful for time series data [2], rainfall prediction [3] and incremental data mining [4].

There exist two categories of mining of data - descriptive and prescriptive. Descriptive mining is to encapsulate or signalize basic properties of data in data sources, whereas prescriptive mining is to executes conclusion on ongoing data, to make forecast based on the old data. There are different types of techniques of data mining like rules of association [5], clustering and classifications. Web mining and SPM are also well researched on the basis of these techniques. One of the major and well researched techniques of data mining which was first introduced is Association rule mining [6]. Association rule mining problem can also be addressed by genetic algorithm [7] and some efficient algorithm is also available [8-9].

II. SEQUENTIAL PATTERN MINING

All In this, the mining of regularly arising sequenced events or patterns is done. In other words, this mining is a set of item sets organized in a sequential database that arises sequentially in a particular order. An example of this mining is the fundamental information regarding mining issue which differentiates different sub groups of arrangement database.

Data is evolving at every minute; especially that data which is on the internet, they are tremendously active. As time goes on, advanced datasets are loaded and old datasets are removed whereas few other datasets are revised. Also, it is quite noticeable that time stamp is a foremost feature of every set of data. It is also vital in the data mining process and it could also provide us even more defined and functional data. A database contains series of events which gets changed with respect to time and it is known as time-series database [10]. Time-series database accounts the correct time of every dataset. These databases are mostly required in storing old information but in diverse areas like- monetary data, medical data, and scientifically data etc. Various techniques of mining have been formulated for mining of time-series data. Fundamentally, four different types of patterns exist that we get from different types of time-series data:

A. Trend Analysis

This is used for finding modification designs of attributes. They can be seasonal movements, cyclic movements, random movements etc.

978-1-7281-9251-2/20/$31.00 ©2020 IEEE

B. Similarity Search

It seeks to discover sequences that vary moderately. It is a blurry matching process which can help in abiding little dissimilarity in a definite threshold. On the basis of length of sequences, sequence matching can be categorized as: subsequence matching and whole sequence matching.

C. Sequential Patterns

It tries to find out the connection between circumstances of sequential patterns in order to discover that if there is any particular sequence of occurrences. We can also discover the sequential patterns of particular individual items and the sequential patterns that cross various items.

D. Periodical Patterns

These are the frequently occurring events in the time series database whose periodicity can be daily, weekly, monthly, seasonal or yearly.

III. Sequential Pattern Mining Algorithms

The algorithms for Sequential Pattern Mining can be classify in two ways [11]:

1. Pattern of sequence generation and store. The agenda of these algorithms is to lessen the quantity of candidate sequences produced so as to limit I/O cost.
2. Second way is based on support computing and the way of candidate sequence testing for frequency.

Based on above mention criteria, SPM algorithms is categorized into two categories -Apriori based and Pattern growth based.

A. Apriori Based Algorithms

This type of algorithms used for discovering frequent item sets in a dataset for Boolean association rule. The name of this algorithm is Apriori because it utilizes preceding knowledge of recurrent item set properties. We have to apply an iterative approach in which k-frequent item sets are used for finding k+1 item set. In order to improve the efficiency of these periodic item sets, a major property is used that is called Apriori property which helps by lessening the space of search.

