

# A Study of Optimal Combination of Information Products by Applying Methodology of Data Mining

薛新光、杜瑞澤

E-mail: 9500948@mail.dyu.edu.tw

## ABSTRACT

The market of Personal Computer has divided into two group in Taiwan after popularity of the computer. One is brand, other is no brand (Do It Yourself :DIY). The computer has entered every home after the Internet as the member of the family. The only one group among users of computer will be insist on the way of assembling the components into the complete computers by themselves instead of products made in factory. They are the special symbol among so many brand computer users. The special group has owned the knowledge and skill, to construct and assemble a whole computer as their hobby and interest, especial for personalized representation different from others. They assemble the personal computer with add-on components between assembling computer with innovation and revolution toward the perfect computer. The more users in the group add more nonfunctional components like led display, control panel to be the totem of the personality from others. Our study hopes to explore the criteria of group and factors of purchase to understand the users of the e-stage about their choices of the products, hobby of usage, requirements and commercial behavior. The study hope to conquer the complexity and variety for assembling computer by the methodology of the Data-Mining to explore the all combinations of DIY computer and permutation of all components and find the optimal combination of information products and factors of design. The final result will derive some rules for designers to follow as their guiding for information products.

Keywords : Data-Mining、e-Stage、Factors of Design、Assembling Computer

## Table of Contents

中文摘要.....	iv	Abstract.....	v
謝.....	vi	目錄.....	vii
錄.....	xiv	表目錄.....	xxvi
錄.....	xxix	第一章 緒論.....	1
景.....	1	1.1 研究背景.....	1
的.....	6	1.2 研究動機.....	4
制.....	12	1.3 研究目的.....	6
程.....	16	1.4 研究重要性.....	8
勘.....	21	1.5 研究範圍與限制.....	12
法.....	30	1.6 名詞解釋.....	13
計.....	42	1.7 研究架構及流程.....	16
則.....	47	第二章 文獻探討.....	20
法.....	50	2.1 資料探勘.....	21
析.....	54	2.2 關聯法則.....	28
析.....	55	2.3 Apriori 演算法.....	30
析.....	56	2.4 決策樹.....	34
析.....	56	2.5 模組化設計.....	42
談.....	57	2.6 綠色設計.....	46
析.....	58	2.7 模組化產品設計準則.....	47
析.....	59	2.8 小結.....	48
析.....	59	第三章 研究方法.....	50
勘.....	60	3.1 研究程序.....	50
勘.....	61	3.2 需求分析.....	54
統設計.....	62	3.2.1 市場分析.....	55
面.....	62	3.2.2 商業資料分析.....	55
	62	3.2.3 廠商網站資訊分析.....	56
	62	3.2.4 實地賣場觀察分析.....	56
	61	3.3 產品特性分析.....	56
	61	3.3.1 廠商產品線分析.....	56
	61	3.3.2 使用者組裝分析.....	57
	61	3.3.3 專家訪談.....	57
	62	3.4 產品特性問卷要素分析.....	57
	62	3.4.1 組裝功能性分析.....	58
	62	3.4.2 組裝特性的分析.....	58
	62	3.4.3 問卷小樣本與專家分析.....	59
	60	3.5 資料探勘資訊系統設計.....	59
	60	3.5.1 Apriori 資料探勘演算法分析.....	59
	60	3.5.2 Apriori 資料探勘系統分析與設計.....	60
	60	3.5.3 決策樹演算法分析.....	60
	61	3.5.4 決策樹系統分析與設計.....	60
	61	3.6 資訊系統資料探勘.....	61
	61	3.6.1 資訊淨化.....	61
	61	3.6.2 資料探勘.....	61
	61	3.6.3 較佳化組合的結果驗證與調整.....	61
	62	3.7 電腦組裝資料探勘系統設計.....	62
	62	3.7.1 廠商資料輸入畫面.....	62
	62	3.7.2 使用者資料問卷輸入畫面.....	62
	63	3.7.3 廠商資料分析.....	63
	63	3.7.4 廠商使用零件分	

