

PRODUCT RECOMMENDATION FOR E- GROCERIES USING APRIORI ALGORITHM

A Project submitted to the Graduate School in partial fulfillment
of the requirements for the degree
Master of Science Intelligent System (IS)
Universiti Utara Malaysia

by

Muziah Muhammad

©Muziah Muhammad, 2004. All rights reserved



JABATAN HAL EHWAL AKADEMIK
(Department of Academic Affairs)
Universiti Utara Malaysia

PERAKUAN KERJA KERTAS PROJEK
(Certificate of Project Paper)

Saya, yang bertandatangan, memperakukan bahawa
(I, the undersigned, certify that)

MUZIAH MUHAMMAD

calon untuk Ijazah
(candidate for the degree of)

MSc. (IS)

telah mengemukakan kertas projek yang bertajuk
(has presented his/her project paper of the following title)

PRODUCT RECOMMENDATION
FOR E-GROCERIES USING APRIORI ALGORITHM

seperti yang tercatat di muka surat tajuk dan kulit kertas projek
(as it appears on the title page and front cover of project paper)


bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan
dan meliputi bidang ilmu dengan memuaskan.
*(that the project paper acceptable in form and content, and that a satisfactory
knowledge of the filed is covered by the project paper).*

Nama Penyelia Utama

(Name of Main Supervisor) : **ENCIK MOHD. SHAMRIE SAININ**

Tandatangan

(Signature)


:  Tarikh (Date): 28/03/04

Nama Penyelia Kedua

(Name of 2nd Supervisor) : **ENCIK AZIZI AB. AZIZ**

Tandatangan

(Signature)

:  Tarikh (Date): 28/03/04

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirement for a postgraduate degree from Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by my supervisor or, in their absence, by the dean of the Graduate School. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole part, should be addressed to:

**Dean of Graduate School
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman**

ABSTRACT

(BAHASA MELAYU)

Kemunculan teknologi dunia tanpa sempadan iaitu Internet, telah membenarkan pelbagai kemungkinan berlaku. Pemasaran barangan secara tidak langsung merupakan satu aspek utama dalam menyumbangkan keuntungan perniagaan. Kebiasaannya, peruncit konvensional terdahulu akan menggunakan susun atur secara fizikal dan kemungkinan pembelian yang terancang melalui paparan barangan. Dalam e-runcit, idea tersebut telah dibangunkan berasaskan cadangan barangan. Kertas kajian ini akan membincangkan tentang rekabentuk dan pembangunan E-Groceries berdasarkan tujuan berkenaan.

ABSTRACT

(ENGLISH)

The emergence of Internet allows many possibilities; one of them is electronic groceries. Indirect marketing is one aspect that contributes to the business profit. Basically, conventional grocer will utilize physical layout and possible buying sequence using product display. In e-groceries, such idea can be implemented using product recommendation basis. This paper will discuss briefly the design and the development of E-Groceries for such purpose.

ACKNOWLEDGEMENTS

I would like to express my thanks and gratitude to Allah, the Most Beneficent, the Most Merciful whom granted me the ability and willing to start and complete this project.

Special thanks to my supervisors, En. Mohd. Shamrie Sainin and En. Azizi Ab Aziz that guides, support, confidence and provides fruitful suggestions during the project development and the writing of this final report.

Especially to my beloved mum, Puan Azizah Jantan and my family whose love and patient make me feels strong in developing this project. I wish to thank for all their help, support, interest and valuable hints to my beloved best friend, Siti Mazura Che Doi.

My former IKBS classmates' batch Nov 2002 and IKBS lecturers who have supported me towards the completion of this project as well as my friends Adib, liza, sue, IKBS and AI juniors.

Lastly but not the least, to my beloved fiancée Ahmad Adnan Pūjiman, thanks for your encouragement, advices and love. May God Bless You All.

