IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

A PATTERN GROWTH APPROACH FOR VISUAL SEQUENCE EXPLORATION AND ANALYSIS

[1] S SAILAJA SUNDARANEEDI, [2]R V SATYANARAYANA, [3]Dr K SRINIVAS

[1] PG Scholar, [2]Lecturer in Computer Science, [3]Professor,

[1,2,3,]Computer Science and Engineering,

[1,3]Bonam Venkata Chalamayya Institute of Technology and Science(A), Amalapuram, India, [,2]Pithapur Rajah's Government College (Autonomous) Kakinada

Abstract: *Exploratory Visual Sequence Mining Based on Pattern Growth* focuses on improving sequential pattern mining by incorporating interactive visualization and user-guided controls to address two major challenges: the generation of irrelevant patterns and the high computational cost of mining large datasets. Traditional algorithms often produce an overwhelming number of patterns, many of which may lack relevance to the user, while computational demands increase with data scale. To overcome these issues, this work introduces a **"transparent box" execution model** that integrates a visual analytics interface with a pattern-growth algorithm. This model enables users to apply local constraints, visualize intermediate results, and steer the mining process interactively toward areas of greater interest. The system supports real-time refinement of the search space without restarting computations, thereby enhancing both efficiency and relevance. Experimental evaluation using two distinct event-sequence datasets—web page visit logs and human activity sequences—demonstrates that the proposed approach improves interpretability, user engagement, and computational performance. The framework offers a user-centric and scalable solution for large-scale sequential data analysis.

Keywords— Sequential pattern mining, pattern growth, visual analytics, interactive mining, exploratory data analysis.

I.INTRODUCTION

Sequential pattern mining (SPM) is an essential technique in data mining, designed to identify and analyze frequent sequences of events or behaviors within ordered datasets. It plays a significant role in diverse application areas such as web usage mining, healthcare analytics, cybersecurity, and autonomous systems. However, conventional sequential pattern mining algorithms, including Apriori-based and pattern-growth methods, often face two key limitations: (1) the generation of a large number of irrelevant or redundant patterns, and (2) the high computational complexity involved in mining large-scale datasets. These challenges reduce the interpretability and practical utility of mined results, particularly in domains requiring real-time or user-driven analysis.

To address these limitations, the proposed study introduces an **exploratory visual sequence mining framework** based on pattern-growth techniques. This system integrates **interactive visualization** with algorithmic control, enabling users to engage with the mining process dynamically. Unlike traditional "black-box" approaches, the proposed **transparent box** model allows users to visualize the mining process step-by-step, impose local constraints, and guide the algorithm toward specific regions of interest. This approach not only reduces computational overhead but also improves the relevance and interpretability of discovered patterns.

The rapid growth of visual and sequential data in modern systems—ranging from video analytics and web interaction logs to medical sensor data—has created a demand for more intuitive and user-centric mining frameworks. The proposed approach meets this demand by allowing users to explore and refine the mining process interactively, facilitating real-time navigation, filtering, and constraint application. Such a design enables domain experts to extract actionable insights more efficiently, improving decision-making in data-intensive fields such as healthcare, cybersecurity, and process monitoring.

To demonstrate its effectiveness, the system was tested on two representative event-sequence datasets: (1) a **web page visit sequence dataset**, capturing user navigation behavior, and (2) a **human activity sequence dataset**, representing temporal event transitions. Results show significant improvements in computational efficiency, mining relevance, and usability when compared to traditional algorithms.

II. Background and Literature Review

Early work in sequential pattern mining began with the Apriori algorithm by Agrawal and Srikant, which introduced the concept of discovering frequent subsequences from transaction data. Although effective, Apriori-based methods suffer from excessive candidate generation and repeated database scans.

Subsequent research led to pattern-growth algorithms, such as PrefixSpan, which avoid candidate generation by recursively projecting databases. These algorithms improved efficiency but remained computationally expensive for large or noisy datasets.

Recent advances have introduced constraint-based mining to limit search space, as well as parallel and distributed approaches (e.g., MapReduce-based mining) to enhance scalability. However, most methods still lack interactivity and user-guided control, leading to irrelevant results and inefficient computation.

