

Curtin Business School

A Study of Cost, Awareness, Knowledge and Perception of Spam 2.0

Farida Hazwani Mohd Ridzuan

**This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University**

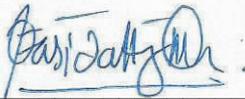
December 2013

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signed: _____

A handwritten signature in blue ink, appearing to read "Fasi Jattar", is written on a light-colored rectangular background. The signature is cursive and includes a horizontal line underneath the main text.

Date: _____

Abstract

Spam 2.0 is defined as “*propagation of unsolicited, anonymous, mass content to infiltrate legitimate Web 2.0 applications*”. The existing research on Spam 2.0 is focused on the detection and prevention of Spam 2.0. Although it has been claimed that Spam 2.0 has its own cost, there has been no research to quantify this cost. At the same time, it has been found out that the extent of proliferation of Spam 2.0 depends on the users’ awareness, knowledge and perception of the issue but the current literature falls short of an adequate exploration of these aspects. Therefore, the objectives of this research are twofold:

- 1) To estimate the cost of Spam 2.0; and
- 2) To assess the public awareness, knowledge and perception of Spam 2.0.

For the purpose of a cost study of Spam 2.0, an extensive review of the current literature on email-spam cost-models and other related models has been carried out, the problems with the existing models highlighted, and the related cost categories and parameters identified to develop a Spam 2.0 cost model. The proposed model specifies the related cost categories and parameters; however, detailed investigations have been carried out for two cost categories only, i.e., the storage cost and the loss of productivity cost. Data collection for storage cost is based on the HoneySpam 2.0 dataset, while data for the Timing dataset have been obtained from a survey carried out to estimate the loss of productivity cost. The proposed Spam 2.0 cost model:

- 1) covers 5 cost categories, i.e., storage cost, loss of productivity cost, labour cost, connectivity cost and software cost;
- 2) proves that Spam 2.0 does have its costs; and
- 3) quantifies the cost of the storage and loss of productivity.

To fulfill the second objective, a comprehensive review of current awareness, knowledge and perception, covering not only computer security issues but also issues from other fields, such as public health, was carried out to develop a 29-item questionnaire for a web-based survey. A set of heuristics to determine the extent of Spam 2.0 was developed and simultaneously gives an in depth understanding of users' perception towards Spam 2.0. The results of the survey give an in-depth understanding of the:

- 1) level of public awareness of Spam 2.0;
- 2) level of public knowledge of Spam 2.0; and
- 3) public perception of Spam 2.0.

Overall, the research carries forward the work on Spam 2.0 and explores some of its unique aspects.

Acknowledgements

In the name of Allah, the Most Gracious, the Ever Merciful.

I am obliged to many people in the voyage of my doctoral thesis. First and foremost, I wish to take this opportunity to express my sincere gratitude for my supervisor, Dr Vidyasagar Potdar. No words would be enough to thank him for his continuous guidance, patience, discernment, and kind support. I am greatly indebted for his time and inputs during the whole period of research. The publication and presentation of my research with him has been a rewarding learning experience for me. I would also like to express my gratitude to Prof. Elizabeth Chang, Prof. Tharam Dillon and Dr Peter Dell.

My thanks also go to my co-supervisor, Wendy Hui. She specifically guided me on quantitative research which boosted my confidence with the work on the survey. I also have to thank Dr Jaipal Singh for his continued kindness and constantly keeping track of my progress. His knowledge and advice have not only guided me but also provided me with encouragement.

My journey to Curtin University would not have been possible without the scholarship offered to me. So, I acknowledge the financial support received from USIM and the Government of Malaysia. I would like to extend my sincere thanks to them for providing me with this opportunity. I am very thankful to my ex-dean who has always been my guiding pillar of strength, Prof Em. Jailani Sukaimi. My special thanks are due to the current dean and colleagues who have supported the extension of my study leave.

The research took longer than I had expected. I am eternally grateful to my parents and siblings for their patience and trust during the ups and downs all these years. Especially, Abah, Mama, Syimah and Miyah have been very supportive, emotionally and financially. They provided me with their much needed and undying support, prayers and unconditional love at a key stage, when it seemed impossible to carry on with the research. I thank them all from the bottom of my heart. I must admit that it is the thought of fulfilling my parents' most cherished dream for their daughter, i.e., to finish this Doctorate, that alone has given me strength all the way through. "*Yeoreobun saranghaeyo!*"

My heartfelt appreciation goes to the close friends I made during these past few years and also those who stood by me during my hardest time. Novita Ikasari, thank you for all the academic advice and personal help. I will always be grateful to you for your wise thoughts and sane advice. K Shima, thank you for always making me laugh and accompanying me to the stress release session in Zumba classes. K Raihana, K Noesma and K Saidatul, thanks for the late night long hours, chat, free food entertainment, care and beautiful friendship.

My warmest thanks go to all for the support I had all these years. Thank you for always comforting me with your constant kind and encouraging words: ex housemate Nani, Fifi, Ecah, Samin, Mashitah, Izza, Anisa, Mashi, K izza, Yana, Adib, GEBArrian, Zumba friends, and the Malaysian community.

Finally, I would like to thank my DEBII colleagues for their invaluable contribution during different phases of the research: Dinusha, Bambang, Haji Binali, Olivia and Zia. Sean and Valencia, thank you for the coffee sessions that always cheered me up. I would also like to acknowledge the initial contribution made by other investigators in the Anti-Spam Research Lab: Pedram Hayati and Kevin Chai.

Truly, *“Allah does not burden a soul beyond more than it can bear (2:286)”*.

I pray that Allah reward you all accordingly.

Contents

Abstract.....	i
Acknowledgements	iii
Contents	v
List of Figures	x
List of Tables	xii
Chapter 1.....	1
Introduction.....	1
1.1 Overview	1
1.2 Spam in Web 2.0 Applications (Spam 2.0).....	2
1.2.1 Spam 2.0 vs. email spam	4
1.2.2 Spam 2.0 vs. web spam	6
1.3 Cost Models.....	6
1.3.1 To measure administrative consequences for business and evaluate the current situation/processes	7
1.3.2 To propose a business solution to improve the current models	7
1.3.3 To develop a basis for benchmarking.....	7
1.3.4 To cut costs strategically	7
1.3.5 To create expectations matching with resources.....	8
1.4 Public Awareness, Knowledge and Perception	8
1.5 Motivations for Research	8
1.5.1 Cost	8
1.5.2 Awareness.....	10
1.5.3 Knowledge.....	11
1.5.4 Perception.....	11
1.6 Significance of the Research	12

1.6.1 Social.....	12
1.6.2 Economic.....	13
1.6.3 Technical.....	14
1.6.4 Psychological.....	15
1.6.5 Environmental.....	16
1.7 Objectives of the Research.....	16
1.8 Plan of the Thesis.....	18
1.9 Conclusion.....	18
Chapter 2.....	20
Literature Review.....	20
2.1 Introduction.....	20
2.2 Cost Model Study.....	21
2.2.2 Other Related Cost Models.....	69
2.2.3 Summary of Literature Review on Cost Models.....	86
2.3 Public Awareness, Knowledge and Perception of Spam 2.0.....	87
2.3.1 Public Awareness.....	88
2.3.2 Knowledge.....	90
2.3.3 Perception.....	94
2.4 Open Issues.....	98
2.5 Chapter Summary.....	100
Chapter 3.....	101
Problem Definition.....	101
3.1 Introduction.....	101
3.2 Problem Definition.....	101
3.2.1 Cost.....	102
3.2.2 Awareness, Knowledge and Perception.....	103
3.3 Research Issues.....	103

3.3.1 Research Issue I: Developing Spam 2.0 Cost Model to Identify Related Costs.	104
3.3.2 Research Issue II: Insufficient Information and Exploration on Public Awareness, Knowledge and Perception Regarding the Topic of Spam 2.0.....	105
3.4 Research Methodology I: Design Science.....	106
3.4.1 Problem Definition.....	106
3.4.2 Conceptual Solution.....	106
3.4.3 Implementation, Test and Evaluation.....	107
3.5 Research Methodology II: Quantitative.....	107
3.5.1 Problem Definition.....	108
3.5.2 Survey Design.....	108
3.5.3 Data Collection and Distribution.....	108
3.5.4 Data Analysis and Assessment.....	108
3.6 Conclusion.....	108
Chapter 4.....	110
Conceptual Solution.....	110
4.1 Introduction.....	110
4.2 Overview of the Solution.....	110
4.3 Solution Description.....	112
4.3.1 Solution I.....	112
4.3.2 Solution II.....	125
4.4 Summary.....	127
4.5 Conclusion.....	127
Chapter 5.....	128
Spam 2.0 Cost Model.....	128
5.1 Introduction.....	128
5.2 Storage Cost.....	129
5.2.1 HoneySpam 2.0 Data Set Statistics.....	130

5.2.2 Storage Cost Survey	138
5.2.3 Estimating the Storage Cost.....	140
5.2.4 Discussion.....	142
5.3 Loss of Productivity Cost.....	143
5.3.1 Timing Data Set	144
5.3.2 Estimating Loss of Productivity Cost.....	150
5.3.3 Discussion.....	152
5.4 Conclusion.....	152
Chapter 6.....	153
Public Awareness, Knowledge and Perception of Spam 2.0	153
6.1 Introduction	153
6.2 Respondent Demographics.....	153
6.3 Descriptive Analysis	156
6.3.1 Awareness of Spam 2.0	156
6.3.2 Knowledge of Spam 2.0	161
6.3.3 Perception of Spam 2.0	172
6.3.4 Justification Comments	178
6.4 Discussion	187
6.5 Conclusion.....	189
Chapter 7.....	191
Conclusion and Future Work	191
7.1 Introduction	191
7.2 Problems and Issues	192
7.2.1 Issues with Spam 2.0 Cost Model	192
7.2.2 Issues with Public Awareness, Knowledge and Perception on Spam 2.0 Survey.....	193
7.3 Dissertation Contributions.....	193
7.3.1 Cost	195

7.3.2 Awareness, Knowledge and Perception 196

7.4 Limitations and Future Works 198

Bibliography 200

Appendices 208

Appendix I Web Survey Questionnaire 208

Appendix II Selected Publications 224

List of Figures

Figure 1.1: Examples of Spam 2.0 3

Figure 1.2: (a) Possible number of viewers in email domain, (b) Possible number of viewers in Web 2.0 domain 5

Figure 2.1: Spam volume received daily by different stakeholders 23

Figure 2.2: Latest email spam statistics as on 4 April 2012 [Source: (Spamcop.net 2012)] 23

Figure 2.3: Costs per employee reported for each type of cost 25

Figure 2.4: Costs reported for each type of cost involved in related organisations 25

Figure 2.5: Costs reported for each type of cost involved for related countries 26

Figure 2.6: Costs reported for each type of cost involved worldwide 26

Figure 2.7: Average time taken by an employee to review and delete an email spam 27

Figure 2.8: Average time taken by an employee to do related activities caused by email spam in a day 28

