利用隱含回饋提供搜尋引擎上的自動查詢修正

摘要

隨著全球資訊網蓬勃的發展，可以幫助使用者根據關鍵字搜尋相關資訊的搜尋引擎也已變成使用者不可或缺的工具之一。但對於搜尋引擎生手而言，往往不知道該如何地輸入適當的關鍵字，導致搜尋結果不如預期。如果搜尋引擎可以提供自動查詢修正(Automatic Query Refinement)的功能，將可以有效地幫助生手在網路上找尋到其想要的資訊。因此，如何得知使用者的資訊需求，如何自動化地達到查詢修正，則成為重要的課題之一。本研究利用使用者的隱含回饋(Implicit Feedback)來分析使用者的資訊需求，並探勘過去具有相同資訊需求的使用者經驗，以幫助搜尋引擎生手有效地搜尋網頁，以達到自動查詢修正的目的。

本研究中，在長期情境資訊方面，我們從查詢日誌中去辨別出以往使用者所查詢的關鍵字以及點選過的網頁，接著，在短期情境資訊的部份，我們也記錄下目前使用者的查詢關鍵字以及未點選之網頁。

最後，我們在長期情境中濾除掉搜尋引擎生手的查詢過程，同時探勘出與目前使用者有相似資訊需求的以往經驗使用者之查詢過程關鍵字集合，藉以推薦給目前使用者，完成自動查詢修正。

Automatic Query Refinement in Web Search Engines Using Implicit Feedback

Abstract

World Wide Web search engines can help users to search information by their queries, but novice search engines users usually don't know how to represent their information need. If search engines can offer query refinement automatically, it will help novice search engine users to satisfy their information need effectively. How to find users' information need, and how to perform query refinement automatically, have become important research issues. In this thesis, we develop the method to help novice search engine users for satisfying their information need effectively by implicit feedback. Implicit feedback in this research is referring to short-term context and long-term context.

In this research, first, long-term context is obtained by identifying each user's queries and extracting conceptual keywords of clickthrough data in each query session from query logs. Then, we also identify current user's queries and extract conceptual keywords of non-clickthrough data for short-term context identification.

Finally, we filter novice sessions from long-term context, and mine frequent itemsets of past experienced users' search behavior to suggest the most appropriate new query to current user according to their information need.