析.....	65	3.7.5 廠商資料探勘分析.....	69	3.7.6 使用者資料分
析.....	69	3.7.7 使用者組裝電腦電腦組件重視的程度.....	74	3.7.8 使用者使用零件分
析.....	77	3.7.9 使用者資料探勘分析.....	80	第四章 資料分析與結
果.....	81	4.1 廠商基本資料分析.....	81	4.1.1 銷售廠家型態家數分
析.....	81	4.1.2 銷售廠家型態家數分析.....	82	4.1.3 銷售主要客戶群分
析.....	84	4.1.4 廠商區域家數一覽表.....	85	4.1.5 銷售廠商訪談職稱分類
表.....	86	4.2 廠商銷售零組件資料分析.....	87	4.2.1 CPU 銷售量比
率.....	87	4.2.2 主機板銷售比率.....	89	4.2.3 記憶體銷售比
率.....	92	4.2.4 硬碟銷售比率.....	95	4.2.5 光碟機銷售比
率.....	97	4.2.6 顯示卡銷售比率.....	100	4.2.7 電腦機殼銷售比
率.....	103	4.2.8 監控面板銷售比率.....	106	4.2.9 冷卻系統銷售比
率.....	107	4.3 使用者基本資料分析.....	110	4.3.1 使用者性別分
析.....	110	4.3.2 使用者年齡分析.....	110	4.3.3 使用者學歷分
析.....	111	4.3.4 使用者居住地分析.....	112	4.3.5 使用者職業分
析.....	114	4.3.6 使用者工作性質與電腦相關性分析.....	115	4.3.7 使用者學生身份分
析.....	115	4.3.8 使用者科系類別分析.....	116	4.3.9 使用者組裝臺數分
析.....	117	4.3.10 使用者購買電腦機殼的材質分析.....	118	4.3.11 使用者購買電腦零組
件分析.....	119	4.4 使用者對電腦組件重視的程度優先權重分析.....	120	4.4.1 使用者選購組
件主要的考慮的重點分析.....	121	4.4.2 選購組件效能表現重點分析表.....	122	4.4.3 選購組件
價格差異重點分析表.....	124	4.4.4 選購組件穩定度差異重點分析表.....	126	4.4.5 選購組
件外觀造型差異重點分析表.....	127	4.4.6 選購組件功能差異重點分析表.....	129	4.4.7 組裝
電腦的原因.....	131	4.4.8 電腦零組件重視程度.....	132	4.5 使用者心中最想要的
的電腦零組件的選擇.....	134	4.5.1 CPU(處理器)使用分析.....	134	4.5.2 主機板使用分
析.....	137	4.5.3 記憶體使用分析.....	141	4.5.4 硬碟機使用分
析.....	145	4.5.5 光碟機使用分析.....	149	4.5.6 顯示卡使用分
析.....	152	4.5.7 電腦機殼使用分析.....	156	4.5.8 監控面板使用分
析.....	159	4.5.9 冷卻系統使用分析.....	161	4.6 專家意見分
析.....	164	4.6.1 專家資格分析.....	165	4.6.2 專家訪談意見分
析.....	168	4.7 設計要素分析.....	170	4.8 資料探勘結果分
析.....	176	4.8.1 使用者關聯法則分析圖.....	176	4.8.2 廠商關聯法則分析
圖.....	179	4.9 小結.....	182	第五章 結論與建
議.....	184	5.1 結論.....	184	5.2 建
議.....	189	參考文獻.....	190	中文部
份.....	190	外文部份.....	191	附錄
一.....	195	使用者問卷.....	195	附錄
二.....	199	廠商問卷.....	199	附錄
三.....	203	專家訪談.....	203	附錄
四.....	205	組裝使用者探勘結果.....	205	附錄
五.....	209	銷售廠商探勘結果.....	209	

## REFERENCES

- 中文部份 1. <http://cindy.cis.nctu.edu.tw/> 交大資訊科學研究 learning lab website 2. 王俊程, 2000, 運用資料挖掘技術協助健保財務管理: 以承保中斷及繳費不正常資料挖掘為例。行政院衛生署委託研究成果報告。 3. 史濟元、戴華山, (2000), 產品綠色設計及應用實例, 工業防治污染第74期, P177-P187 4. 周欽凱, 2004, 利用『資料探勘技術』探討急診高資源耗用者之特性, 國立台灣大學醫療機構管理研究所碩士論文。 5. 官政能, 陳源德, 1993, 應用綠色設計觀念於產品開發之策略研究, 工業設計技術及學術研討會 論文集, p11-p19。 6. 施進發, 梁錫昌, “機械模塊學理論,” 中國機械工程, 第 8 卷, 第 6 期, pp.53-55, 1997。 7. 張琬婷, 吳孟潔, 許容溶, 李昱慧, 陳羽瑩, 李芝儀, 謝邦昌, 2003, 資料探礦在行銷策略上之 應用---以運動鞋市場為例。中華資料探礦協會, CDMS-Newsletter (7) , p1-p32。 8. 張衡閣, 2002, 一個資料庫多維度序列法則探勘方法, 朝陽科技大學資訊管理系, 碩士論文。 9. 黃仁貴, 2002, 以叢集分析技術探討病患就診屬性與看診時間之關係。台北醫學大學醫學資訊研究所碩士論文。 10. 黃勝崇, 2001, 資料探勘應用於醫療院所輔助病患看診指引之研究。南華大學 11. 黃裕哲, 2001, 綠色模組化設計之評估架構研究.以電動自行車組裝與拆解為例, p26-p30, 大葉大學設計研究所碩士論文。 12. 劉仁傑, 陳國民, 模組化趨勢下的產品創新策略新動向, 機械工業雜誌154 期。 pp.181。 13. 劉家煒, 2001, 應用人工智慧方法於產品設計之研究 國立成功大學機械工程學系碩士論文。 外文部份 14. A Distance-Based Attribute Selection