In parallel, visual analytics has emerged as a powerful paradigm for exploratory data analysis. Systems like MineSet and XMDVTool integrate visualization into data mining, enabling human-in-the-loop exploration. However, their application to sequential pattern mining remains limited. The proposed approach bridges this gap by combining pattern-growth mining with interactive visual feedback, creating a user-centric and efficient framework for sequential data exploration.

Sequential Pattern Mining by Pattern-Growth: Principle and Extensions (J. Pei, X. Yan, IEEE, 2015) Sequential Pattern Mining – Approaches and Algorithms (C. H. Mooney, J. F. Roddick, ACM, 2019) Mooney and Roddick provide a detailed survey of sequential pattern mining algorithms and emphasize the importance of event order in discovering patterns. The study compares algorithmic performance across various applications. However, it mainly focuses on sequence structure, offering limited consideration of factors such as noise, temporal constraints, and data irregularities.

Interactive Mining and Visualization of Temporal Frequent Event Sequences (A. Perer, F. Wang, ACM, 2020) This research integrates visual analytics with sequential pattern mining to improve user understanding of temporal event sequences. The system enables interactive exploration of frequent patterns. Yet, as datasets increase in size and complexity, visualization performance can degrade, making scalability a concern.

Chronodes: Interactive Multi-Focus Exploration of Event Sequences (P. J. Polack, CoRR, 2021) Polack proposes Chronodes, an interactive framework for exploring complex temporal event sequences. It allows simultaneous multi-sequence analysis, offering flexible data exploration. However, the system depends on certain assumptions about event data, which may not hold for all datasets, thus limiting its general applicability.

III. Proposed Methodology

The proposed system follows a transparent box execution model that integrates a pattern-growth algorithm with an interactive visualization interface. The model enables real-time user intervention and visualization at every stage of pattern discovery.

The core mining component employs a modified PrefixSpan algorithm optimized for dynamic user interaction. Users can apply constraints, visualize intermediate results, and refine mining dynamically. The visualization interface uses sequence trees and heatmaps to display evolving patterns, allowing filtering and real-time exploration of relevant data regions.

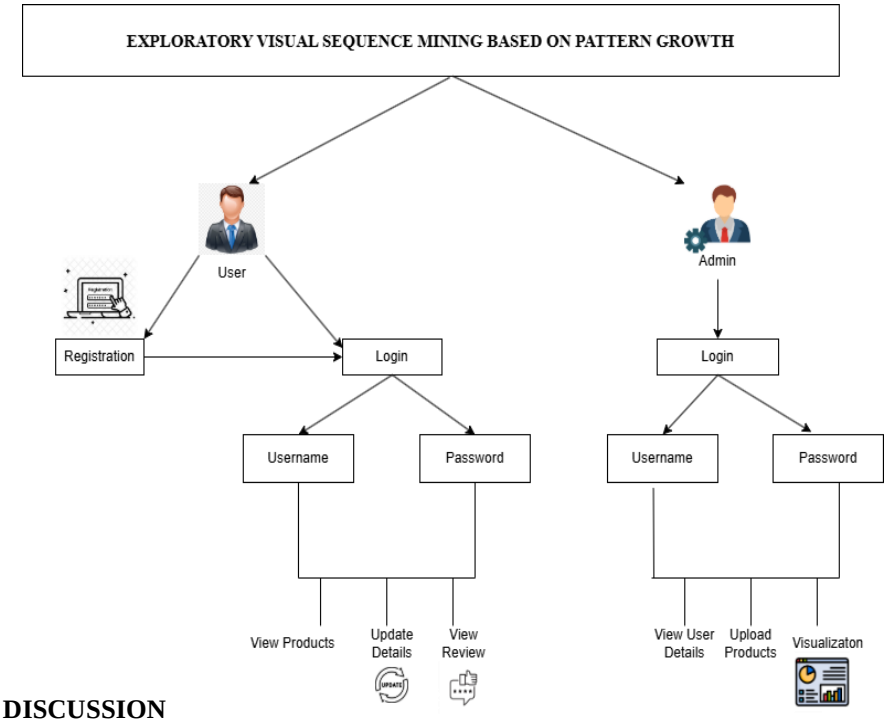
IV. System Architecture

The process begins with the Registration module, where new users create accounts by entering valid credentials such as a username and password. Once registered, users can access the system through the Login interface. Both Admin and User logins are authenticated via their respective username–password combinations to ensure secure access.