Figure 2.9 : Parasitic economics of spam [Source : (Cobb 2003)] 29

Figure 2.10: Email spam carbon footprint [Source: WebpageFX Team (2011)] 33

Figure 2.11: Generic cost model development methodology 34

Figure 2.12: Graphical summary of literature review on cost models 87

Figure 3.1: Multimethodological approach [Burstein and Gregor (1999) adapted from Nunamaker and Chen (1990)]. 109

Figure 4.1: Overview of the conceptual solution 111

Figure 4.2: Screenshot of survey 117

Figure 4.3: Screenshot of Question 1 118

Figure 4.4: Screenshot of Question 2 119

Figure 4.5: Screenshot of Question 3 120

Figure 4.6: Screenshot of Question 4 120

Figure 4.7: Screenshot of Question 5 121

Figure 4.8: Screenshot of Question 6 121

Figure 4.9: Screenshot of Question 7 122

Figure 4.10: Screenshot of Question 8 122

Figure 4.11: Screenshot of Question 9 123

Figure 4.12: Screenshot of Question 10..... 124

Figure 4.13: Survey design 126

Figure 5.1: Spam units involved in a discussion forum..... 130

Figure 5.2: Total number of spam units with its percentage according to month 135

Figure 5.3: Size for each spam units with its percentage according to month 136

Figure 5.4: Percentage of spam posts sent by spammers divided into six categories 137

Figure 5.5: Percentage of spam posts sent by spammers for the first category (0–10) 137

Figure 5.6: Steps to estimate time used for Spam 2.0 identification 147

Figure 6.1: Comparison between perceived awareness and actual awareness. 160

Figure 6.2: Comparison between perceived knowledge and level of knowledge 171

Figure 6.3: Basic statistics for comments 179

List of Tables

Table 2.1: Literature review on email spam cost models	38
Table 2.2 Unspecified method literature review on email spam cost models	43
Table 2.3 Frequency analysis of stakeholders.....	50
Table 2.4: Frequency analysis for cost categories	53
Table 2.5: Literature review on related cost models	70
Table 2.6: Frequency analysis for the cost categories	79
Table 4.1: Summary of proposed Spam 2.0 cost model.	113
Table 4.2: Questions' category and source.	118
Table 5.1: Spam profile	131
Table 5.2: Personal message(pm) spam.....	132
Table 5.3: Spam post	133
Table 5.4: Total spam in the discussion forum	134
Table 5.5: Storage cost	141
Table 5.6: Storage cost per MB per month	141
Table 5.7: Average size of spam units.....	142
Table 5.8: Average storage cost for 100,000 spam in a year	142
Table 5.9: Commercial filtering services.....	143
Table 5.10: Relationship between researcher and respondents' view of spam and non-spam and its costs.....	145
Table 5.11: Basic statistics of timing data set for each example.....	149
Table 5.12: Total time for related costs for spam identification	150
Table 6.1: Summary of respondent demographics	154
Table 6.2: Perceived awareness of online spam.....	156
Table 6.3: Actual awareness of online spam.....	158
Table 6.4: Respondents' score for actual awareness	159

Table 6.5: Respondents level of awareness 161

Table 6.6: Perceived knowledge of online spam..... 162

Table 6.7: Respondents’ data to assess knowledge on online spam-related activities 163

Table 6.8: Respondents’ data to assess knowledge on online spam. 167

Table 6.9: Respondents’ data to assess knowledge on spam identification 169

Table 6.10: Respondents’ level of knowledge 171

Table 6.11: Percentage of respondents’ perception towards crime 173

Table 6.12: Percentage of respondents’ perception on crime’s punishment..... 173

Table 6.13: Respondents’ perceived vulnerability to spam 174

Table 6.14: Respondents’ perceived likelihood of being spammed 175

Table 6.15: Respondents’ perception on seriousness of computer security problem 176

Table 6.16: Respondents’ perception on crime’s motivation..... 177

Chapter 1

Introduction

This chapter:

- ▶ Provides an introduction to Spam 2.0;
- ▶ Describes the differences between Spam 2.0, email spam and web spam;
- ▶ Gives an overview of cost models;
- ▶ Gives a short summary of the public awareness, knowledge and perception of Spam 2.0;
- ▶ Explains the significance and importance of this research;
- ▶ Explains the motivation for and objectives of this research; and
- ▶ Presents the organisation of this thesis.

1.1 Overview

Reliance on the Internet in today's daily life has brought the web to function as an important platform for generating and obtaining information and knowledge. However, the quality of content created on the web depends solely on the users who post content without being administered. In particular, the content that is unsolicited, inappropriate and irrelevant, called *spam*, has emerged as an important issue and a major concern for the Internet community as it can be manipulated by unscrupulous elements to their advantage. For example, *blogs* and *forums* can be used to post fake advertisements or misleading facts. Worse, these advertisements may lead to other security concerns and more serious problems. Additionally, spam posted on the Internet wastes storage and network resources. Because of all these, spam has a cost of its own and imposes costs on the Internet users and resources as well. Modelling all the costs associated with spam will enlighten the Internet community about the extent of the problem and enable necessary action to overcome it. By far, spam has seriously decreased the quality of information on the web. Though research on enhancing the techniques to combat spam has its importance, it is worthwhile to note that spam so easily infiltrates the web because the Internet users lack awareness and knowledge of the problem, and have differing perceptions about handling it.

1.2 Spam in Web 2.0 Applications (Spam 2.0)

The Web 2.0 platform provides for interactive information collaboration among the users and enables them to play the role of active content contributors. This platform, associated with web applications, offers openness and freedom to the users to add value to the web application. Such web applications include social networking sites, media sharing sites, wikis, blogs, forums and web-based communities. However, this platform, which promotes multi way collaboration instead of the traditional one way interaction, has also opened the doors for a problem called spamming. Traditionally, the term ‘spamming’ is closely related to email spam, which is an exploitation of the email domain. However, spam in Web 2.0 applications, called Web 2.0 Spam or simply Spam 2.0, is defined as “*propagation of unsolicited, anonymous, mass content to infiltrate legitimate Web 2.0 applications*”. Thus, Spam 2.0 is different from other types of spam as it is distributed through a legitimate website. Figure 1.1 shows some examples of Spam 2.0 found in Web 2.0 applications.

Spamming can be done using automated bots or manual spammers. The bots try to imitate the real users by manipulating information which the users are allowed to add and hosting it on Web 2.0 applications in the form of *spam units*. A spam unit is defined as an “*attribute that can be manipulated by spammers to embed their spam content*”. Such manipulation includes embedding unnecessary texts, images, hyperlinks, sounds, videos and file attachments. This information, or spam unit, neither enhances the quality nor increases the value of a page. In the email domain, spam is embedded in the email content. Therefore, the spam unit in the email domain is the email itself and its attachments. On the other hand, the spam units involved explicitly in different Web 2.0 applications are user profiles, posts, polls, tags, comments and personal messages.

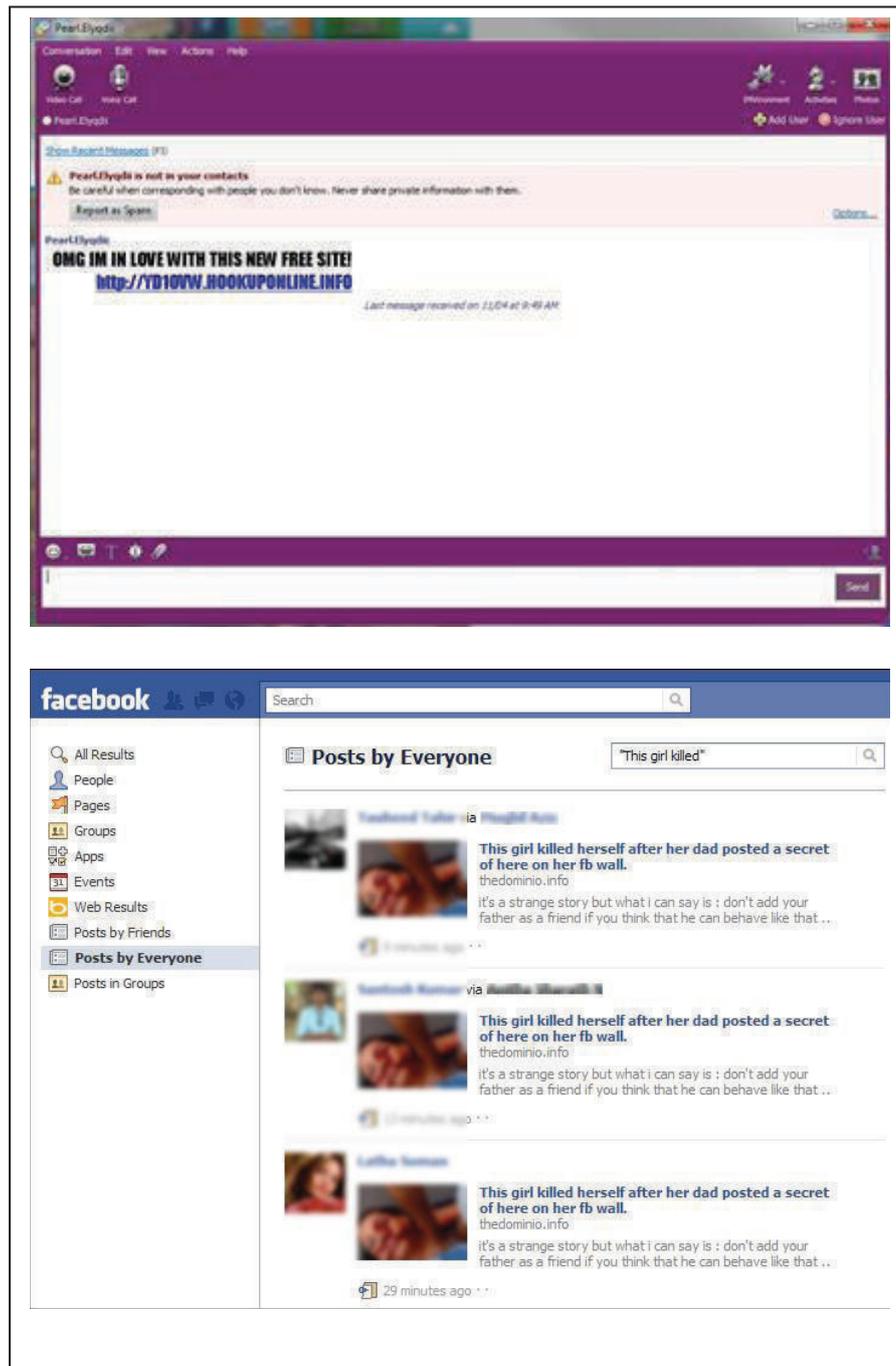


Figure 1.1: Examples of Spam 2.0

Basically, the purposes of both types of spamming activities, i.e., spamming through email and web domain, are similar. Spammers are driven to spam to generate revenues by obtaining a higher traffic to their sites, advertising their products and services, providing false information to the users, and stealing valuable information from them. Spam 2.0 has also been found to be a medium to spread spyware – a malware that leads to other security attacks, such as scams, phishing and fraud. Furthermore, Spam 2.0 is intended to infiltrate legitimate websites that provide genuine content including government organisations, universities and companies. Due to this reason, unknowledgeable users can easily be deceived by the spam content found on these websites, and thus, the reputation of

these legitimate websites may suffer. In the long run, Spam 2.0 could cause users to develop distrust for the information available on the Internet, which can affect the quality of the Internet services.

