IMPLEMENTATION OF CARD+ ALGORITHM FOR
INTERNET USER CATEGORIES IDENTIFICATION IN
ITS SERVER

Student’s Name : NOOR FITRIA AZZAHRA
Student’s ID : 5108 100 025
Department : Informatics, FTIF-ITS
First Advisor : Diana Purwitasari, S.Kom., M.Kom.
Second Advisor : Abdul Munif, S.Kom., M.Kom.

Abstract

Nowadays, the spread of information via the internet is growing fast. This increases the need for internet. ITS is one of the largest university in Indonesia is also experiencing this. Therefore we have to be evaluate the use of internet in order to find user behaviors in using the ITS. One of the evaluation methods that can be applied is personalization. Personalization is an evaluation method that mine user behavior based on server proxy profile of the user. In order to solve problem of session dimension that are too large, the data is converted into a relational data from one session to another session.

In this final project we use CARD+ (Competitive Agglomeration Relational Data) algorithm for doing personalization. CARD+ is a modification of CA (Competitive agglomeration) algorithm for relational data. CA combines the advantages of two clustering, i.e. fuzzy c-means (FCM) in achieving the minimum value of objective function and hierarchical clustering in getting the optimal number of clusters. CARD+ is a development of CARD algorithms with a post-clustering stage that merge overlapping clusters.

In this final project, the identification is based on the interest degree of the user to the URL. The density f clusters for CARD and post-clustering stage will be calculated by Dunn index. The
experiment results Dunn index after post-clustering is better than before this stage. So that the CARD+ can be implemented for identification user categories of user ITS server.

*Keywords: log file, session, behaviour matriks, relational data, CARD+, post-clustering, interest degree, Dunn index*