Abstract

Web data mining is an important research topic because it attains a significant amount of interest from both academic and industrial environments. Web sequential pattern mining is an imperative for analyzing the access behavior of web users. Recently, few research works have been designed for mining the web sequential patterns. However, performance of existing techniques was not effectual. In order to overcome such limitation, Hilbert Space clustering based Chronological Backward Search (HSC-CBS) Technique is proposed. HSC-CBS Technique is designed in order to improve the performance of web sequential patterns mining. The HSC-CBS Technique at first used Hilbert space clustering in order to group the similar user’s interest web patterns in web log database which resulting in improved clustering accuracy. The clustering of frequent web patterns in web log database helps for minimizing the space and time complexity of web sequential pattern mining. After clustering, HSC-CBS Technique applied chronological backward search algorithm in order to efficiently mine the web sequential patterns and improving true positive rate of web sequential pattern mining. The HSC-CBS Technique conducts the experimental works on the parameters such as execution
time, space complexity, clustering accuracy, true positive rate of mining and scalability. The experimental results show that the HSC-CBS Technique is able to improve the true positive rate of pattern mining and also reduces the execution time as compared to state of the art works.

References


12. Xiuming Yu, Meijing Li, Kyung Ah Kim, Jimoon Chung and Keun Ho Ryu, “Emerging Pattern-Based Clustering of Web Users Utilizing a Simple Page-Linked Graph”, Sustainability, Volume 8, Pages 1-18, 2016


Hilbert Space Clustering based Chronological Backward Search for Effective Web Sequential Pattern Mining


21. Amazon Commerce reviews set Data Set: https://archive.ics.uci.edu/ml/datasets/Amazon+Commerce+reviews+set

Index Terms

Computer Science | Pattern Recognition

Keywords

Chronological backward search, Hilbert Space clustering, mining, web user, web log database, web pattern.