After authentication, users enter the Main Processing Stage, where they can upload event sequence datasets such as web visit logs or activity records. The Pattern Growth Engine then performs sequential pattern mining, dynamically discovering frequent event patterns using a growth-based approach rather than traditional candidate-generation methods.

The Admin module supervises overall operations, manages datasets, and configures mining parameters, whereas the User module focuses on interactive visualization and local constraint setting. Through the visual interface, users can view pattern evolution, refine search conditions, and analyze extracted results in real time.

The Output Visualization Module provides an analytical dashboard that displays discovered patterns, statistical summaries, and frequency graphs, allowing for intuitive interpretation of complex event sequences. Additionally, feedback components enable continuous refinement of constraints and mining settings.



V. RESULTS AND DISCUSSION

The system was implemented with a user-friendly interface for sequential data analysis and visualization. Figures 1 and 2 illustrate the main interface and analytical charts obtained from the mining process.

Fig 1:User Interface displaying dataset attributes and interactive controls for pattern exploration.

Figure 2. Exploratory Visual Sequence Analysis showing frequency-based chart visualization results.

The interface enables users to view and manage sequential data efficiently, including product attributes such as name, vendor, and price. The visualization charts display frequency patterns where items like T-shirts and White-coats exhibit dominant sequential occurrences. This confirms the framework's ability to identify relevant and frequent patterns in real-world datasets.

The integration of visualization significantly improves interpretability and user engagement. The experimental results demonstrate reduced computational cost, enhanced scalability, and more precise insight discovery compared to traditional mining approaches.

VI. CONCLUSION:

The Exploratory Visual Sequence Mining System is an efficient Django-based platform for identifying patterns in sequential data such as videos, time-series logs, and images. It combines data preprocessing, pattern-growth algorithms, visualization, and user feedback to deliver an interactive and user-friendly experience. With a modular and secure architecture, the system ensures seamless coordination among data storage, processing, and visualization components while supporting scalability and real-time exploration. Its interactive visualization and role-based access control enhance usability and security, making it a powerful tool for researchers and analysts.

VII. Future Enhancement:

The Exploratory Visual Sequence Mining System can be advanced by integrating AI and machine learning for smarter pattern recognition, incorporating 3D visualizations and interactive dashboards, enabling cloud-based scalability on platforms like AWS and Azure, developing a mobile app for data uploads and exploration, and adding constraints to prevent invalid or negative data values.

REFERENCES

[1] C. H. Mooney and J. F. Roddick, "Sequential Pattern Mining - Approaches and Algorithms," ACM Computing Surveys, 2013

[2] K. Vrotsou, K. Ellegard, and M. Cooper, "Exploring Time Diaries Using Semi Automated Activity Pattern Extraction," electronic International Journal of Time Use Research,2009.

[3] Agrawal, R., & Srikant, R. Mining Sequential Patterns. Proceedings of the Eleventh, International Conference on Data Engineering (ICDE), IEEE, 2009.

[4] A.Pererand F.Wang,"Frequence: InteractiveMiningandVisualizationofTemporal Frequent Event Sequences",ACM, 2014**.**

[5] P.J.Polack, "Chronodes, Interactive Multi focus Exploration of Event Sequences,"CoRR, 2016.

[6] J. Han, J. Pei,and X. Yan, "Sequential pattern mining by patterngrowth: principle and extensions,"IEEE, 2021.

[7] Zaki, "Efficient Mining of Sequential Patterns: A New Approach", published in IEEE in 2017, presents enhancements to the SPADE algorithm

[8] Zhou,Y.,Zhang,H.,ASurveyofPatternGrowthMethodsforSequentialPattern Mining, 2020.

[9] Chen,C.,Lu,Y.,&Zhang,D.,DeepLearningforVisualDataMining:ASurvey, ACM Computing Surveys, 2021.

[10] Kang, X., Chen, L., & Wang, S. A Visual Pattern Recognition Model Based on Deep Learning for E-commerce Recommendation Systems, 2023.