1) *Generalized Sequential Pattern (GSP) Algorithm:* The GSP algorithm proposed in [12], doesn't need to discover all the periodic item sets at the first place. GSP Algorithm is used in finding solutions of many issues present in sequence mining. It can be called as a level wise algorithm in which all the items that are frequent are generally seen level-wise and also all singleton items are computed. Then the items that are not frequent are removed. Finally, each and every transaction holds only the frequent items which it was holding initially. This database is sent to the GSP algorithm as an input. The GSP algorithm makes multiple passes. At first pass, it discovers the continual sequences which have the least support. At every pass, each sequence of data is analyzed to upgrade the occurrence number which is holding the sequence. The benefits of GSP is, it reduces the search space whereas the bottlenecks of GSP are multiple scanning of database and a huge set of candidate sequences is generated
2) *Prefix-Tree for Sequential Patterns (PSP) Algorithm:* This algorithm is built around GSP. In PSP [13], end user accesses the sequences that are graded in the web log with reference to IP address. The end-user is permitted to give a period of time “t” through which will access sequences which are temporarily adjoining to each other is assembled. With the help of prefix-tree, mining procedure can be handled that is similar to GSP. This algorithm is in a way similar to Apriori but it is executed without the Apriori-generate join. The database still has to be examined many times although this algorithm is more systematic in comparison with GSP.
3) *Sequential Pattern Discovery using Equivalence classes (SPADE) Algorithm:* This algorithm involves the highlights of a search space splitting in which the vertical database design is included by the inquiry search space [14]. In this algorithm, first lattice is decomposed into sub-lattices, with the goal that every sub-lattice can be totally handled utilizing either a depth-first or breadth- first pursuit strategy. The bitwise and logical operations are included in the counting of the candidate sequence method which is supported by SPADE. Directed exploratory outcomes demonstrate that SPADE is about twice as quick as GSP. It is because this algorithm utilizes a progressively systematic support counting method dependent on the idlest structure. Furthermore, a linear scalability is demonstrated by SPADE regarding the number of arrangements. Following examples elaborate the functioning of the SPADE algorithm. SPADE algorithm completes the mining in three passes of database scanning.

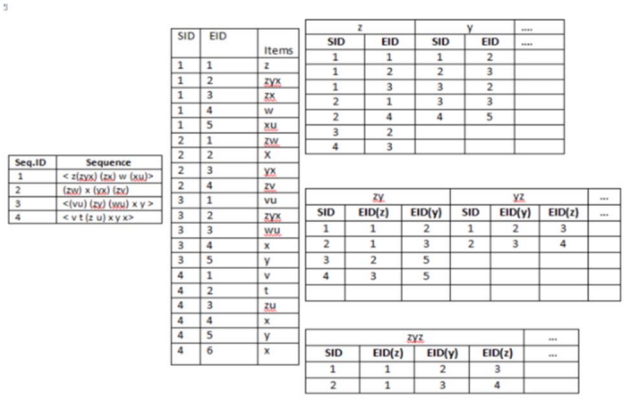

Fig. 1. An example of Sequential Pattern Discovery using Equivalence Classes (SPADE

4) *Sequential Pattern Mining (SPAM) Algorithm:* This is a latest strategy for discovering all the persistent sequences inside a transactional database. This mining is effective when the sequential patterns are very long in the database [15]. There is a strategy called as depth first search strategy which is used to engender candidate sequences and also many pruning mechanisms that are applied in order to diminish the search space. The data in the transactional database is stored which permits for effective support counting and also significant bitmap compression by using a vertical bitmap representation. A major property of SPAM algorithm is that this strategy progressively outputs latest frequent item sets which is in an online fashion.

5) *Sequential Pattern Mining with Regular Expression Constraints (SPIRIT) Algorithm:* In SPIRIT algorithm [16], there is involvement of regular expressions which are supposed to be used as a flexible constraint specification tool. It includes a generic user specified regular expression constraint on the mined examples, in this manner empowering significantly adaptable and powerful restrictions. So as to push the constraining inside the mining procedure, practically speaking the calculation utilizes a suitably loose, that is less prohibitive, variant of the limitation. There exist a few variants of the calculation, varying in how much the limitations are authorized to prune the inquiry space of pattern during calculation. Decision of regular expressions as a constraint specification tool is roused by two significant variables. To start with, REs gives a straightforward, regular linguistic structure for the compact determination of groups of successive patterns. Secondly, REs has adequate expressive control for indicating a wide scope of fascinating, non-insignificant design limitations.

B. Pattern-Growth Based Algorithm

This algorithm was emerged in the early 2000s, problem of generate-and-test. . This algorithm works on the idea according to which it escapes the candidate generation step altogether, it mainly concentrates on searching for a restricted portion of the initial database. One property which plays an important role in pattern-growth is search space partitioning.

