Abstract

Clustering is an important task for any recommendation system. Clustering methods are suggested by many researchers for search engine optimization. Search engines help users for better searching by user's query recommendation. Clustering is helpful for finding actual relation between different queries which are not the same as they seem. But do clustering of user query is also a difficult task because of user enters lots of type and varying queries. Many times these queries may very short to get their real meaning and also can generate different meanings. Any single query may have various meaning on one hand many different query words may have common meaning for searching contents. Lots of clustering methods are given in last decades for search engine optimization but these methods unable to proper utilization various information hidden in user query log. This paper gives a novel clustering approach based on to identify query similarity and apply SOM clustering for effective clustering results. We propose a novel similarity matrix for user queries by uses of URL clicked by user through searching results. Text similarity and time similarity are also measure for calculating similarity between two queries. This method shows good results within clustering performance to compare with other existing methods.
Top-K Search Query Grouping using SOM Clustering for Search Engine

References

- Dr. G. K. Gupta, "Introduction to Data Mining with Case Studies", PHI Publication, 2005.

**Index Terms**

Computer Science

Algorithms

**Keywords**

Query Logs  Query Process  SOM Clustering