Abstract

Nowadays, spam in email is starting to use image media. Spam that once use to be in a text form, is beginning to change into an image form so that it can't be unidentified by an anti-spam filter. In 2007, almost 50% of email spam is made in a form of an image and it is increase by 5% every month.

In this final project, software was build using Microsoft Visual C# to detect image spam with the use of a method called decision tree C4.5. This method is one of the many learning method that classified the data from the training data. Image information that is made into an attribute predictor can come from a file attribute, file header, and histogram image. The learning result is a rule in a form of a tree.

Analysis process is done by finding the most reliable rule model that can classified an image spam and ham. The rule model is produce in the training process from some training data and will be tested and compared with other data, so that it accuracy can be calculated base on the rule model.

Keyword: Decision Tree C4.5, Image Spam filter, Histogram Image