

結合詞彙關聯度及語句關聯度之潛在語意分析

丁泓丞, 周智勳

資訊工程學系

資訊學院

chc@chu.edu.tw

摘要

本論文針對文件自動摘要進行研究，文中結合潛在語意分析(Latent Semantic Analysis)與文章詞彙關聯度及文章句子關聯度的概念，建構出關聯性矩陣，用來加強與擷取文件中的概念結構，以得到語意層面的分析，進而篩選最適合之句子，作為文件

摘要的依據。實驗針對13 大類(生活、地方、社會、政治、科技、旅遊、財經、健康、國際、教育、運動、戲劇、藝文)共1300 篇網路新聞進行測試。在效果評估方面，提出兩個評估指標，以達到較客觀的摘要評估結果。實驗結果顯示，我們所提之方法能在較低相似度與重複性下，獲得較具代表性的摘要句。

關鍵字：中文文件摘要，潛在語意分析，奇異值分解，中文斷詞