Measure for Decision Tree Induction. Machine Learning, 6, 81-92. Paul E. Utgoff (1989). 15. Agrawal, R. and Srikant, R., "Fast Algorithm for Mining Association Rules," In Proceedings of the 20th International Conference on Very Large Databases, pp. 487-499, 1994. 16. Agrawal, R., Imilienski, T. and Swami, A., "Mining Association Rules between Sets of Items in Large Databases," In Proceedings of ACM SIGMOD International Conference on Management of Data, pp. 207-216, 1993. 17. Brin, S., Motwani R. and Silverstein, C., "Beyond market baskets: Generalizing association rules to correlations," In Proceedings of ACM SIGMOD Conference on Management of Data, pp. 265-276, 1997. 18. Ericsson A. and Erixon G., Controlling Design Variants: Modular Product Platforms, ASME Press, U.S.A., 1999. 19. Erixon G. and Ostgren B., "Synthesis and Evaluation Tool for Modular Designs," International Conference on Engineering Design, Hague, pp.898-905, August 17-19, 1993. 20. Erixon G., von Yxkull A. and Arnstrom A., "Modularity-the Basis for Product and Factory Reengineering," Annals of CIRP, Vol.45, pp.1-6,1996. 21. F Han, J. and Kamber, M., "Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers," San Francisco, 2000. 22. Fayyad, U.M., "Data Mining and knowledge Discovery: MakingSense Out of data," IEEE Expert, Volume 11, Issue 5, pp. 20-25, 1996. 23. Schlesinger, G., "Inspection tests on machine tools", Machinery Publishing Co. Ltd,1932. 24. Incremental Induction of Decision Trees. Machine Learning, 4, 161-186. 25. Jiawei Han , Micheline Kamber , Data Mining – Concepts and Techniques , Morgan Kaufmann Publish , 2001, 1rd 26. Kim, Sung-Min, Jong-Dal Kim, Jeong-Hee Hong, Do-Won Nam, Dong-Ha Lee, Jeon-Young Lee , "A System for Association Rule Finding from an Internet Portal Site , " 2000. 27. Kleissner, C., "Data mining for the enterprise," In Proceedings of the Thirty-First Hawaii International Conference on, Volume 7, pp. 295-304, 1998. 28. Michael, J.A. and Linoff, G., "Data Mining Technique: for Marketing, Sales and Customer Support," Wiley Computer Publishing, New York, 1997. 29. Quinlan, J.R., 1999, "Simplifying Decsion Trees" , International Journal of Human Computer Studies 51,p497-510. 30. R. Agrawal and R. Srikant, "Mining Sequential Patterns," Proc. of the Int'l Conference on Data Engineering, Mar. 1995. 31. R. Agrawal and R.Srikant, "Fast Algorithm for Mining Association Rules in Large Databases," Proceedings of The 20th International Conference on 64 Very Large DataBases, Santiago, Chile, pp. 487-490 (1994). 32. Ragel, B. Cremilleux, "MVC—a preprocessing method to deal with missing values," Data & Knowledge Engineering Volume: 18, Issue: 3, pp. 189-223 (1996). 33. Simoudis, E., "Reality check for data mining," IEEE Expert, Volume 11, Issue 5, pp. 26-33, 1996. 34. Stuart Russell, Peter Norvig (1995). Section 18.3 - Learning Decision tree. Artificial Intelligence - A Modern Approach. R.Lopez De Mantaras (1991).