1) *FreeSpan Algorithm:* The GSP algorithm proposed in [12], doesn't need to discover all the periodic item sets at the first place. GSP Algorithm is used in finding solutions of many issues present in sequence mining. It can be called as a level wise algorithm in which all the items that are frequent are generally seen level-wise and also all singleton items are computed. Then the items that are not frequent are removed. Finally, each and every transaction holds only the frequent items which it was holding initially. This database is sent to the GSP algorithm as an input. The GSP algorithm makes multiple passes. At first pass, it discovers the continual sequences which have the least support. At every pass, each sequence of data is analyzed to upgrade the occurrence number which is holding the sequence.The benefits of GSP is, it reduces the search space whereas the bottlenecks of GSP are multiple scanning of database and a huge set of candidate sequences is generated

2) *PrefixSpan Algorithm:* Prefix is a further structured algorithm based on projection, proposed by [17]. Observation of the prefix subsequences and their relating postfix subsequences, are anticipated into anticipated databases, as opposed to project sequencing database in Prefix algorithm. A direct implementation of the Apriori property is used so that the candidate sequences are reduced alongside anticipated databases. Furthermore, PrefixSpan is proficient in light of the fact that in Prefixspan, the total arrangement of patterns is mined and it has a fundamentally quicker execution than both, GSP calculation and FreeSpan. Development of anticipated databases is the significant cost of PrefixSpan which is similarly to FreeSpan. PrefixSpan needs to build an anticipated database at worst, for every consecutive database. After the projection of database is completed, usage of bi-level projection which is represented in FreeSpan. The utilization of successive

prefixes to divide the search space and to project arrangement databases is the fundamental thought of this algorithm. The focus is to look for the admissible sequences. The functioning of the PrefixSpan algorithm can be understood by the following example.

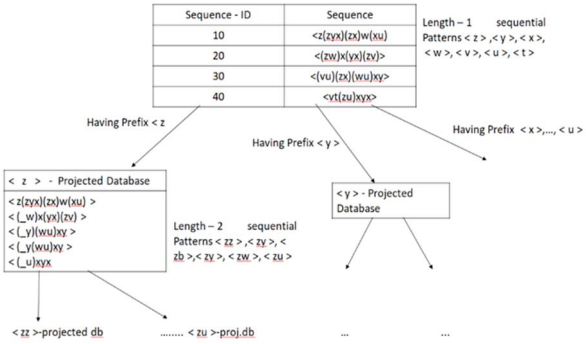

Fig. 2. An example of PrefixSpan Algorithm

3) *Web Access Pattern-Miner (WAP-Miner) Algorithm:* It is a tree structure-mining strategy alongside its WAP-tree structure [18]. An example development and tree structure-mining procedure utilized in WAP-mine calculation [13] with its WAP-tree structure. To build the WAP-tree from frequent groupings alongside their help, the sequence database is scanned only twice here. At first occurrence, a "header table" is kept up to point for each item in a frequent itemset, that is kept under observation in a threaded route afterwards so as to mine the tree for frequent sequences, expanding on the postfix. Frequent 1 sequence are found by the first scan of the database and the WAP-tree with only frequent sub-sequences built by the second scan.

4) *FS-Miner Algorithm:* Inspired by FP-tree and ISM [19], this is a technique which is based on tree projection design development concept that takes after WAP-mine and assists intelligent and gradual mining [20]. The noteworthiness of FS-Miner is that it begins mining promptly with 2 subsequences from the second (which is additionally the last sweep) of the database (at k = 2). It can do as such because of the packed portrayal in the FS-tree, which takes into consideration the header table of edges as opposed to single hubs and things, contrasted with WAP-tree and PLWAP-tree. It is likewise viewed as a variety of a tire, as it keeps bolster which are present in hubs similar to the edges of the tree that speaks to 2-groupings which is necessary for this steady mining process.

IV. Comparative Analysis of Sequential Pattern Mining Algorithms

A detailed comparative analysis is given in below Table1. The comparative analysis includes the key features of each algorithm with advantages and disadvantages.