Previously, spamming in other domains, such as emails and instant messengers, was targeted at specific private spaces. Therefore, only the receivers were vulnerable to the spam content. However, the Web 2.0 platform offers the users publicly shared spaces where all the content can be viewed by all the users. This provides a bigger and larger reach for the spam content posted on the Web 2.0 applications, at a lower cost. Apart from a successful manipulation of the regular setting of the Web 2.0 platform, Spam 2.0 distribution rates are growing also because of the inadequacy of the existing anti-spam techniques, devised basically for email spam detection and prevention, in stopping the distribution of Spam 2.0. Thus, Spam 2.0 is able to cause several types of direct and indirect costs. First of all, the users are annoyed because of the inconvenience caused by such spam. Then, the reputation of businesses is put at stake and loss of business becomes very probable because of it. And finally, valuable computing resources, such as storage, network bandwidth, and human resources are wasted in storing, dealing and handling this type of spam, as its contents cannot usually be deleted by normal users and needs the site administrators to be extra vigilant and efficient.

Therefore, Spam 2.0 is more overwhelming than email spam or any other type of spam. The next sections further elaborate the differences between email spam, web spam and Spam 2.0.

1.2.1 Spam 2.0 vs. email spam

Compared to email spam, Spam 2.0 has a higher impact factor. The impact factor of each spam unit can be evaluated on the basis of two characteristics – the possible number of viewers and the attribute creation flexibility. The possible number of viewers is defined as “the likelihood of each attribute being viewed by the users” and is categorised into two types. High possible number of viewers means that the attribute is viewed by more than one user, while low possible number of viewers means that the number of viewers of the attribute is one or none. In the same way, the attribute creation flexibility is categorised into high, medium and easy, depending on the ease of creation and manipulation of each attribute by the spammers. An attribute is placed under the category of high attribute creation flexibility if it is easy to be created as well as manipulated. On the other hand, an attribute is categorised under medium attribute creation flexibility if it is easy to be created but hard to be manipulated. An attribute is placed under the category of low attribute creation flexibility if it is hard to be created as well as manipulated.

Theoretically, email spam has a 1:1 relationship, which means that an email is usually read by just one user, but Web 2.0 applications provide sharing services and are community based. Therefore, Spam 2.0 has a 1:M relationship which means that 1 spam unit in any Web 2.0 application can be read by many users, thus having a higher viewer impact. Figure 1.2 depicts these relationships for some most

popular Web 2.0 applications. For instance, profiles, wall posts, comments on wall posts, photos, comments on photos and personal messages are spam units in Facebook. Thus, it is clear that the built-in features of the Web 2.0 applications themselves allow the spammers to manipulate spam units to reach out to a wider viewership/readership faster and more easily.

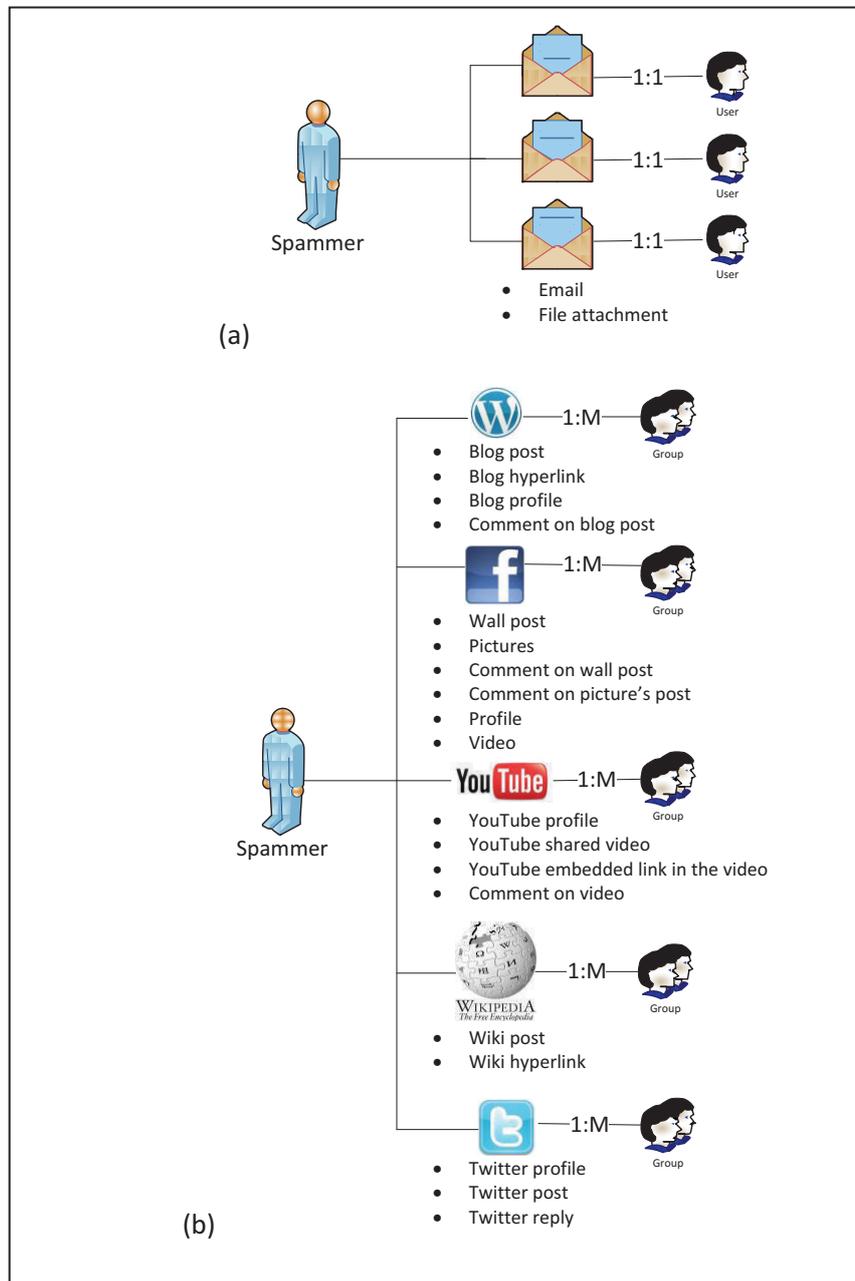


Figure 1.2: (a) Possible number of viewers in email domain, (b) Possible number of viewers in Web 2.0 domain

Although the attribute creation flexibility is pretty high in both, meaning that there are no limits to sending spam in both email and Web 2.0 domains, yet email spam is targeted only at active email addresses, while a single Spam 2.0 content embedded in a spam unit may reach multiple legitimate websites leading to a high possible number of viewers. A single email message, on the other hand, can reach only one potential victim, that too only if it is not filtered out, and is read only if the addressee

chooses to open it. This is one of the reasons why Web 2.0 applications offer an attractive platform for spamming activities.

The lifetime of a spam unit in both the domains is also different. In email spam, the email owner can directly delete the spam mail from the inbox. However, in a Web 2.0 application, general users are not allowed to delete others' posts. Hence, removing spam depends on the administrator. This could take a longer time and the spam could be overlooked hence impacting more viewers, as the spam content can be seen by all until it is deleted. Nonetheless, spammers nowadays are spamming both the domains as both are easily accessible and often linked to each other.

1.2.2 Spam 2.0 vs. web spam

Overall, the methods of spamming in web spam are also used in Spam 2.0. However, Spam 2.0 is targeted at legitimate websites, while web spam is commonly targeted at search engines to improve page ranks. Thus, Spam 2.0 goes beyond web spam; the information provided in it can mislead and trick the users.

This section explained Spam 2.0 and compared it with email spam and web spam. As mentioned earlier, Spam 2.0 has its own costs, imposes costs on others, and wastes resources. Therefore, the next section goes on to describe various cost models for spam.

1.3 Cost Models

Although it is clear that Spam 2.0 has its costs, no research has been carried out so far to make a clear estimate of these costs. Therefore, a cost model is required to estimate the costs involved in this type of spamming activity, including its business, economic and financial aspects. This section, therefore, first describes the importance of a cost model and relates it to the issue under consideration, i.e. Spam 2.0. A cost model is required to:

1. measure the administrative consequences for businesses and evaluate the current situation/processes;
2. propose a business solution in order to improve the current models;
3. help build a basis for benchmarking;
4. cut costs strategically;
5. create expectations matching with resources.

1.3.1 To measure administrative consequences for business and evaluate the current situation/processes

Spam 2.0 is a business for spammers. The possible expenditure on the spammers' side includes the cost of the software used to send automated spam, or the wages of the manual spammers hired. Their business model depends on the revenues generated from a successful spam campaign. A cost model will, therefore, help in identifying the real costs incurred by the spammers and understanding their economic model, bringing out their profits against the expenditure. Additionally, a cost model will also help in measuring the investments made in anti-spam software and its importance in handling spam. Thus, a cost model will help us evaluate the current overall situation of Spam 2.0.

1.3.2 To propose a business solution to improve the current models

A cost model also proposes a business solution in case any improvement is required on the current models. It helps in identifying the costs involved in any process defined in the cost model. Hence, a cost model for Spam 2.0 will help in identifying the costs involved and proposing an improved solution for each process.

1.3.3 To develop a basis for benchmarking

A cost model is an organised step-by-step approach to create a benchmark. Till date, to the best of our knowledge, no benchmark has been created to handle Spam 2.0. In order to create a benchmark, a study of several cost models is required. Thus, creating a cost model for Spam 2.0 is essential to help prepare groundwork for the benchmarking process.

1.3.4 To cut costs strategically

An evaluation of each business process will help ascertain if the costs incurred in the process are really necessary, or can be cut down. For the Spam 2.0 cost model, the target is to reduce the costs of the anti-spam sides to the lowest minimum, while creating an environment where it would be too costly for the spammers to send spam. This approach has already been suggested in several researches on email spam; however, implementing it without affecting the general users has been found to be quite difficult. Nonetheless, a cost model will help in identifying the excessive costs and possibly cut down on them strategically.

1.3.5 To create expectations matching with resources

A cost model is expected to help the customers to be clear about the products and services that can be delivered with a given level of funding, thus matching their expectations with the resources. In Spam 2.0, anti-spam products are needed to filter and handle spam, adding to the costs of the companies. Hence, a cost model will help them find out whether the amount of money they are paying to get the services is acceptable.

1.4 Public Awareness, Knowledge and Perception

Despite the grave problems caused by Spam 2.0, as described in Section 1.2, the issue is underrated or taken lightly. Protecting the Internet against Spam 2.0 attacks, however, is becoming increasingly important due to the threats it poses to the innocent web users. Nevertheless, the spammers' revenues rely heavily on the users' participation and decisions, and the users' decisions are influenced by their awareness, knowledge and perception of spam campaigns.

1.5 Motivations for Research

Though the current literature admits that Spam 2.0 has its costs, no such costs have officially been reported except by us (Ridzuan, Potdar et al. 2011). In the meantime, a lot of studies undertaken to combat Spam 2.0 have highlighted some key problems, leading to the need for a cost model. These problems have provided motivation for this research.

