URN etd-0407109-183716-1315
Author Hsiao Wei Wang
Author's Email Address No Public.
Statistics This thesis had been viewed 39 times. Download 0 times.
Department
Year 2005
Semester 2
Degree Master
Type of Document Master's Thesis
Language zh-TW.Big5 Chinese
Title Using Efficient Algorithms to Cluster Transaction Data
Date of Defense 0000-00-00
Page Count 60

Keyword
data mining
cluster
transaction data
Boolean computation
similarity function

Abstract
Along with the information technology development, the enterprise may the record, store up consumer's transaction data easily and analyze consumer's expense tendency. To improve relations between the enterprise and customer, promotes customer's loyalty and satisfaction and expands the market advantage base all to have the remarkable influence. Data mining may widely apply in each domain, the correlation research of the cluster always is the recent for several year quite important research topic of the data mining area, after clustering the object, then may the characteristic which reveals for each cluster institute carry on the further the analysis. At present most often uses cluster algorithm（ex: PAM or K-means） all is calculates distance of the object and the central point of the cluster （Euclidean distance）to do for the ownership basis of the cluster.
In this thesis, we in view of the transaction databases in each transaction data, aim at two aspects to mining. First, using the PAM algorithm of the cluster technology to make the traditional string compared and Boolean expression, then all transaction data will cluster according to the similarity, will favor we to understand the characteristic of each cluster. The use cluster support is using to express the scale of this cluster, the transaction item confidence is using to express in the identical cluster, a transaction contains relations between the items, these two kinds of measures appraised each cluster after clustering. Next, uses the hierarchical agglomerative algorithm to make traditional string compared and Boolean expression, each transaction data regards as one cluster, if satisfies the similarity function biggest, then these two cluster to make the merge until finally can form one cluster. After the computation, looks for the characteristic of the transaction item of the cluster in each cluster. For the enterprise when makes the decision, the forecast and the personalized recommendation analysis may provide help. Finally, this research basis on proposing the method, hoped can penetrate the method which Boolean expression, in experiment for doing the traditional and improvement algorithms to appraise, hoped will be able to improve and accelerate the efficiency and time of data mining.
Advisory Committee Cheng Chui Chen - advisor

Files No Any Full Text File.
Date of Submission 2006-06-28