

Article Abstract

Title:	Constraint based frequent pattern mining for generalized query templates from web log
Author(s):	Ramachandra. V. Pujeri ¹ , G.M. Karthik ²
Address(es):	¹ KGiSL Institute of Technology, Coimbatore, Tamil Nadu, INDIA e-mail:sriramu_psg@yahoo.com, Tel +91-98431-20515, +91-422-6619929 ² Department of Computer Science and Engineering, SACS MAVMM Engineering College, Madurai, Tamil Nadu, INDIA e-mail:gmkarthik16@gmail.com, Tel +91-94432-85673, +91-452-2389873
Journal:	<i>International Journal of Engineering, Science and Technology</i> , Vol. 2, No. 11, 2010, pp. 17-33.
Abstract:	The World-Wide Web provides every Internet citizen access to an abundance of information, but difficulty increases in identifying the relevant piece of information. Popular Search engine uses log for keeping track of user activities including user queries, click-through and their behavior. Research in web mining tries to address this problem by discovering knowledge from user logs. We propose an approach to discover patterns that can predict user's search, without aid of remote server. Our method analyses user's interactions by constructing FP-tree, which facilitates in producing templates from the user logs. <i>Consensus tree</i> growth is restricted and templates are obtained from leaves, which assist user's searching process with precision. We show the effectiveness of our method on realistic web logs and explore the tradeoff between prediction's accuracy and usefulness. Test results show the improved algorithm has lower complexity of time and space, and fit the capacity of memory.
Keywords:	frequent pattern mining, web log mining, CBFP mining, FP-tree, data mining.