The main purpose of the research is to address the issue of the costs of Spam 2.0 and develop a cost model to estimate these. It also assesses the Internet users' awareness, knowledge and perception of Spam 2.0. In a nutshell, it addresses problems pertaining to the following four issues:

1) Cost, 2) Awareness, 3) Knowledge, and 4) Perception.

1.5.1 Cost

This section describes the problems that contribute to the need for a cost model for Spam 2.0. The problems are:

1. The costs of Spam 2.0 for all stakeholders are unclear.
2. The existing cost models are not fully applicable to Spam 2.0.

3. Spam 2.0 facilitates identity theft, leading to online frauds and scams and causing monetary loss.
4. Spam 2.0 damages reputation of legitimate websites with potential loss of business.
5. Spam 2.0 causes inconvenience to users.
6. Spam 2.0 is increasing at a rapid pace and its distribution seems to have gone out of control.
7. The number of Internet users is constantly increasing.

1.5.1.1 Ambiguity about the cost of Spam 2.0 for all stakeholders

Spam 2.0 burdens the stakeholders with costs which they have to bear. Even spammers have to invest their money to spam. Nonetheless, their economic model still enables them to gain profit (Hayati et al. 2012). However, this economic model does not work similarly for other stakeholders. Vast amounts of money, resources and energy are wasted to combat Spam 2.0, yet it is vague exactly how much the stakeholders have to spend for it. Therefore, it is important to produce a cost model to identify where the burdens come from and how to reduce them.

1.5.1.2 Inapplicability of the existing cost models to Spam 2.0

A number of cost models have been developed in previous researches (e.g. Nucleus Research 2003, 2004; Rockbridge Associates Inc. 2005; Nucleus Research 2007; Rockbridge Associates Inc. 2009). Nonetheless, the existing literature focuses mainly on the costs of email spam, as explained in Section 2.2.1.5, and the models proposed in it do not fully fit in with Spam 2.0 cases. Hence, there is a need to modify and combine the existing models to estimate the cost of Spam 2.0.

1.5.1.3 Identity theft facilitated by Spam 2.0, leading to online fraud and scams and causing monetary loss

It has been shown that a huge amount of money is lost due to online fraud and scams. These activities are often carried out by stealing identities through Spam 2.0 campaigns (Hinde 2002). This cost can be accounted under indirect costs caused by Spam 2.0.

1.5.1.4 Loss of reputation and business due to Spam 2.0

Spammers often disseminate fake advertisements through their spam campaigns (Hayati and Potdar 2009). To spread these fake advertisements, spammers often use the websites of trustworthy sources and companies. This could cause loss of reputation and affect the future business of these trustworthy sources and companies, even though they have nothing to do with the spam campaigns. In the long run, this could lead to reduced trust in Web 2.0 applications, and affect the credibility and quality of

information available on the Internet in general. This reputation cost and potential loss of business could also be considered as the indirect costs caused by Spam 2.0.

1.5.1.5 User inconvenience because of Spam 2.0

In the ongoing battle against spamming, many techniques have been introduced to stop spams from being propagated in Web 2.0 applications. However, most of these techniques cause user inconvenience. For example, flagging and notifying the administrator in the forums requires users to click the flag button and explain why they are reporting it as spam. CAPTCHA wastes the users' time in entering the characters. Moreover, the CAPCHAs are getting increasingly tougher and users usually have to redo them over and over again (Ridzuan and Potdar 2012; Potdar et al. 2010). This not only consumes the users' time but also causes annoyance to them.

1.5.1.6 Uncontrollably increasing rate of Spam 2.0 distribution

The distribution of Spam 2.0 is increasing every year (Akismet 2013; Sophos 2008, 2010, 2011). With the availability of new technologies, such as auto submitters and web spambots (Hayati et al. 2009), it has become easier for spammers to spread their spam campaign, further increasing their success rate and profits. As these technologies add to the spammers' revenues, it will be an interesting study to find out how these costs affect other stakeholders.

1.5.1.7 Increasing number of Internet users

There are new users being introduced to the Internet every day, and thus, the number of Internet users keeps on growing by the day. With the increase in their number, the number of Web 2.0 applications users is also increasing every day. The new users are naturally more vulnerable to spamming and can easily become the spammers' victims. Therefore, a cost model to quantify the costs per person is the need of the hour.

1.5.2 Awareness

The thesis now comes to describing the problems related to awareness, in order to bring out the need for a study of the awareness, knowledge and perception of Spam 2.0. The problems are:

1. Spam 2.0 is not fully understood by the Internet users; and
2. Spam 2.0 proliferates because of the users' lack of awareness.

1.5.2.1 Lack of knowledge of Spam 2.0 on the part of the Internet users

Not all Internet users have the basic knowledge of cyber security threats or are aware of the threat posed by Spam 2.0. Most of them are general users but still using Web 2.0 applications. If the Internet users have a better understanding of the problem, they would be empowered to avoid its trap. Moreover, they could contribute to combating spam instead of becoming its victims. However, most Internet users are unaware that Spam 2.0 could be the source of scam, fraud, virus and other security threats.

1.5.2.2 Role of the users' lack of awareness in proliferation of Spam 2.0

Spam 2.0, because of its very nature, can quickly proliferate to a wide network of users. For example, a spam post in Facebook can proliferate in the network in just a few minutes after being posted, depending on the number of viewers (Jeffries 2008; Brandt 2010; Scam Sniper 2011; Cluley 2012). It is likely to be seen by many until it is reported as spam or removed by the person who posted it. However, in many cases, even the persons who post spam are not aware that the content they have posted is a spam. Thus, it is undeniable that a broad awareness is the only realistic way to counter spam. Even though it cannot directly stop spam, at least the users could avoid inadvertently spreading it.

1.5.3 Knowledge

A study of the knowledge of Spam 2.0 on the part of the Internet users is also required. One of the main reasons for the proliferation of Spam 2.0 is the users' lack of knowledge of the issue. Therefore, it is important to check the perception of the Internet users on spam. Due to an insufficient knowledge, some of them might be under the impression that spam itself is not a significant problem (Ridzuan, Potdar, and Hui 2012). Knowledge could enable the users to circumvent spam and avoid contributing to spam campaigns as well as motivate them to report spam activities, thus speeding up the identification and deletion of spam entries in Web 2.0 applications. The wild proliferation of Spam 2.0 is due in no small measure to the inadequacy of user knowledge.

1.5.4 Perception

The need to address the Internet users' perception of Spam 2.0 arises because of the following problems:

1. Spam 2.0 is not usually recognised as a new type of spam.

2. Users' perception of the trustworthiness, credibility and quality of information on Web 2.0 platforms is compromised because of Spam 2.0.

1.5.4.1 Not recognising Spam 2.0 as a new type of spam

The common perception seems to be that Spam 2.0 works similarly as email spam, and the techniques used to combat email spam can also be used to combat Spam 2.0. However, compared to email spam, Spam 2.0 has a higher viewer impact and cannot be tackled by the techniques used to combat email spam (Ridzuan, Potdar, and Singh 2011). Thus, this problem is underrated.

1.5.4.2 User's perception of the information on Web 2.0 platforms getting compromised because of Spam 2.0

The rapid increase in online fraud and scams originating from spam-related campaigns underlines the need to check the users' perception of the trustworthiness, credibility and quality of information available on Web 2.0 applications. Trust is the basis of the functioning of Web 2.0 applications; users engage in Web 2.0 applications-based activities with a certain level of confidence in the credibility and quality of information provided by other Internet users. Spam 2.0 could decrease this level of confidence and put the very existence of this trust-based relationship at stake.

Thus, it is clear that there is a need for a cost model study specifically for Spam 2.0, and to assess the awareness, knowledge and perception of Spam 2.0 among the Internet users.

1.6 Significance of the Research

Section 1.5 listed the motives for this research. This section elaborates the significance of the research in greater detail. The problems the research addresses can be divided into five categories as follows:

(1) Social, (2) Economic, (3) Technical, (4) Psychological, and (5) Environmental.

1.6.1 Social

1.6.1.1 Providing a comprehensive and unbiased report for anti-spam communities

The aim of this research is solely to provide an unbiased report for educational purposes in order to help the anti-spam communities to combat spam. Till date, most spam cost reports have focused on email spam. This research, on the other hand, focuses on the cost of Spam 2.0.

1.6.1.2 Getting the severity of the problem of Spam 2.0 acknowledged

Findings of this study are expected to provide a Spam 2.0 cost model to estimate the total cost of Spam 2.0. This will give an idea of how much cost Spam 2.0 imposes on the stakeholders, and have people acknowledge the seriousness of the problem. Having recognised the seriousness of the problem, users will not hesitate to contribute in helping to identify and report spam posts when they come across one.

1.6.1.3 Enhancing the awareness and knowledge of the Internet community of Spam 2.0

Most prior research in the subject has focused on strengthening the techniques to combat Spam 2.0, and very limited work has been done on awareness and knowledge of Spam 2.0. Although these techniques would certainly help prevent Spam 2.0 from being propagated, without a sufficient level of awareness and knowledge on the part of the users, the problem cannot easily be solved. Therefore, this research will benefit the Internet community by educating the users on Spam 2.0.

1.6.1.4 Reducing spam proliferation rate and improving the quality of the content on the web

The rate of proliferation of Spam 2.0 depends on the users' involvement by clicking. Therefore, the proliferation rate can be reduced if the users are made aware of the problem and educated on dealing with spam posts. This will also induce users to notify the spam to the administrators of the application they are using, hence indirectly maintaining and improving the quality of the available content on the web.

1.6.1.5 Contribution to Australian National Broadband Network (NBN) initiative

As pointed out above, if people are aware of and knowledgeable about spam and spam related problems, they will be able to distinguish spam with non-spam easily. This could indirectly reduce spam proliferation, or at least slow down its rate. Thus, the resources provided by the Australian NBN will not be wasted on spams.

1.6.2 Economic

1.6.2.1 Combating spam economically

The research addresses the spam cost issues from different perspectives with a focus on each stakeholder, and can then further be used in developing a cost model. This cost model can be utilised

to make the costs for spammers higher than those for anti-spammers, so that the spammers would get no profits from spamming and thus lose interest in it, which will be perhaps the best solution to the problem. However, even if that does not happen, this cost model can be utilised to make it affordable to combat Spam 2.0.

1.6.2.2 Changing economic model of the spammers

Spammers are said to send spam without having to spend too much money. This research will investigate the matter in depth. Having identified the real costs for the spammers, the research could further be used to change the spammers' economic model to increase the costs of sending spam for them.

1.6.2.3 Preventing users from becoming spammers' victims and suffering monetary loss

As a result of this research, more Internet users are expected to become aware of and knowledgeable about Spam 2.0. Therefore, they will be able to deal with spam effectively and would not easily be led by such campaigns or become their victims. The research will also enhance their awareness and knowledge of spam-related problems; thus, they would not get involved in scams and fraud, and would be able to avoid monetary losses.

1.6.2.4 Reducing loss of productivity

Spam causes the users to spend their time to identify and delete spam posts. If a user is an employee, online in the working hours, this amounts to a loss of productivity for the company. This research is expected to increase the users' awareness and knowledge of spam, so that they wouldn't fall easily into the spammers' tricks. Thus, the costs of the loss of productivity will decrease.