TABLE OF CONTENTS

	Page
PERMISSION TO USE	i
ABSTRACT (BAHASA MELAYU)	ii
ABSTRACT (ENGLISH)	iii
ACKNOWLEDGMENTS	iv
LIST OF TABLES	v
LIST OF FIGURES	vi
CHAPTER 1: INTRODUCTION	
1.1 Problem Statement	3
1.2 Project Objective	3
1.3 Scope of the Study	3
1.4 Significance of the Study	4
1.5 Thesis Structure	4
CHAPTER 2: LITERATURE REVIEW	
2.1 E-Commerce and E- Groceries	5
2.2 Product Recommendation System	5
2.3 Apriori Algorithm	7
2.4 Summary	9
	10
CHAPTER 3: METHODOLOGY	

3.1	Problem Identification	12
3.2	Theory Building	12
3.2.1	System Design	12
3.2.2	Apriori Algorithm	15
3.3	System Development	16
3.3.1	Database Module	17
3.3.2	Apriori Algorithm Module	18
3.3.3	Generated Rule Module	18
3.3.4	E-Groceries Page Module	18
3.4	Experimental Result	18
3.5	Frequent Similarity Between Two Products	22
3.6	Summary	23
CHAPTER 4: FINDING		
4.1	The Example of the Results	24
4.2	Summary	33
CHAPTER 5: CONCLUSION AND RECOMMENDATIONS		
5.1	Conclusion	34
5.2	Recommendations	35
	REFERENCES	36
	APPENDIXES	
	Appendix A	
	Appendix B	
	Appendix C	
	Appendix D	

LIST OF TABLES

Table No.	Name of Table	Page No.
3.1	Experimental Result	21
3.2	Example of Transaction	22
3.3	Result of Support	22
4.1	The Recommended Product	24

LIST OF FIGURE

Figure No.	Name of Figure	Page No.
2.1	Example Of Product Recommendation System (www.amazon.Com)	7
2.2	Example Of Product Recommendation System (www.cdnw.Com)	8
3.1	Research Methodology (Nunameker <i>Et Al.</i> , 1991)	11
3.2	UML Use Case View	12
3.3	Sequence Diagram For Product Recommendation	13
3.4	Sequence Diagram For Generate Rule	14
3.5	Sequence Diagram For Product Display	15
3.6	General Flow Of Apriori Algorithm	16
3.7	System Architecture	17
3.8	Process For Product Recommendation	19
3.9	Example Of Product Recommendation	20
3.10	Calculation Of Confidence Of $A \rightarrow C$	22
3.11	Calculation Of Confidence Of $C \rightarrow A$	23
4.1	Case One	26
4.2	Recommendation Page For Case One	26
4.3	Case Two	27
4.4	Recommendation Page For Case Two	28

4.5	Case Three	29
4.6	Recommendation Page For Case Three	29
4.7	Case Four	30
4.8	Recommendation Page For Case Four	31
4.9	Case Five	31
4.10	Recommendation Page For Case Five	32

CHAPTER 1

INTRODUCTION

The continuous growth in the size and use of the World Wide Web imposes new methods of design and development of online information services. Most Web structures are large and complicated and users often miss the goal of their inquiry, or receive ambiguous result when they try to navigate through them. On the other hand, the e-business sector is rapidly evolving and the need for web marketplaces that anticipate the needs of the customers is more evident then ever.

There is no exact definition of e-commerce. E-commerce is known as a method of buying and selling products, services and information electronically. The main methods of e-commerce are the Internet and WWW although e-mail, fax and telephone orders are also prevalent. According to Timmers (1999) e-commerce including electronic trading of physical goods and intangibles such as information. This encompasses all the trading steps such as online marketing, ordering, payment and support for delivery. E-commerce includes the electronic provision services, such as after-sales support or online legal advice. It also include electronic support for collaboration between

The contents of
the thesis is for
internal user
only

REFERENCES

- Aldridge, A., Forcht, K., and Pierson, J. (1997). *Get linked or get lost: marketing strategy for the Internet*. Internet Research: Electronic Networking Applications and Policy, Vol. 7 No. 3, pp.161-169.
- Agrawal, R., Imielinski, T and Swami, A. (1993). *Mining Association Rules Between Two Sets of Items in Large Databases*, ACM SIGMOD Conference on the Management of Data, pp. 207-216.
- Borgelt, C. and Kruse, R., (2002). Induction of Association Rules: Apriori Implementation. URL: <http://www.citeseer.nj.nec.com> Date Accessed: 19th Oct 2003
- Bloch, M., Pigneur, Y., and Segev, A. (1996). *On the Road of Electronic Commerce: A Business Value Framework, Gaining Competitive Advantage and some Research Issues*. http://inforge.unil.ch/yp/pub/ROAD_EC/EC.HTM. Accessed date: 18 Jan, 2003.
- Brin, S., Motwani, R., & Silverstein, C.,(1997).Beyond Market Basket: Generalizing Association Rules To Correlations.ACM SIGMOID-'97.
- Cunningham SJ. and Frank E, (2000), Market Basket Analysis of Library Circulation Data, in Proc. Of the Sixth International Conference on Neural Information Processing, Vol. 2, Perth, Australia, IEEE Service Canter, pp. 825-830.

