

Title	決定木を用いた質的データからのグラフ構造学習
Author(s)	川崎, 隆史
Citation	
Issue Date	2014-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/11980
Rights	
Description	Supervisor: Dam Hieu Chi, 知識科学研究科, 修士

Graph Structure Learning from Qualitative Data by Decision Tree

Takafumi Kawasaki

School of Knowledge Science
Japan Advanced Institute of Science and Technology
March 2014

Keywords: Decision Tree, Graph Structure Learning, Qualitative Data.

Recently, the number of studies has been increasing in which a knowledge discovery by means of graph and network was motivated. A community extraction from SNS and prediction for link of protein reciprocal action are examples of them.

The most important issue in this field is how to determine a graph structure because the information of the analytic target is needed to be contained in the graph.

N.Meinshausen and P.Bühlmann proposed a graph structure learning method in 2006 (MB-method). A graph structure is obtained by multiple regression analysis which is applied for all variables of data. It succeeded in extracting a graph structure which is composed by important relationships from data by using Lasso regularization method for variable selection.

MB-method employs framework of actual value and regression analysis motivated by quantitative analysis, and doesn't appropriate for analyzing qualitative data which is composed by discrete value. As a qualitative data-oriented method, there was logistic regression analysis modified by the strategy of MB-method. However, the major problem is that logistic regression doesn't examine relations between explanatory variables to a response variable.

In this study, I propose the graph structure learning method considering relations between explanatory variables. It employs decision tree which fits both for qualitative data processing and for examination of relations between explanatory variables. The backward elimination method was used for variable selection. A graph structure learning algorithm was modified by the idea of MB-method.

The experiment were executed to validate the proposed method. The backward elimination method was effective to decrease the number of variables in selection process of generating decision trees. The graph created by this method shows a relationship which reflects the characteristic of each variables, suggesting our graph is consistent with common sense. The graph obtained in this method principally contains Hi-dimensional complex information. In contrast to the graph by the relation analysis which is to be build by pear to pear data, own method is suggested to has a advantage in knowledge discovery.