1.6.3 Technical

1.6.3.1 Proposing a new cost model for Spam 2.0

As mentioned earlier, many existing studies have developed email spam cost models which, however, cannot easily be applied to Spam 2.0. To the best of our knowledge, this research will be the first work to deal with Spam 2.0 costs.

1.6.3.2 Embedding new time-factor questions in the survey to assess the users' knowledge

The methodology used in this research includes a web survey for data collection. However, unlike traditional questionnaires and other prior researches, the timing function has been embedded in the questions to assess the users' knowledge, and the results have been presented comparatively using several questions

1.6.4 Psychological

1.6.4.1 Enhancing confidence in the use of Web 2.0 applications

The research will address the issues of user awareness, knowledge and perception. Predictably, users will be more confident in managing Spam 2.0 once they come to know what it is all about. Hence, by making them knowledgeable, this research will contribute to their confidence of not becoming the victims of spamming activities.

1.6.4.2 Reducing user annoyance

Users are often annoyed with spam. As pointed out in Sections 1.5.1.3 and 1.5.1.4, this research is expected to reduce the propagation of Spam 2.0, thus decreasing user annoyance while using Web 2.0 applications.

1.6.4.3 Enhancing concentration in the work environment

If the users are aware of Spam 2.0, they will not fall for the spam campaigns and not be drawn to waste their time on them. This will enhance their concentration during working hours. Psychologically, users will be able to better focus on their work and thus enhance their productivity.

1.6.4.4 Enhancing trust in the services provided by Web 2.0 applications

At present, users are in the process of building their trust in Web 2.0 applications. This study will predictably reduce Spam 2.0 proliferation on the network as also users' clicks on spam campaigns, as the Internet users' awareness and knowledge increase. Thus, it will help speed up the process of building trust and enable the users to use the services provided on Web 2.0 applications without any sense of doubt.

1.6.4.5 Enhancing user confidence in the quality and credibility of information available on Web 2.0 applications

As mentioned earlier, Spam 2.0 reduces the quality and credibility of information on the Internet. However, once the users are aware and have adequate knowledge, they will be able to differentiate between the fake and real services on Web 2.0 applications. As mentioned in Section 1.5.1.4, this research is expected to enhance the quality of information on the web, and thus enhance user confidence in such information.

1.6.5 Environmental

1.6.5.1 Reducing the carbon footprint of spam

As the knowledge and awareness level of the Internet users increases, the spammers will expectedly lose their motivation for spamming. This will reduce the impact of the carbon footprint of spam by reducing the energy unnecessarily wasted on spam activities. At the very least, users will stop clicking on or promoting spam campaigns; hence the carbon footprint generated from these activities will certainly be reduced.

1.7 Objectives of the Research

Section 1.6 discussed the significance and contributions of this research. This section enlists the main goals to be achieved from it as follows:

- To develop a cost model for Spam 2.0.
- To study the public awareness of Spam 2.0.
- To study the knowledge of the Internet users of Spam 2.0.
- To study the public perception of Spam 2.0.

To achieve the goals and solve the problems, as mentioned in Section 1.4, 6 objectives have been derived as follows:

Objective 1: To cull out and evaluate the existing cost categories and parameters from the existing cost studies

The prior literature mainly focuses on the cost of email spam. Furthermore, it specifically covers several countries but not Australia. The cost studies have usually been carried out by commercial

companies. Even so, these studies can provide valuable input for our cost model. Thus, the first objective is to cull out and evaluate all cost categories and parameters from the literature to be modified, combined and used in this research.

Objective 2: To develop a cost model for Spam 2.0 by estimating the storage cost and loss of productivity cost

Under this objective, the thesis first proposes a cost model based on the existing research works to fit all types of Web 2.0 applications. The dataset provided by the previous researches in the form of a discussion forum has been used and the cost of storage has been estimated on the basis of current market prices. The thesis then estimates the cost of loss of productivity to identify Spam 2.0 from the analysis of a survey of the public users. This survey determines the time taken to identify Spam 2.0 based on the knowledge of the users. All cost categories have been combined into a cost function which is then generalised to fit in with all types of Web 2.0 applications and form a cost model for Spam 2.0. This cost model covers all stakeholders and the cost calculations enable a study of the impact of Spam 2.0.

Objective 3: To validate the cost model

Under this objective, the proposed cost models have been validated using other spam datasets and expert reviews

Objective 4: To produce a public report for the Spam 2.0 cost model

Under this objective, the cost model having been validated, has been reproduced in the form of a public report.

Objective 5: To study the existing literature on the awareness, knowledge and perception of Spam 2.0 on the part of the Internet users.

The current literature lacks in studies of the awareness, knowledge and perception of Spam 2.0. Therefore, an online survey questionnaire has been created and circulated among general Internet users in an attempt to address these three issues. This survey assesses the users' knowledge using the timing function questions. Nonetheless, the assessment does not rely on these timing function questions alone.

Objective 6: To provide a public report on the awareness, knowledge and perception of Spam 2.0 on the part of the general users

Under this objective, the data obtained from the survey have been analysed and reproduced in the form of a public report.

1.8 Plan of the Thesis

The thesis comprises seven chapters. The remaining six chapters are structured as follows:

Chapter 2 provides an in-depth review of the related research on cost models, including email spam and other related cost models, such as those for the IT department, cloud computing and data centre. This is followed by a detailed comparison of all cost models and evaluation of the cost categories, parameters and attributes used in them. A comprehensive literature on the issues of awareness, knowledge and perception has also been included in the second part of the chapter, along with brief definitions of the three terms.

Chapter 3 defines the main research problems identified after the literature review. Having identified the clearly-defined research issues, the chapter goes on to describe the two approaches used for the research.

Chapter 4 proposes a conceptual framework for research based on the two approaches. It presents two solutions and constructs the conceptual model for each.

Chapter 5 presents the cost model for Spam 2.0 based on the results, including those of the storage and loss of productivity costs.

Chapter 6 presents the results obtained from the survey on the awareness, knowledge and perception of the Internet users and reports the main descriptive statistics.

Chapter 7 summarises the key findings of the thesis. This includes a discussion of the results and the major benefits of the research, along with a conclusion and suggestions for future work in the field.

1.9 Conclusion

This thesis studies the costs of Spam 2.0 and assesses the awareness, knowledge and perception of the users with a view to harnessing these to reduce its distribution rate. Accordingly, this chapter provided a brief introduction to the concept of Spam 2.0 and pointed out the differences between Spam 2.0 and other types of spam. It also explained why Spam 2.0 is an attractive way for spammers to run their spamming activities and generate revenues, and discussed the importance of a cost model to make it unprofitable or less profitable for them. Finally, the chapter touched on the aspects of public awareness, knowledge and perception which play an important role in Spam 2.0 proliferation.

The next chapter will carry out an extensive literature review to study different cost models in order to come up with a cost model for Spam 2.0. A detailed literature review on the awareness, knowledge and perception issues will also be included in it.

Chapter 2

Literature Review

This chapter carries out:

- ▶ A survey and evaluation of cost model studies, including email spam cost models and other related cost models like those of the IT department, data centre and cloud computing;
- ▶ A survey and evaluation of the awareness, knowledge and perception of the Internet users regarding Spam 2.0; and
- ▶ A discussion of the open issues in both – cost model studies and awareness, knowledge and perception of Spam 2.0.

2.1 Introduction

This chapter carries out a review of the previous literature on spam related issues and evaluates the cost models proposed in it. It also presents an overview of the issues related to awareness, knowledge and perception of spam. Thus, the literature surveyed in this chapter can be categorised into two:

- **Cost model studies** – This section provides a review of the related cost models, such as email spam cost model and others, and identifies the issues making them inapplicable to Spam 2.0.
- **Awareness, knowledge and perception studies** – Most of the prior literature has focused on improving methods to identify spam. However, awareness, knowledge and perception of the Internet users are also important in reducing spam 2.0 proliferation. Therefore, this section provides a survey of the existing literature on awareness, knowledge and perception of spam.

The preliminary concepts used in this thesis are defined below:

- **Cost category:** Cost category is a generic classification for a value that is wasted/spent in dealing with spam. Schadler (2009) describes his cost categories as hardware, server software, client software, storage message filtering, message archiving, mobile messaging, staffing and financing.
- **Cost parameters:** Cost parameters are the basic variables that are measurable, act as an input for the cost category function, and affect the total value for the related cost category.

- **Cost function:** Cost function is a function of input parameters and output value for the cost category. Each cost category has a different cost function.
- **Cost model:** Cost model is the sum or combination of several cost functions that mix all the different cost categories measured into a total single value. In this thesis, the main objective is to produce a cost model that will account for the cost of Spam 2.0.
- **Stakeholder:** The classic definition of stakeholder as given by Freeman is, “*A stakeholder in an organisation is (by definition) any group or individual who can affect or is affected by the achievement of the organisation’s objectives*” (Freeman 1984). In this thesis, a stakeholder is defined as any group or individual who can affect or is affected/impacted by spamming activities. Stakeholders can be individuals (such as spammers or users) or larger groups (such as organisations, countries or the whole world). Stakeholders can be grouped into two, spammers and non-spammers. Non-spammers can be personal users, organisations, countries or the whole world.

2.2 Cost Model Study

This section provides a definition of email spam in subsection 2.2.1.1, followed by the basic background studies including statistics and cost of email spam in Section 2.2.1.2. Subsection 2.2.1.3 focuses on the impact of email spam, including economic, social, psychological, technical and environmental impacts. Subsection 2.2.1.4 goes on to describe the generic methodology used in the existing cost models and subsection 0 further explains this methodology. Finally, subsection 2.2.1.6 carries out a critical analysis comparing all the cost models found in the literature.

2.2.1.1 Definition of email spam

Email spam, commonly just called spam, is defined as unsolicited bulk email. Technically, spam has been defined by Spamhaus (2012) as any email received “*if (A) the recipient's personal identity and context are irrelevant because the message is equally applicable to many other potential recipients; AND (B) the recipient has not verifiably granted deliberate, explicit, and still-revocable permission for it to be sent.*” Thus, it is evident that spam is mainly about consent.

Nevertheless, it is very difficult to distinguish between spam and non-spam, because a spam for one might not be spam for others since it depends on the content of the mail and needs of the individual (Ridzuan and Potdar 2012; Potdar et al. 2010). For example, an individual who receives an email advertisement from an unknown source selling things they want to buy, or otherwise useful to them, might not consider that email as spam. On the other hand, others who are not interested in the advertisement, or do not find the mail otherwise beneficial, may just discard it as spam. Hence, there is a grey area between spam and non-spam.

2.2.1.2 Background studies

This section first presents the statistics on email spam. Next, some of the cost values stated in the previous literature have been pointed out. These costs are valued in time and/or monetary aspects.

Volume

Statistics on email spam are mostly reported by companies that sell products and services dealing with spam. One of the earliest cost studies in the field of email spam was done by Gartner Group in 1999. Their survey reported how the recipients felt about spam and whether they lodged complaints about it. In all, 31% had received 11 or more email spams per week and 91% at least one email spam per week (Gartner Group 1999). According to the survey, 42% of the respondents disliked spam chiefly because it required time to read and discard it, 35% considered it an assault on privacy, and 15% found it offensive (Gartner Group 1999).

