Towards Eradication of SPAM: A Study on Intelligent Adaptive SPAM Filters

Tarek Hassan
(B.Sc. Computer Science, Egypt)

This thesis is presented for the degree of Master of Computer Science

Murdoch University, Western Australia

2006

I declare that this thesis is my own account of my research and contains as its main content work, which has not previously been submitted for a degree at any tertiary educational institution.

_____

Tarek Hassan

ABSTRACT

As the massive increase of electronic mail (email) usage continues, SPAM (unsolicited bulk email), has continued to grow because it is a very inexpensive method of advertising. These unwanted emails can cause a serious problem by filling up the email inbox and thereby leaving no space for legitimate emails to pass through. Currently the only defense against SPAM is the use of SPAM filters. A novel SPAM filter GetEmail5 along with the design rationale, is described in this thesis. To test the efficacy of GetEmail5 SPAM filter, an experimental setup was created and a commercial bulk email program was used to send SPAM and non-SPAM emails to test the new SPAM filter.

GetEmail5's efficiency and ability to detect SPAM was compared against two highly ranked commercial SPAM filters on different sets of emails, these included all SPAM, non-SPAM, and mixed emails, also text and HTML emails.

The results showed the superiority of GetEmail5 compared to the two commercial SPAM filters in detecting SPAM emails and reducing the user's involvement in categorizing the incoming emails.

This thesis demonstrates the design rationale for GetEmail5 and also its greater effectiveness in comparison with the commercial SPAM filters tested.

TABLE OF CONTENTS

Page No.

ACKNOWLEDGEMENTS

The work presented in this thesis was carried out under the supervision of Mr. Peter Cole and Associate Professor Lance C.C. Fung, to whom I extend my gratitude for their valuable and friendly guidance, encouragement and helpful suggestions throughout the progress of the research and in the preparation of this thesis.

I would like to extend my special thanks and indebtedness to Kashif Saleem for his help and support.

My gratitude and heartfelt thanks to Dr. Karin Strehlow and Anthony Horton, for their support, friendship, help and encouragement, God Bless.

Finally, my thanks and gratitude go to my wife, Eman and daughter, Amira who have been understanding, patient, supportive and making *duaa(supplication)* for the success of my study. Thank you also to my parents, sisters and brother for their encouragement and support during my stay in Australia.

PUBLICATIONS

Hassan, T., Cole, P. and Fung C.C. (2006), "An Intelligent SPAM filter – GetEmail5", Proceedings of the 2nd IEEE International Conference on Cybernetics and Intelligent Systems (CIS 2006), Bangkok Thailand, 7-9 June, 2006, (ISBN 1-4244-0023-6) pp. 86-90.

Hassan T., Cole P. and Fung C.C. (2005), "Development and evaluation of an Intelligent SPAM filter", Proceedings of the 6th Postgraduate Electrical and Computing Symposium (PEECS 2005), (ISBN 0-7298-06090-X), Edith Cowan University, Perth, WA, pp. 142-145

Hassan T., Cole P., and Fung C.C., (2004) "Towards Eradication of Spam: A Study on Intelligent Adaptive Spam Filter", Proceedings of the Fifth Postgraduate Electrical Engineering and Computing Symposium (PEECS 2004), 28th September 2004, Perth, Western Australia, pp 203-206.