TABLE I. A Comparative Analysis of Various Sequential Pattern Mining Algorithms

Algorithm	Key Feature	Advantage	Disadvantage
GSP	Apriori based approach, uses downward closure property	Reduces the search space	Database is scanned multiple times
PSP	End- user accesses the sequences that are graded in the web log with reference to the IP address	More systematic in comparison with GSP	Database still has to be examined many times
SPADE	Vertical format sequential pattern mining method	2 times faster than GSP	Multiple database scans
SPAM	Discovers all the persistent sequences inside a transactional database	Effective when the sequential patterns are very long in the database	Complete algorithm with its data structure fits in main memory.
FreeSpan	Pattern Growth based method	Effective as it is mines the complete set of patterns	Costly
PrefixSpan	Structured algorithm based on projection	Finds the frequent items after scanning the sequence	Cost of memory space might be high
WAP-Miner	Tree structure-mining strategy	Sequence database is scanned only twice	Tree structure may use more memory
FS-Miner	Intelligent and gradual mining	Steady mining process	Extra memory consumption by header table.

V. Reasearch Domain Address by Sequential Pattern Mining

A. Health and Medical Domain

Sequential pattern mining is a compelling method to recognize transient connections among medications and can be utilized to foresee subsequent stages in a patient's medicine routine. Precise predictions can be made without utilizing the patient's whole drug history.

Right now, acquired cases information for patients endorsed at any rate one diabetes medicine between year 2008 and 2011, and partitioned these into a training set and test set. After this author applied the CSPADE algorithm to mine sequential patterns of diabetes medicine solutions both at the medication

class and nonexclusive medication level and positioned them by the support statistic [21]

B. ICT Risk Assessment and Management

ICT chance appraisal and the executives depend on the examination of information on the joint conduct of an objective framework and its aggressors. The devices in the Haruspex suite model wise, objective situated aggressors that arrive at their objectives through arrangements of assaults. The instruments artificially produce these arrangements through a Monte Carlo technique that runs numerous reproductions of the aggressor conduct.

This paper exhibits a successive example mining examination of the assault arrangement database to separate an elevated level and compact comprehension of the aggressor methodologies against the framework to evaluate. Such a comprehension is communicated as a lot of successive examples that spread, and potentially segment, the assault arrangements. [22]

C. Behavior Model

Understanding the human conduct during the creation procedure would be a significant issue for satisfying by and large operational greatness in software development. Among approaches proposed to find the human conduct dependent on the succession exercises, process mining is one of which has gotten considerations of late. While latest procedure mining approaches in the space of human conduct address process revelation and post-examination, not many of them have paid considerations on pre-investigation. The pre-examination is one of the approaches to create a dependable and high-calibre of occasion log which intentionally impacts on finding a day by day normal conduct and dismissing unpredictable successive conduct.

This study intends to propose another method for pre-examination utilizing sequential pattern mining. The key commitments of this examination first, is to decide the potential neighbourhood practices utilizing sequential pattern mining considering time requirement.[23]

VI. Research Challenges in Sequential Pattern Mining

At present time, various strategies are available for uncovering sequential patterns with respect to the initial definition. These types of patterns are immensely befitting for a huge number of applications. Yet there still are several research challenges in this field of mining of data. Some of them are:

- To be tremendously systematic, scalable, only a small number of database scans are included.
- To be able to engulf several types of user-specific constraints. [24]
- Algorithm must be able to tackle large search space.
- Algorithm must abstain from redoing inspection of database during the course of process of mining.
- To examine constraints like Frequency and Monetary constraints and also scan their result with respect to time of execution, usage of memory and scalability.
- To review target aligned sequential pattern mining and also its application in some real dataset.[25]
- To initiate the concept of object-aligning in this sequential pattern mining, through which there can be a flexibility of mining, focused fragments of the database.
- For a larger set of databases there can be a probability of having dispersed sequential pattern mining in order to impart scalability.

Some more research opportunities suggested by [26] are -

- Develop new algorithms those are faster and memory efficient
- Design new interactive algorithms those are more user friendly.
- Proposed new algorithms that support more constraints and more visualization power.
- Design and develop new pattern mining activities with new challenges,
- Offering new applications for established algorithms,
- Recommending modifications of existing issues that support big data mining trends, use parallel architectures, etc.