Later, several organisations gave a more detailed insight into spam propagation rate by revealing spam filtered by them. For example, Microsoft filtered 2.4 billion email spams daily to stop them from reaching its clients' inboxes in MSN and Hotmail servers (Gates 2003). Another ISP, AOL, also blocked a similar amount of spam (Rowland 2003) coming to 67 emails per inbox per day. These spam mails accounted for 80% of the incoming email traffic of their servers (Rowland 2003). According to Radicati Group, in 2004, approximately 15 billion spam emails were transmitted every day worldwide (Rowland 2003). Nucleus Research reported twice the number of spam messages per employee received in 2004 (7,500) compared to 2003 (3,500) (Nucleus Research 2003, 2004).

Nevertheless, in 2007, on an average, the Nucleus Research reported each user receiving 21 spam messages per day, showing a reduction from 29 in 2004 (Nucleus Research 2007). A comparison of the volume of spam received daily by each employee, company, and worldwide, as reported by several companies such as Microsoft, AOL, Radicati and Nucleus Research is presented in Figure 2.1.

Sophos revealed that, in the first quarter of 2008, 92.3% of their email was spam (Net-security 2008). In 2009, a report from Ferris Research stated that more than 75% of the emails sent on a daily basis were spam (Ferris Research 2009). This figure aligns with the report of Radicati Group which put spam at about 81% of the total email traffic (Sara Radicati 2009). Commtouch reported that, on an average, throughout the first quarter of 2010, 80% of its email traffic was spam (Commtouch 2010). However, Symantec, in its January 2012 report, stated that the amount of spam in its email traffic had decreased significantly and was only 69% now (Symantec Intelligence 2012b). In its February 2012 report, it reported a further decrease by 1%, to 68% of its total email traffic (Symantec Intelligence 2012a). The latest statistics on email spam provided by SpamCop have been given in Figure 2.2.

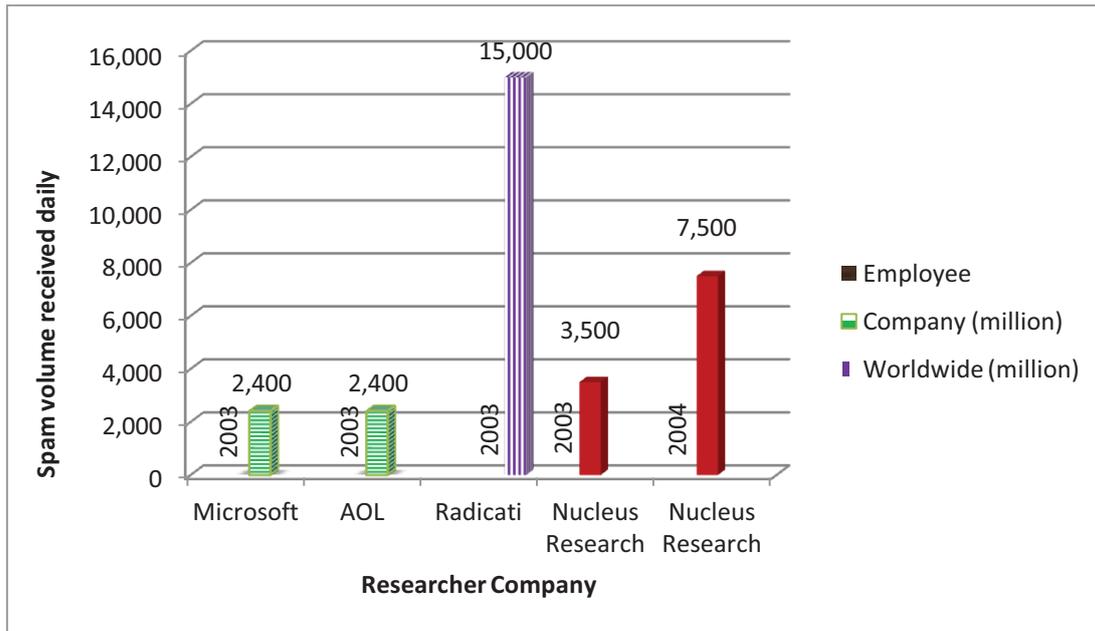


Figure 2.1: Spam volume received daily by different stakeholders

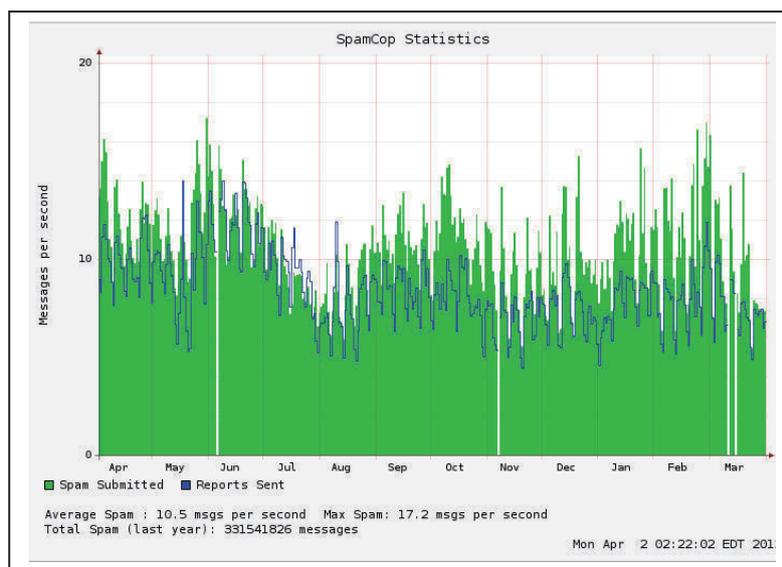


Figure 2.2: Latest email spam statistics as on 4 April 2012 [Source: (Spamcop.net 2012)]

Cost

As revealed by the statistics above, the propagation of email spam never ends. Therefore, concerns have arisen about the impact of spam and its costs for organisations. The Windows & .NET magazine, in its study covering storage cost, bandwidth cost, support cost and loss of productivity cost, stated that the total annual cost of email spam is nearly \$2.4 million, with 98% of it coming from the loss of productivity cost (Windows & .NET Magazine 2003).

According to Nucleus Research, the average annual cost of spam for each employee was \$874 in 2003 (Nucleus Research 2003). Using similar methods for data collection, the organisation reported that, compared to 2003, the costs for average productivity loss per employee per year doubled to roughly

\$1934 (Nucleus Research 2004). In 2007, its survey found that the cost of the loss of employees' productivity due to email spam was \$71 billion per year (Nucleus Research 2007). On the other hand, in 2004 and 2009, the National Technology Readiness Surveys (NTRS) prepared by Rockbridge Associates, estimated the cost value of email spam based on the time taken by users to delete spam mails, rating it on the basis of the average US salary, quantified the total loss of productivity cost at \$21.58 billion per year (Rockbridge Associates Inc. 2005, 2009)

As reported by (Morrissey 2003), Ferris Research found that the monetary cost of email spam for corporate organisations in 2002 was US\$8.9 billion for US and US\$2.5 billion for Europe respectively. Ferris Research also stated that the worldwide productivity cost of spam and other expenditures had been estimated at \$50 billion in 2005, with \$17 billion accounted for by US organisations (Ferris Research 2005). Compared to the 2003 figure of \$10 billion accounted for by US organisations, this number shows an increase of \$7 billion. Another research by a similar company taking the worldwide view showed that email spam costs in 2009 were \$130 billion (Ferris Research 2009; Jennings 2009).

In 2004, as reported by Ukai and Takemura (2007) and Takemura and Ebara (2008), Japan bore the cost of email spam in terms of labour and capital losses to the tune of US\$ 17 billion and US\$ 22 billion respectively. Although the impact of reduction in Japan's GDP by 500 billion yen due to these losses comes to only 0.1% of the GDP in 2004, yet if sufficient actions are not taken, this could lead to a greater impact (Ukai and Takemura 2007). Ukai and Takemura used a production function and found that the efficiency of labour was affected because of processing email spam (Takemura and Ebara 2008). They further produced results showing that email spam also reduced labour productivity and affected Japan's economy directly and indirectly.

It is notable that some of the cost values mentioned above have been generated through simple calculations. Additionally, some of the surveys were done in commercial companies that have not revealed their methods to the public. Furthermore, these cost values are either company based, or country based taking only two countries – the US and Japan. Nevertheless, the values above have been generated through surveys and a combination of several reports, estimates and adjustments. Thus, these statistics do prove that email spam is a big nuisance to the Internet community. The next section will further describe how it affects the community.

A summary of all the annual costs involved for different stakeholders has been presented in Figures 2.3-2.6 below. Figure 2.3 shows costs based on the existing reports from Nucleus Research and Radicati Group with different types of costs from 2003 to 2010. This figure focuses on the cost for each employee, hence taking each employee as a stakeholder. Figure 2.4 shows the values of different types of costs reported by Windows & .NET in 2003. This figure focuses on organisations as stakeholders. Values for the related costs involved, as reported by several companies, have been

illustrated in Figure 2.5. The values were reported for different countries; hence, this figure focuses on countries as stakeholders. Figure 2.6 shows the values reported for worldwide costs by Ferris Research.

