

# WEB PAGE ACCESS PREDICTION BASED ON MARKOV MODEL AND PAGE RANK ALGORITHM

Phyu Thwe

*University of Technology (Yatanarpon Cyber City), Pyin Oo Lwin, Myanmar*

*pthwe19@gmail.com*

## **Abstract**

The Web now becomes one of the main sources of the information. At the same time, it also makes the user to get lost in the millions of information. One way to assist the users who need their applicable information is to predict a user's future request and use the prediction for pre-fetching, caching and recommendation. The purpose of this paper is to explore ways to exploit the information from web logs for predicting users' web page access.

When the user browses the web pages, user leaves some valuable information in web log. This web log information is very helpful to find out web navigation behavior of user. Through this behavior we can find out what kind of information user wanted from the web sites. Web usage mining has assumed importance in learning about web user's behavior and user interactions with the website. It uses data mining techniques to discover non-trivial user behavior patterns. These patterns can then be used to make the predictions of next page to be accessed by the user. We proposed to use Markov model to determine next page accesses on a particular website. Markov models are commonly used in the identification of the next page to be accessed by the user based on the previously accessed pages. Then, we use popularity and similarity-based page rank algorithm to make prediction when the ambiguous results are found. Page Rank is a numeric value that represents how important a page is on the web. When one page links to another page, it is effectively casting a vote for the other page. The more votes for a page, the more important the page must be.

Web page access prediction can be useful in many applications. The improvement of accuracy can make a change in the web advertisement area. Using web page access prediction, the right advertisement will be added according to the users' browsing patterns. Also, it helps web administrators restructure the websites to improve site topology and user personalization as well as market segmentation. It is also helpful for caching the predicted page for faster access and for improving browsing and navigation orders.

