

記錄編號	3320
狀態	NC088FJU00392005
助教查核	
索書號	
學校名稱	輔仁大學
系所名稱	資訊工程學系
舊系所名稱	
學號	487516058
研究生(中)	何仁傑
研究生(英)	Ren-Jie Ho
論文名稱(中)	從大型資料庫中挖掘感興趣的型樣

論文名稱(英)	Mining Interested Information from Large Database
其他題名	
指導教授(中)	顏秀珍
指導教授(英)	Show-Jane Yen
校內全文開放日期	
校外全文開放日期	
全文不開放理由	

電子全文送交國圖.	
國圖全文開放日期.	
檔案說明	
電子全文	
學位類別	碩士
畢業學年度	88
出版年	
語文別	中文
關鍵字	資料探勘 序列型樣 相關規則 查詢

(中)	
關鍵字 (英)	data mining sequential patterns association rules query
摘要 (中)	<p>挖掘 (Mining) 相關規則 (Association Rules) 是找出大部分顧客，在每一筆交易中所採購之物品的相關性，挖掘序列型樣 (Sequential Patterns) 則是找出大部分顧客在其不同交易中採購物品的順序行為，兩者的目的都是從顧客的交易資料中，找出有用的資訊。由於顧客的交易不斷在進行，交易資料會不斷的增加，為了提供使用者最新的資訊，系統必須週期性地找出所有的相關規則與序列型樣，再者，使用者可能僅對某部分的資訊有興趣。然而，若每次都要重新找出所有這類的資訊，需花費相當多的時間。為了依照使用者的需求來找出最新的資訊。我們定義了一套資料探勘語言 (Data Mining Language)，使用者可以利用這套語言自行定義條件及需求，而系統可以針對這些條件及需求，快速的獲得使用者感興趣的資訊。</p>
摘要 (英)	<p>The destination of mining association rules is to discover the associative purchasing behaviors from each transaction of most customers. Mining sequential patterns is to discover the sequential purchasing behaviors from a large amount of the transactions of customers. Both of these are to discover the useful information. Because the transactions are increased and updated frequently, the system must discover all of the association rules and sequential patterns in a period of time. Besides, browsing through all of the information is not efficient if the users are only interested in part of the information. For these reasons, we design a data mining language for the users to define what they are interested. Then the system can look for the interested information rapidly according to the definition defined by the users.</p>
論文 目次	<p>摘要.....i 內容.....ii 圖表目錄.....iv</p> <p>1. 導論 1.1.1 問題說明 1.1.2 相關工作 4 2. 資料探勘語言 5 3. 挖掘感興趣之相關規則 8 3.1 資料庫型態的轉換 8 3.2 查詢相關規則：明確指定前項與後項 9 3.3 查詢相關規則：可能指定前項或後項包含的項目 11 4. 挖掘感興趣之序列型樣 20 4.1 挖掘 1-頻繁序列 20 4.2 查詢序列型樣：明確指定所有項目集之項目 24 4.3 查詢序列型樣：查詢序列中指定前 N 項或後 N 項項目集中的項目，或查詢序列中的項目集皆未指定(N³1) 26 4.4 查詢序列型樣：第三類查詢 34 5. 實驗結果 42 5.1 產生資料庫 42 5.2 實驗數據與討論 43 6. 結論與未來工作 50 參考資料 52 圖表目錄 表一 交易序列資料庫(TSD) 8 表二 交易位元資料庫(TBD) 9 表三.1-項目集資料庫 15 表四.2-項目集資料庫 16 表五. 2-項目集資料庫 18 表六. 3-項目集資料庫 19 表七 交易項目集其購買人數 21 表八 2-項目集資料庫 22 表九 2-候選項目集與其購買人數 22 表十 頻繁項目集資料庫 23 表十一 1-序列資料庫 24 表十二 1-項目集資料庫 30 表十三 1-序列資料庫 30 表十四 2-序列資料庫 31 表十五 1-序列資料庫 32 表十六 2-序列之資料庫 33 表十七 3-序列資料庫 33 表十八 1-項目集資料庫 39 表十九 1-序列資料庫 39 表二十 2-序列資料庫 40 表二十一 參數表 42 表二十二.資料庫</p>