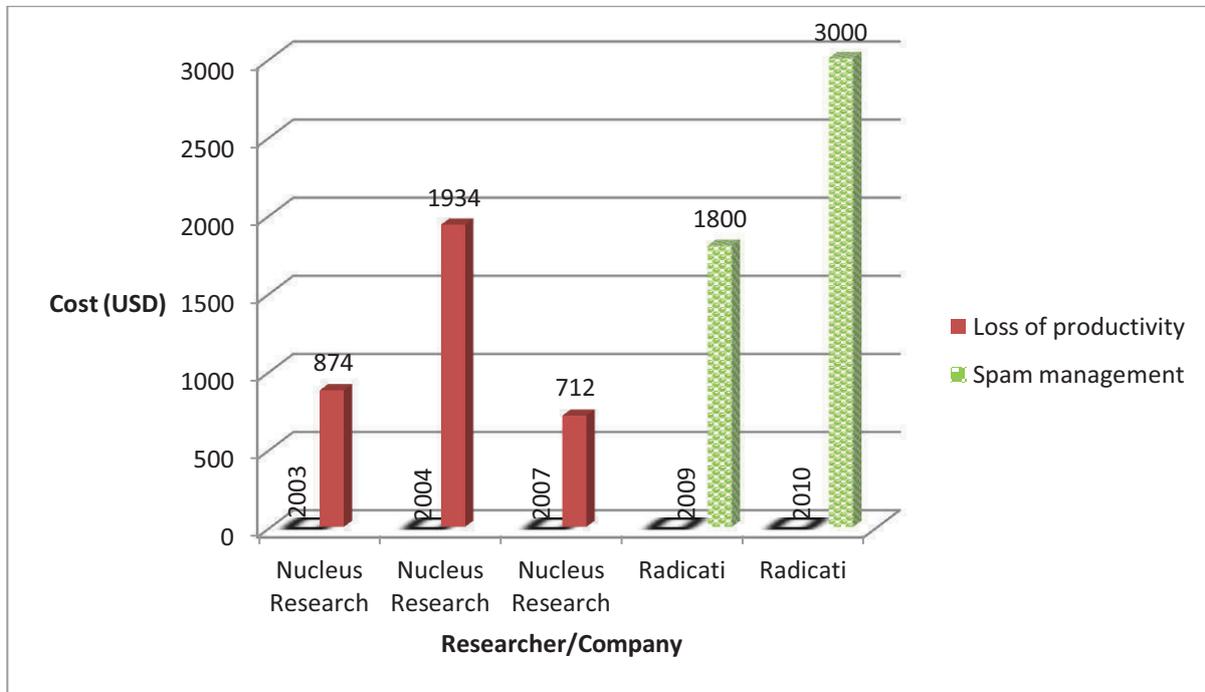


Figure 2.3: Costs per employee reported for each type of cost

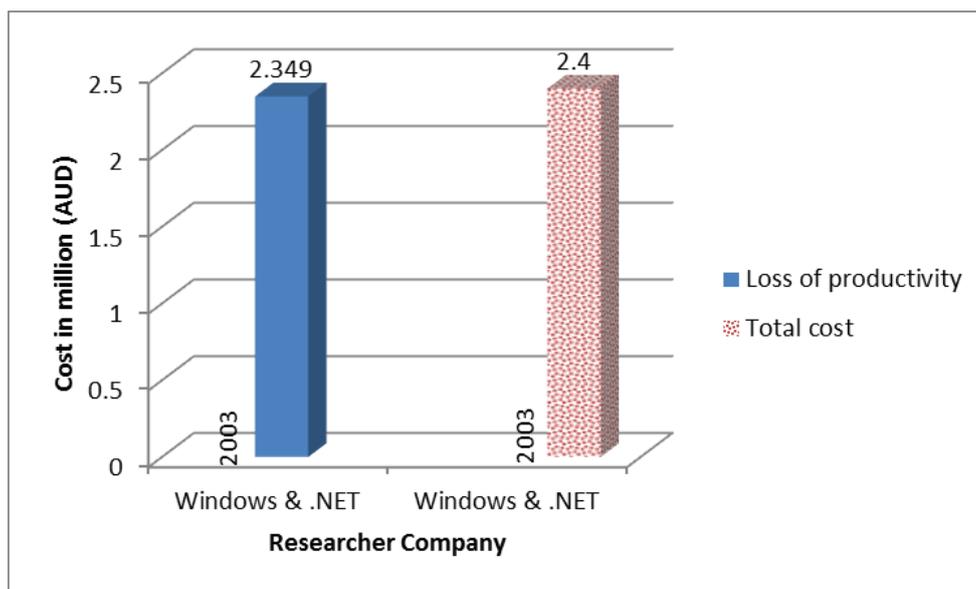


Figure 2.4: Costs reported for each type of cost involved in related organisations

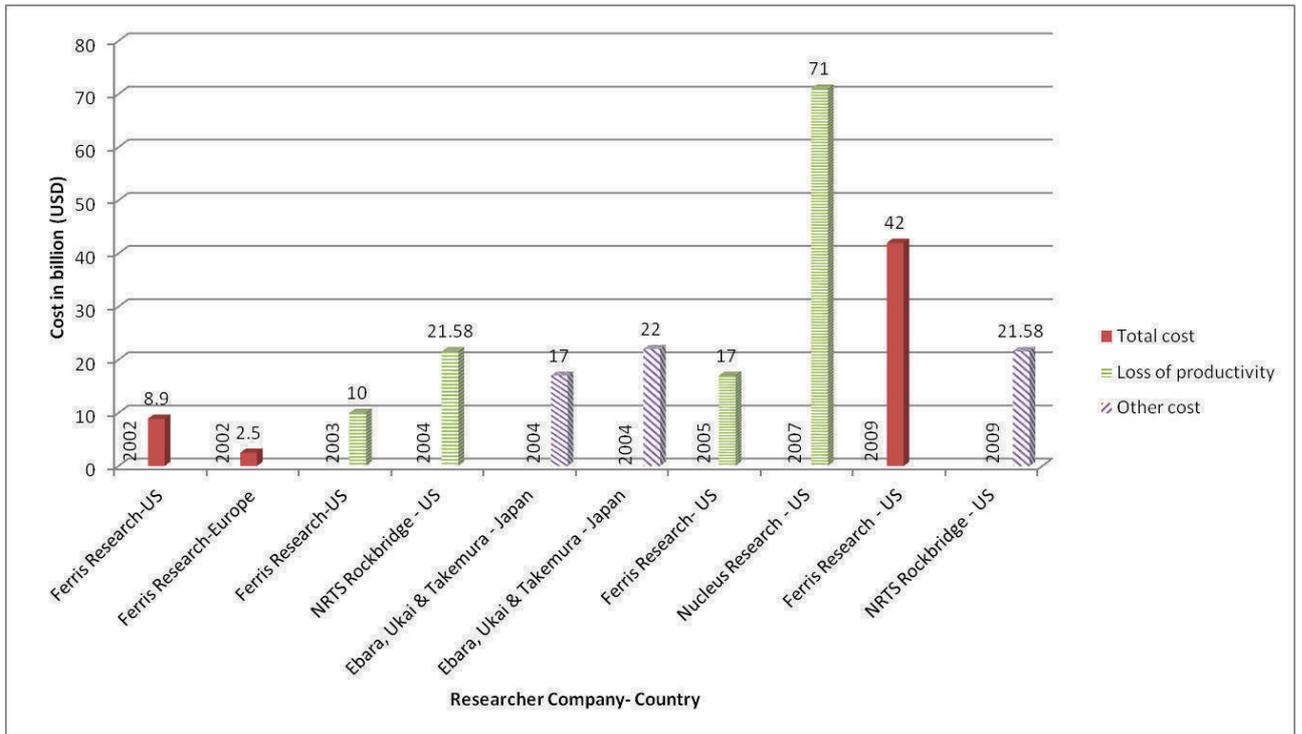


Figure 2.5: Costs reported for each type of cost involved for related countries

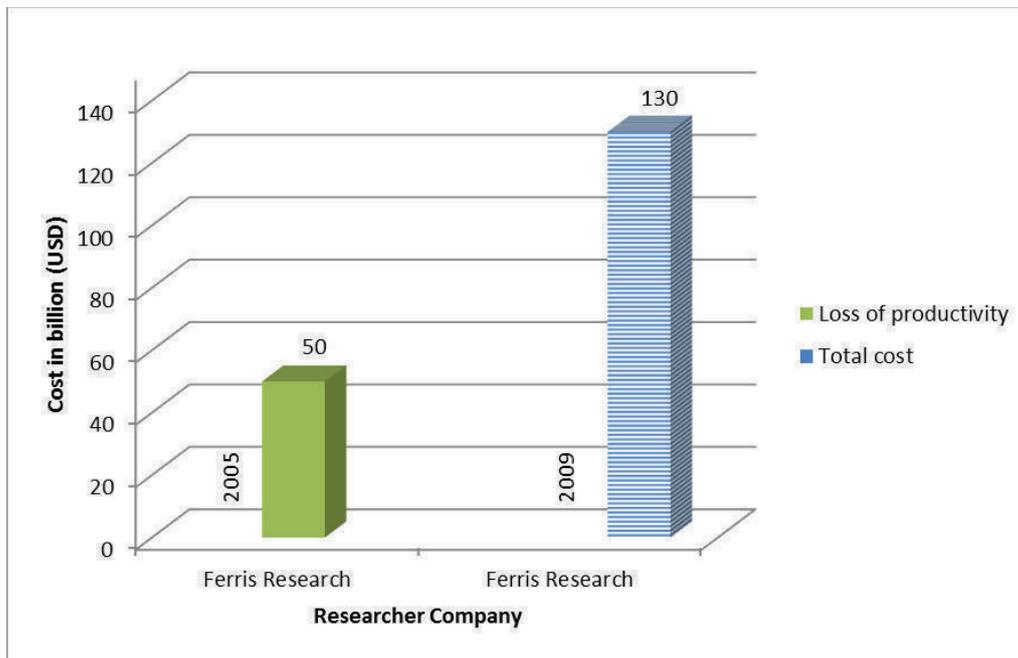


Figure 2.6: Costs reported for each type of cost involved worldwide

Time

Several studies have reported the time taken by Internet users to review and delete spam emails received by them. The value has been reported either in seconds taken by each user to manage each email spam, or total minutes taken by each user for email spam management in a day. This value then becomes one of the input parameters to calculate the cost of productivity loss based on the assumption

that, if this amount of time was not used by the employee to review and delete spam, it would actually have been allocated to work.

Windows & .NET used 5 seconds in its survey as the average time taken by 12,000 employees to process each spam email (Windows & .NET Magazine 2003). Nucleus Research, in its 2003 survey, reported that each employee took a total of 6.5 minutes per day to review and delete email spam messages. In 2004, Nucleus Research reported that, on an average, an employee used 30 seconds to review and delete an email spam (Nucleus Research 2004). This value declined in its 2007 survey as it was reported that, on an average, an employee took 16 seconds to review and delete an email spam (Nucleus Research 2007), possibly because employees are getting smarter in dealing with spam.

According to NTRS 2004 survey, the average time taken to manage email spam messages by the Internet users was 2.8 minutes per day (Rockbridge Associates Inc. 2005). This value decreased in its 2009 report which stated that, on an average, 1.4 minutes per day was spent on managing email spam (Rockbridge Associates Inc. 2009).

Figure 2.7 shows the average time taken per employee to review and delete an email spam. These values were reported by Windows & .NET and Nucleus Research in 2003, 2004 and 2007. On the other hand,

Figure 2.8 shows the average time taken by an employee to do related activities caused by email spam in a day. These values were reported by Nucleus Research and NRTS Rockbridge surveys. However, it is important to take note of the fact that these values were not generated on the basis of a similar average number of email spam received by an employee in a day.