VII. Future Trends in Sequential Pattern Mining

Today few techniques are accessible for effectively finding sequential patterns according to the underlying definition. Patterns of such type are generally relevant for numerous applications. Particular strategies, generally roused from past algorithms, exist in a wide range of areas. Today, according to the colossal volume of information accessible, stream data mining showcases a rising class of data-concentrated applications where the data streams in and out progressively. These types of applications additionally

require quick responses. To grow the functionality of sequential rules, it is foremost important to consider more data. Thus, by linking sequential patterns with the category of client, the aim of sequential pattern mining is to give the end-user with more prominent patterns. Some potential application domain where in future the sequential pattern mining will play an important role includes sequential pattern mining with uncertain data [27], high utility item set mining, determining medical prescription [28] or pathology test for patient, activity pattern of tourist [29] or some campus placement pattern of the students. Some kind of recommendation algorithm and user behavior analysis might another hot area where sequential pattern mining play a vital role.

VIII. CONCLUSION

Although the idea of Sequence Data Mining is new but it has attained substantial progress in recent years. Various approaches that are concerned by sequential pattern mining have been lodged to deal with the effectiveness of the algorithms improvement either with the help of latest structures, latest approaches or by the database management in the computer memory. Thus, based on the proposed norms, this paper classifies sequential pattern mining into two major classes – Apriori and pattern-based algorithms. There are certain requirements which we have to be examined for a veritable sequential pattern-mining algorithm. Firstly, a search space should be created by a method which must be as small as possible. The features that allow this contain pruning of early candidate sequence and partitioning of search space. Secondly, it is essential that this search space is narrowed within the search space. There also exist an algorithm which can have a narrow searching procedure like depth-first search. Thirdly, techniques other than tree projection should be examined for discovering the authentic techniques sequential pattern-mining. This paper also describes the future trends and research challenges in sequential pattern mining.

REFERENCES

[1] H. K. Soni, S. Sharma, and M. Jain, “Frequent pattern generation algorithms for Association Rule Mining : Strength and challenges,” in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, 2016, pp. 3744–3747.

[2] N. Mishra, H. K. Soni, S. Sharma, and A. K. Upadhyay, “A Comprehensive Survey of Data Mining Techniques on Time Series Data for Rainfall Prediction,” J. ICT Res. Appl., vol. 11, no. 2, pp. 167–183, 2017.

[3] N. Mishra, H. K. Soni, S. Sharma, and A. K. Upadhyay, “Development and Analysis of Artificial Neural Network Models for Rainfall Prediction by Using Time-Series Data,” Int. J. Intell. Syst. Appl., vol. 10, no. 1, pp. 16–23, Jan. 2018.

[4] M. Kumar and H. K. Soni, “A Comparative Study of Tree-Based and Apriori-Based Approaches for Incremental Data Mining,” Int. J. Eng. Res. Africa, vol. 23, pp. 120–130, Apr. 2016.

[5] H. K. Soni, S. Sharma, and M. Jain, “Plausible characteristics of association rule mining algorithms for e-commerce,” in 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), 2017, pp. 36–39.

[6] R. Agrawal, T. Imieliński, and A. Swami, “Mining association rules between sets of items in large databases,” in Proceedings of the 1993 ACM SIGMOD international conference on Management of data - SIGMOD ’93, 1993, pp. 207–216.

[7] H. K. Soni, “Peculiarly Approaches for Association Rule Mining using Genetic Algorithm,” in 4th International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB-18), 2018, pp. 227–232.

[8] H. K. Soni, S. Sharma, and U. A.K, “Two Novel Pioneer Objectives of Association Rule Mining for high and low correlation of 2-varibales and 3-variables,” Int. J. Eng. Technol., vol. 9, no. 2, pp. 695–703, Apr. 2017.

[9] H. Soni, S. Sharma, and A. Upadhyay, “An Empirical Algorithm for High and Low Correlative Association Rule Mining,” Int. J. Intell. Eng. Syst., vol. 11, no. 3, pp. 223–232, Jun. 2018.