之參數設定值 46 表二十三 C5-T10-I10-R10 所產生的候選序列個數ICKI與頻繁序列個數ILKI 47 表二十四 C10-T10-I10-R10 所產生的候選序列個數ICKI與頻繁序列個數ILKI 47 表二十五 C20-T10-I10-R10 所產生的候選序列個數ICKI與頻繁序列個數ILKI 47 表二十六 C10-T10-I10-R20 所產生的候選序列個數ICKI與頻繁序列個數ILKI 48 表二十七 C10-T10-I10-R40 所產生的候選序列個數ICKI與頻繁序列個數ILKI 48 圖一 挖掘所有頻繁項目集—Apriori 演算法與我們演算法之比值 44 圖二 挖掘所有頻繁序列—AprioriAll 演算法與我們演算法 48 圖三 挖掘所有頻繁序列—AprioriAll 演算法與我們演算法之比值 49

參
考
文
獻

[1]Ashok Savasere, Edward Omiecinski, Shamkant Navathe. An Efficient Algorithm for Mining Association Rules in Large Databases. In Proceedings of 21st VLDB Conference Zurich, Swizerland,1995 [2]H. Mannila and H. Toivonen. Discovering frequent episodes in sequences. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD' 96), pages 146-151, Portland, Orgon, Aug. 1996. AAAI Press. [3]H. Mannila, H. Toivonen, and A.I. Verkamo. Efficient Algorithm for Discovering Association Rules. In Proceedings of AAAI Workshop on Knowledge Discovery in Databases, pages 181-192, 1994. [4]J.S. Park, M.S. Chen, and P.S. Yu. An Effective Hash-Based Algorithm for Mining Association Rules. In Proceedings of ACM SIGMOD,24(2):175-186, 1995 [5]Ming-Syan Chen. Efficient Data Minging for Path Traversal Patterns. In IEEE Transactions on Knowledge and Data Engineering 1998, pages: 209-220. [6]Rakesh Agrawal, and et al. Mining Sequential Patterns. In Proceedings of International Conference on Data Engineering, pages 3-14,1995. [7]Rakesh Agrawal, and et al. Mining Sequential Patterns: Generalizations and Performance Improvements. In Proc. of the Fifth Int'l Conference on Extending Database Technology (EDBT), Avignon, France, March 1996. [8]R. Agrawal and et al. Mining Association Rules Between Sets of Items in Large Databases. In Proceedings of ACM SIGMOD, page 207-216,1993 [9]R. Agrawal and et al. Fast Algorithm for Mining Association Rules. In Proceedings of International Conference on Very Large Data Bases, pages 487-499, 1994. [10]R. Agrawal and et al. Database Mining: A Performance Perspective. In IEEE Transactions on Knowledge and Data Engineering, page 914-925,1993. [11]Ramakrishnan Srikant and Rakesh Agrawal. Mining Generalized Association Rules. In Proc. of the 21st Int'l Conference on Very Large Databases, Zurich, Switzerland, September 1995. [12]R. Agrawal, K. Lin, H. S. Sawhney, K. Shim. Fast Similarity Search in the Presence of Noise, Scaling, and Translation in Time-Series Databases. Proc. of the 21st Int'l Conference on Very Large Databases, Zurich, Switzerland, September 1995. [13]R. J. Bayardo Jr. and R. Agrawal. Minig the Most Interesting Rules. In Proc. of the 5th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining, August 1999. [14]R. Agrawal, R. J. Bayardo Jr. and R. Srikant. Athena: Mining-based Interactive Management of Text Databases. IBM Research Report RJ10153, July 1999. [15]Rakesh Agrawal, Johannes Gehrke: Dimitrios Gunopulos, Prabhakar Raghavan. Automatic Subspace Clustering of High Dimensional Data for Data Minig Applications. In Proc. of the ACM SIGMOD Int'l Conference on Management of Data, Seattle, Washington, June 1998. [16]R. Meo, G. Psaila, and S. Ceri. A New SQL-like Operator for

	<p>Mining Association Rules. In Proceedings of the International Conference on Very Large Data Bases, pages 122-133, 1996. [17]R. Agrawal and et al. An Interval Classifier for Database Mining Applications. In Proceedings of International Conference on Very Large Data Bases, pages 560-573, Vancouver, British Columbia,1992. [18]Show-Jane Yen and Arbee L.P. Chen. An Efficient Approach to Discovering Knowledge from Large Databases. In PDIS, page 8-18,1996 [19]Show-Jane Yen, Arbee L. P. Chen: An Efficient Data Mining Technique for Discovering Interesting Association Rules. DEXA Workshop 1997: 664-669 [20]Show-Jane Yen. Mining Frequent Traversal Patterns in a Web Environment. In Proceedings of International Sympostum on Intelligent Data Engineering and Learning. Pages: 219-224,1998.</p>
論文頁數	53
附註	
全文點閱次數	
資料建置時間	
轉檔日期	
全文檔存取記錄	
異動	M admin Y2008.M7.D3 23:17 61.59.161.35

記
録

--