
Dissertation

Present to

the Computing Department of

the Hong Kong Polytechnic University

in partial fulfilment

of the requirement for the Degree of

Master of Science in Information Technology

by Andy C.F. Lai

January 1998

Dissertation

Present to

the Computing Department of

the Hong Kong Polytechnic University

in partial fulfilment

of the requirement for the Degree of

Master of Science in Information Technology

by Andy C.F. Lai

January 1998

Abstract of dissertation entitled:

A Personalized WWW Agent

submitted by Andy Chun Fai Lai

for MSc in Information Technology

at The Hong Kong Polytechnic University in Jan 1998.

Information Retrieval has long been an interesting subject that leads to many studies. In general, its objective is to search for better and faster algorithms to assist information searching. It is commonly agreed that, the efficiency of information gathering will result with a better decision making and quality of life.

Day after day, information are being posted in the World Wide Web. In such a huge amount of un-categorised raw information, tools to help gathering relevant information like search engines and Web robots are being developed. However, a single method may not be enough to serve all kinds of people. Here, it has left some room for us to enhance the technologies and make the Web more useful.

In this paper, we present our design of a PWA (Personalized WWW Agent) to provide information searching for Internet users via the WWW. The system can be personalized by letting the users to choose their own methods to analyse Internet documents found. A neural network is used to help in predicting documents most likely will meet the users' needs.

ACKNOWLEDGEMENTS

It is my pleasure to express my appreciation to Dr. Vincent Ng, in the Department of Computing, the Hong Kong Polytechnic University for his precious advice and guidance on this dissertation. Also, I would like to thank my parents that gave me spiritual support and encouragement for my completion of this paper.

Table of Contents

Abstract	p.1
Acknowledgements	p.2
1. Introduction	p.5
1.1 Problem Domain	p.5
1.2 Contribution of The Study	p.8
2. Review of Related Works	p.9
2.1 Information Retrieval	p.9
2.2 Searching and Ranking Algorithm	p.11
2.3 Features comparison between Internet Indexing search engines	p.17
2.4 Case studies of similar applications	p.19
3. Use of Artificial Intelligence	p.22
3.1 An overview of Artificial Intelligence	p.22
3.2 Backpropagation Network	p.26
4. Design of the PWA	p.30
4.1 Tools used	p.30
4.2 The choice and settings of the Neural Network	p.32
4.2.1 Settings of the Neural Network	p.32
4.2.2 The choice of the Neural Network (Neural Network Topology)	p.37
4.3 Network Architecture	p.38
4.4 Client and Server architecture	p.39
4.5 Functionality and GUI designs	p.41
4.6 Data Flow Diagram of the system	p.53
4.7 Classes design of the system	p.59

5. Testing and Evaluation	p.60
5.1 How we conduct the testing	p.60
5.2 Results	p.63
5.3 Discussion	p.68
6. Enhancement	p.69
7. Conclusion	p.71
8. References and bibliography	p.72
9. Appendix	p.74