- Connolly, T., and Begg, C Eds. 2002. *Database Systems: A Practical Approach to Design, Implementation and Management*. Edinburg Gate, Hawlow: Addison-Wesley.
- Chen, S. (2001). *Strategic Management of e-Business*. Sussex: John Wiley & Sons, Ltd..
- Doddi, s.,Marathe, A., Ravi,S.S. and Toney,D.C.,(2002). Discovery of Association Rules in Medical Data. Online: <http://www.c3.lanl.gov>
- Damanpour, F. (2001). *E-business E-commerce evolution: Perspective and Strategy*. *Managerial Finance*, Vol. 27 No.7, pp.16-31.
- Dyche, J. (2000) *e-Data: Turning Data Into Information with Data Warehouse*, Addison Wesley Longman, Second Edition, Canada.
- Eid, R., Trueman, M., and Ahmed, A.M. (2002). *A cross-industry review of B2B critical success factors*. *Internet Research: Electronic Networking Applications and Policy*, Vol. 12 No. 2, pp.110-123.
- Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P. and Uthurusamy, R. (1996) "Advance in Knowledge Discovery and Data Mining".AAAI press/ MIT press.
- Hamuro, Y., Katoh, N., Matsuda, Y. and Yada, K.,"Mining Pharmacy Data Helps To Make Profits", *Data mining and Knowledge Discovery*, Vol. 2, pp.391-398.
- Han, J.& Kamber, M.,(2001). *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers.
- Jennings, S and Higuchi, E. (1993) E-commerce Recommendation Application. *Journal of Data Mining and Knowledge Discovery*. Vol. 5. pp. 115-152.

- Lee, K. B. and Suh, S.C., (1998). The Efficient Algorithm in the Volume of Market Basket Data for Association Rules, *Expersys-98, The 10th International Conference*.
- Mannila, H., Toivonen, H. and Verkamo, A. (1998). *Efficient Algorithms for Discovering Association Rules*, *IEEE Transactions on Knowledge and Data Engineering* 8(6), pp. 962-969.
- Ming, M. V., Karypis, G. & Kumar, V., (1997). Parallel Algorithms for Mining Sequential Associations: Issues and Challenges. <http://www.citeseer.nj.nec.com/>
Date Accessed: 12 October 2003.
- Nor Fadilah Tahar @ Yusoff (2003), Association Rule Mining Using Market Basket Analysis, Master Project MSc. Intelligence Knowledge Based System, Universiti Utara Malaysia.
- Nunamaker, J. F., Chen, M. and Purdin, T. D. M. (1991) System Development in Information Systems Research, *Journal of Management Information Systems*, Vol. 7, No. 3, pp. 89-106.
- Schafer, J.B., Konstan, J.A. and Reidl, J. (2001). "E-commerce recommendation Application", *Data Mining and Knowledge Discovery*, Kluwer Academic, pp. 115-153.
- Sarwar B.M., Karypis, G., Konstan, J. and Riedl, J. (2000). "Analysis of Recommendation Algorithms for E-Commerce," *ACM Conf. Electronic Commerce*, ACM Press, pp.158-167.
- Sarwar B.M., Karypis, G., Konstan, J. and Riedl, J. (2000). Application of dimensionality reduction in recommender systems-a case study. *ACM WebKDD 2000 Web Mining for E-Commerce Workshop*.

- Timmers, P. (1999). *Electronic Commerce: Strategies and Models for Business to Business Trading*. Chichester: John Wiley & Sons Ltd.
- Turban, E., King, D., Lee, J., Warkentin, M., and Chung, H.M. (2002). *Electronic Commerce: A Managerial Perspective*. 2nd Edition, New Jersey: Pearson Education
- United Nation Conference on Trade and Development (UNTAD) (2002). *E-commerce and Development Report 2002*. <http://r0.untad.org/>. Accessed date: 11 July, 2003.
- Webb, G. I.,(2000). Efficient Search For Association Rules. To Appear In Proc. KDD-2000, Boston.
- Wen, H. J., Chen, H. G., and Hwang, H. G. (2001). *E-commerce Web site design: strategies and models*. Information Management & Computer Security, Vol 9 No.1, pp.5-12.