Using Petri Nets to Enhance Web Usage Mining[1]

Shih-Yang Yang

Department of Information Management
Kang-Ning Junior College of Medical Care and Management
Nei-Hu, 114, Taiwan
Shihyang@knjc.edu.tw

Po-Zung Chen, Chu-Hao Sun

Department of Computer Science and Information Engineering
Tamkang University
Tamsui, 251, Taiwan
pozung@mail.tku.edu.tw, steven.sun@fubon.com

Abstract: Precise analysis of the web structure can facilitate data processing and enhance the accuracy of the mining results in the procedure of web usage mining. Many researchers have identified that pageview identification and path completion are of great importance in the result of web usage mining. Currently, there is still a lack of an effective and systematic method to analyze and deal with the two steps.
In the present study, we propose the application of Petri Nets (PN), a model used to analyze the framework of webpages in a website. We adopt Place in the PN model to represent webpage on the websites and use Transition to represent hyperlink. The study explores how to undergo the pageview identification after we use the PN model to conduct the analysis of the framework and then obtain incident matrix. Likewise, we use reachability property in the model to undergo path completion.

Keywords: Petri nets, Web usage mining, Data Preprocessing

1 Introduction

This study introduces a method to enhance a web usage mining using Petri Nets (PN) in modeling a web structure. We also observe that PN can help resolve pageview identification and path completion, particularly in a complex webpage comprising many frames in one single pageview.

[1] This work is supported by the National Science Council, Taiwan, ROC, under Grant # NSC 93-2213-E-032-016.