Expectation maximization clustering algorithm for user modeling in web usage mining system

ABSTRACT

To provide intelligent personalized online services such as web recommender systems, it is usually necessary to model users' web access behavior. To achieve this, one of the promising approaches is web usage mining, which mines web logs for user models and recommendations. Web usage mining algorithms have been widely utilized for modeling user web navigation behavior. In this study we advance a model for mining of user's navigation pattern. The model is based on expectation-maximization (EM) algorithm and it is used for finding maximum likelihood estimates of parameters in probabilistic models, where the model depends on unobserved latent variables. The experimental results represent that by decreasing the number of clusters, the log likelihood converges toward lower values and probability of the largest cluster will be decreased while the number of the clusters increases in each treatment. The results also indicate that kind of behavior given by EM clustering algorithm has improved the visit-coherence (accuracy) of navigation pattern mining.