[10] J. Han, J. Pei, B. Mortazavi-Asl, Q. Chen, U. Dayal, and M.-C. Hsu, “FreeSpan,” in Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’00, 2000, pp. 355–359.

[11] N. R. Mabroukeh and C. I. Ezeife, “A taxonomy of sequential pattern mining algorithms,” ACM Comput. Surv., vol. 43, no. 1, pp. 1–41, Nov. 2010.

[12] R. Srikant and R. Agrawal, “Mining sequential patterns: Generalizations and performance improvements,” 1996, pp. 1–17.

[13] Masseglia, F., Poncelet, O., and Cicchetti, R. 1999. An efficient algorithm for web usage mining. Network Inform. Syst. J. 2, 571—603

[14] Zaki, M. J., 2001. "Spade: An efficient algorithm for mining frequent sequences". In: Machine Learning. Vol. 42. pp. 31–60.

[15] J. Ayres, J. Flannick, J. Gehrke, and T. Yiu, “Sequential PAttern mining using a bitmap representation,” in Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’02, 2002, p. 429.

[16] Minos N. Garofalakis, Rajeev Rastogi, and Kyuseok Shim. 1999. SPIRIT: Sequential Pattern Mining with Regular Expression Constraints. In Proceedings of the 25th International Conference on Very Large Data Bases (VLDB '99), Malcolm P. Atkinson, Maria E. Orlowska, Patrick Valduriez, Stanley B. Zdonik, and Michael L. Brodie (Eds.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 223-234.

[17] Jian Pei et al., “PrefixSpan,: mining sequential patterns efficiently by prefix-projected pattern growth,” in Proceedings 17th International Conference on Data Engineering, pp. 215–224.

[18] J. Pei, J. Han, B. Mortazavi-asl, and H. Zhu, “Mining Access Patterns Efficiently from Web Logs,” 2000, pp. 396–407.

[19] S. Parthasarathy, M. J. Zaki, M. Ogihara, and S. Dwarkadas, “Incremental and interactive sequence mining,” in Proceedings of the eighth international conference on Information and knowledge management - CIKM ’99, 1999, pp. 251–258.

[20] M. El-Sayed, C. Ruiz, and E. A. Rundensteiner, “FS-Miner,” in Proceedings of the 6th annual ACM international workshop on Web information and data management - WIDM ’04, 2004, p. 128.

[21] A. P. Wright, A. T. Wright, A. B. McCoy, and D. F. Sittig, “The use of sequential pattern mining to predict next prescribed medications,” *J. Biomed. Inform.*, vol. 53, pp. 73–80, 2015.

[22] M. D’Andreagiovanni, F. Baiardi, J. Lipilini, S. Ruggieri, and F. Tonelli, “Sequential pattern mining for ICT risk assessment and management,” *J. Log. Algebr. Methods Program.*, vol. 102, pp. 1–16, 2019.

[23] F. Setiawan and B. N. Yahya, “Improved behavior model based on sequential rule mining,” *Appl. Soft Comput. J.*, vol. 68, pp. 944–960, 2018.

[24] J. Pei, J. Han, and W. Wang, “Constraint-based sequential pattern mining: the pattern-growth methods,” J. Intell. Inf. Syst., vol. 28, no. 2, pp. 133–160, Mar. 2007.

[25] H.-E. Chueh, “Mining Target-Oriented Sequential Patterns With Time-Intervals,” Int. J. Comput. Sci. Inf. Technol., vol. 2, no. 4, pp. 113–123, Aug. 2010.

[26] http://data-mining.philippe-fournier-viger.com/the-future-of-pattern-mining/
[27] http://data-mining.philippe-fournier-viger.com/the-future-of-pattern-mining/
[28] A. P. Wright, A. T. Wright, A. B. McCoy, and D. F. Sittig, "The use of sequential pattern mining to predict next prescribed medications," J. Biomed. Inform., vol. 53, pp. 73–80, Feb. 2015.
[29] A. Talpur and Y. Zhang, "A Study of Tourist Sequential Activity Pattern through Location Based Social Network (LBSN)," in 2018 International Conference on Orange Technologies (ICOT), 2018, pp. 1–8.