

# 在XML資料倉儲中實體化資料方體選取之研究

李義偉、邱紹豐

E-mail: 9601127@mail.dyu.edu.tw

## 摘要

傳統資料庫存放的資料異動性高，並不利於資料分析。為了有效分析資料庫的資料，企業都會建立資料倉儲（Data Warehouse），以滿足資料分析的需求。藉由分析出來的資訊，協助公司經營的決策方向。查詢資料倉儲中的資料時，其匯總運算（Aggregation）耗費的資源很高，為了提升查詢能，以預先建置資料方體（Data Cube）的方法，降低匯總運算所耗費得資源。但在儲存空間有限之情況下，該如何挑選資料方體建置，以提升查詢效能，非常的重要。因此，本文利用Lattice架構挑選資料方體建置，儘量滿足每一種查詢方式為目標。此外，隨著網際網路的興起，許多的資料藉由網路傳遞，若公司間欲交換資料倉儲中的資料時，會有異質性資料庫間相容的問題。因此，本文以XCube架構為基礎，用XML（eXtensible Markup Language）來建立資料倉儲，以解決此問題。本文同時在XML資料倉儲中，加入Lattice架構，以方便使用者快速的查詢資料方體。

關鍵詞：資料方體

## 目錄

封面內頁 簽名頁 授權書.....	iii	中文摘要.....	iv	英文摘要.....	v
要.....	v	誌謝.....	vi	目錄.....	vii
圖目錄.....	ix	第一章 緒論.....	1	第一節 研究動機.....	2
第二節 研究目的.....	2	第三節 論文架構.....	3	第二章 相關研究.....	4
第一節 資料倉儲基本概念.....	4	第二節 XML相關技術.....	10	第三節 XML資料倉儲相關技術.....	13
第四節 Lattice相關技術.....	19	第三章 系統流程.....	23	第一節 資料方體大小之預估.....	24
第二節 挑選資料方體實體化之演算法.....	23	第三節 支援Lattice之XCube架構.....	28	第四節 搜尋資料方體之演算法.....	30
第四章 系統實作.....	33	第一節 實作挑選資料方體.....	34	第二節 實作資料方體查詢.....	37
第五章 效能評估.....	38	第六章 結論.....	43	參考文獻.....	45

## 參考文獻

- [1] W.H. Inmon, "Building the data warehouse", 1996.
- [2] Surajit Chaudhuri and Umeshwar Dayal, "An Overview of Data Warehousing and OLAP Technology", SIGMOD Record Volume 26., pp.65-74, 1997.
- [3] 曾守正和周韻寰, "資料庫系統進階實務", 華泰出版社, 2003.
- [4] XML, <http://www.w3.org/XML>.
- [5] Bergholz A., "Extending Your Markup: An XML Tutorial", IEEE Internet Computing, Vol 4, No.4, pp.74-79, July-August 2000.
- [6] XSLT, <http://www.w3.org/TR/xslt>.
- [7] XQuery, <http://www.w3.org/XML/Query>.
- [8] XML Schema, <http://www.w3.org/XML/Schema>.
- [9] DOM, <http://www.w3.org/DOM>.
- [10] Roy, J. and A. Rammanujan, "XML: Data's Universal Language", IEEE IT Professional, Vol.2 No.3, pp.32-36, May-June 2000.
- [11] Jason McHugh, Serge Abiteboul, Roy Goldman, Dallan Quass and Jennifer Widom, "Lore: A Database Management System for Semistructured Data", ACM SIGMOD Record Volume26, pp54-66, 1997.
- [12] Roy Goldman, Jason McHugh and Jennifer Widom, "From Semistructured Data to XML: Migrating the Lore Data Model and Query Language", WebDB99, pp25-30, 1999.
- [13] Roy Goldman, Sudarshan Chawathe, Arturo Crespo and Jason Mchugh, "A Standard Textual Interchange Format for the Object Exchange Model", Stanford university, 1996.
- [14] Andreas Bauer, Gunnar Harde and Wolfgang Hu"mmer, "XCube - XML For Data Warehouses", The 6th ACM international workshop

on Data warehousing and OLAP., pp.33-40, 2003.

[15] Laura Irina Rusu, Wenny Rahayu and David Taniar, " A Methodology for Building XML Data Warehouses " , International Journal of Data Warehousing & Mining., pp.23-48, 2005.

[16] Laura Irina Rusu, Wenny Rahayu and David Taniar, " On Building XML Data Warehouses " , IDEAL 2004., LNCS 3117, pp.3293-299, 2004.

[17] Andrew Nierman, Divesh Srivastava, H.V. Jagadish, Laks Lakshmanan, shurug Al-Khalifa, Stelios Paparizos and Yuqing Wu, " Grouping in XML " EDBT 2002 Workshops., LNCS 2409, pp.128- 147, 2002.

[18] Stelios Paparizos, Shurug Al-Khalifa, Adriane Chapman, H. V. Jagadish, Laks V. S. Lakshmanan, Andrew Nierman, Jignesh M. Patel, Divesh Srivastava, Nuwee Wiwatwattana, Yuging Wu and Cong Yu, " TIMBER: A native system for querying XML " , ACM SIGMOD., pp.672-672, 2003.

[19] Anand Rajaraman, Jeffrey D. Ullman and Venky Harinarayan, " Implementing Data Cubes Efficiently " , SIGMOD Record, 25:2, pp.205-227, 1996.

[20] 林志麟與邱承凡, " 資料倉儲實體化視域選取之研究 – 以資料方體之建置為例 " , 元智大學, 資訊研究所碩士論文, 2000.

[21] 陳耀輝與劉宇昌, " 在資料倉儲中針對查詢選擇實體化視域之研究 " , 屏東技術學院, 資訊管理技術研究所碩士論文, 1997。

[22] Amit Shukla, Jeffrey F. Naughton, Karthikeyan Ramasamy, Kristin Tufte, Prasad Deshpande and Yihong Zhao, " Cubing Algorithms, Storage Estimation, and Storage and Processing Alternatives for OLAP " , IEEE Data Eng.Bull.20(1). pp.3-11.,1997.