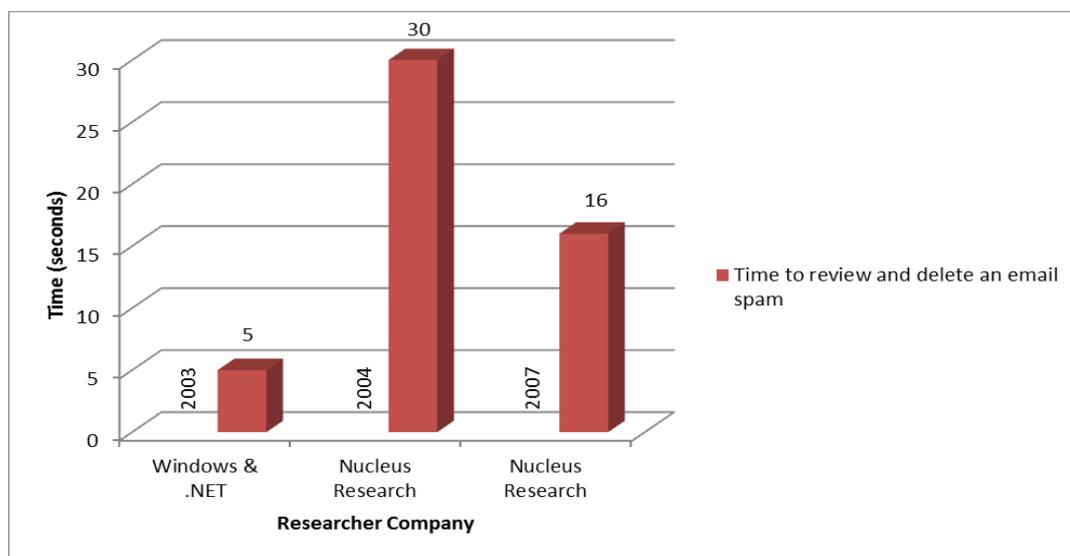


Figure 2.7: Average time taken by an employee to review and delete an email spam

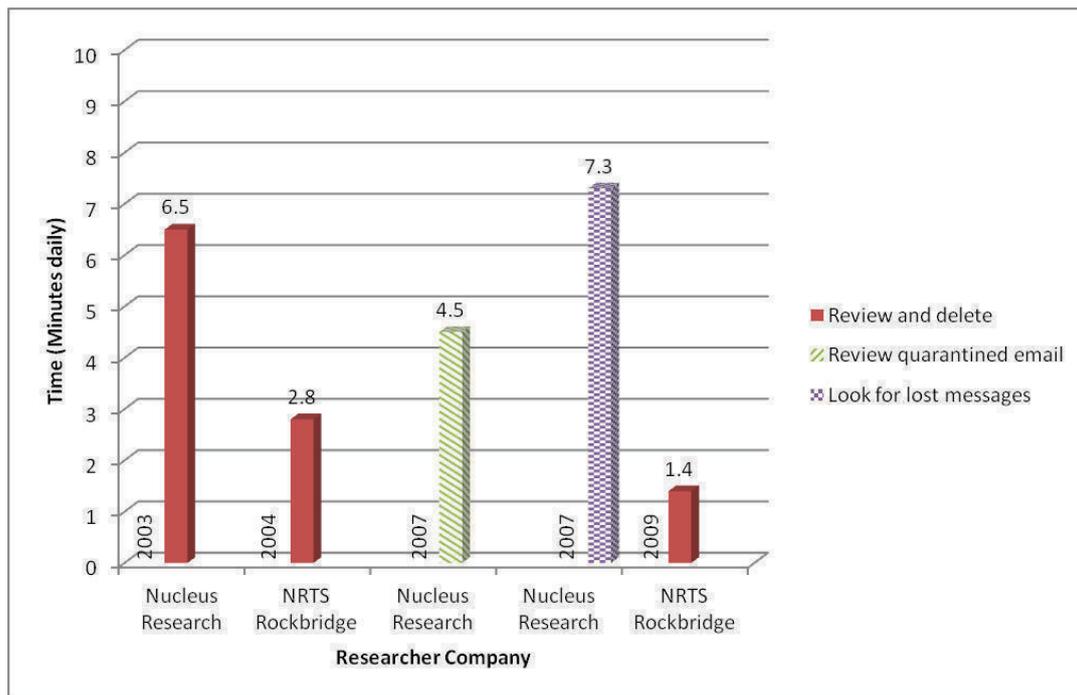


Figure 2.8: Average time taken by an employee to do related activities caused by email spam in a day

2.2.1.3 Impact of email spam

The previous section gave an estimate of the cost of email spam. This section describes the impact of email spam divided into five categories – social, psychological, technical and environmental.

Economic impact

Economically, email spam results in several issues, such as 1) monetary loss, 2) cost of anti-spam filtering products, 3) the need for user training, and 4) productivity loss.

Monetary Loss

Spam gives monetary benefits to the spammers. However, this benefit comes at the cost of monetary loss to individuals, organisations and countries, as stated in Subsection 2.2.1.3. Spam necessitates significant infrastructure investment for software, storage, human resources, network bandwidth and hardware causes the wastage of these resources as well as time. Furthermore, spam causes additional burden on the ISP in handling and filtering it. It can be seen from Figure 2.9 that the cost for spammers to spam is decreasing over time, but the cost for combating spam is increasing over time.

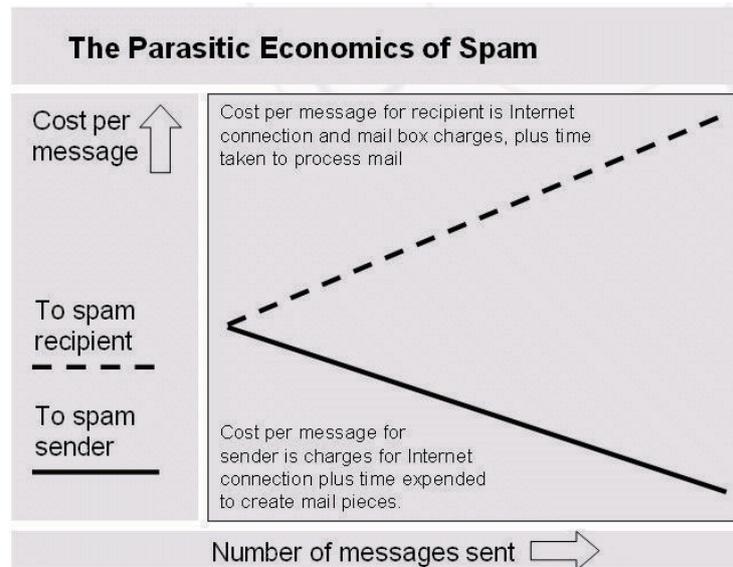


Figure 2.9 : Parasitic economics of spam [Source : (Cobb 2003)]

Cost of anti-spam filtering products

In order to avoid spam attacks, organisations need to acquire latest anti-spam filtering products. However, these products usually do not come for free. Furthermore, the organisation has to invest on extra storage and archive for its employees' email. This extra storage is used to keep the filtered potential spam mail for further review by the employees. Hence, the spam mails are kept until the employees review them and decide to delete them. Storage is also needed for frequent updates for the anti-spam filtering products, in order to enhance the effectiveness of the filter.

The need for user training

Employees need to be trained on using anti-spam products. It is useless to implement and invest in a good anti-spam filtering product if the end users do not know how to use them. Since these anti-spam filtering products keep coming out with new versions, frequent training of the end users becomes necessary.

Productivity loss

Spam causes loss of or decrease in productivity. In the quarantine strategy, the spam messages are filtered and all filtered messages are reviewed which obviously requires considerable time. In contrast, in the delete strategy, the email messages flagged as spam are deleted without a review. As a result, however, the users need to look for accidentally deleted legitimate messages among thousands of other deleted messages, again consuming a lot of time. Nucleus research reported that, in 2007, 4.5 minutes per week was spent to review and delete filtered messages, while users spent even more time, i.e., 7.3 minutes, to look for lost legitimate messages (Nucleus Research 2007). Different strategies result in different values of costs and the report indicates that the delete strategy results in higher costs than the quarantine strategy (Nucleus Research 2007). Obviously, the estimate of loss of productivity

is higher and the impact bigger when a company has more employees. Based on the Nucleus Research reports, on an average, the annual loss of productivity per employee can range from 1.2%-3.1% (Nucleus Research 2003, 2004, 2007).

Social impact

Socially, email spam results in privacy abuse and causes loss of reputation as well as potential loss of business. It is also used to commit fraud, scam and harassment, and spread pornography, while leaving the door open to other security threats.

Privacy abuse

Sending email spam can be considered as an act of privacy abuse because email inbox is a confidential and private property of an account user. This is the reason behind the definition of spam given earlier stating that the recipient's consent is the main factor in determining whether an email message is considered as spam or not. Spam also causes people to be guarded and hesitant to reveal their email addresses in public. This may cause them to avoid using any applications that require email address for account registration.

Loss of reputation

In the long run, spam can make the users suspicious of any application or business that is new and unknown to the community. Thus, it could contribute to creating distrust, endangering the very existence of the online community which is based on trust. Non-spamming companies could suffer from damage to reputation as a result of others' wrongdoing. Even reputed brands could be affected and damaged by email spam if the spammers use their name in their spam campaign. This could degrade the quality of information on the Internet. For example, there is a fake Trojan using Microsoft Security Essential (Abrams 2010). This Trojan gives a fake alert to the user to install it and users do not hesitate to install it thinking that it is from Microsoft. Later, these users begin to receive information that Trojan was detected in their computer. Acting on this information, the users will not be able to remove this infection from the computer so easily. Consequently, later even genuine updates from Microsoft Security will cause users to be more cautious in trusting this product. It is also possible that they subscribe to some other anti-spam product.

Potential loss of business

False positive email messages could also cause legitimate email loss due to filtering. Important and legitimate emails not delivered to the recipient could result in loss of business opportunities.

Fraud and scam

There is also a possibility of fraud and scam campaigns to be escalated through email spam. Email could lead to cases of identity theft. Users could be deceived by receiving an email from a trustworthy-looking entity requesting confidential information and gullible users might fall for the

trick. Symantec, in its February 2012 report, stated that there was an increase of 0.01% in phishing emails since January, and one in 358.1 emails was identified as phishing mail (Symantec Intelligence 2012a).

Harassment and pornographic material

Email spam content could include harassing material (Pfleeger and Bloom 2005) or pornographic and adult content included in the message or as attachment. Sometimes, users are redirected to other webpages from the link embedded in the mail that request for personal information, such as credit card number. Because of this, parents are becoming unduly cautious of letting their children use the Internet and, sometimes, children are not allowed to use computer unsupervised.

Open door to other security threat

On the security note, email spam could mislead users to the web pages with virus, worms, Trojans or malware, or to phishing websites. Such mails target the inboxes of vulnerable users and provide an open door to other cyber security threats. These threats could also be embedded in attachments or drive-by downloads. It is obvious; therefore, that email spam is contributing to the propagation of virus, worms and Trojans on the Internet. A basic user who has little knowledge of spam could easily be deceived by spammers.

Psychological impact

Psychologically, email spam causes annoyance to the users, negatively affects their confidence regarding the quality and credibility of the available information and decreases the employees' concentration.

Annoyance

Spam annoys almost every Internet user especially when the amount of spam received is very large. Such spam could be X-rated, commercial, friendly-looking, or of other types. It is also infuriating to look for email coming from legitimate sources in the spam folder when one has to. Webmasters or application developers usually implement software to prevent spammer's registration on their applications. A common method used for it is the prevention approach, called *Completely Automated Public Turing test to tell Computers and Humans Apart* (CAPTCHA). CAPTCHA is proven to be the most effective and popular method used during registration. Although it works well to keep bots from registering on Web 2.0 applications, often it is also difficult for human users to enter it. Some of the CAPTCHA texts cannot be seen clearly and strain the eyes. It is extremely irritating for the users when they have to redo it over and over again until they get it right. Besides, some questions provided in the CAPTCHA are too difficult, language dependent or requiring some specific knowledge, and not all users can solve these. Hence, the method which is used to prevent spammers has also increased user inconvenience while using the Internet.

User confidence

In the view of the continuing onslaught of spam, many users are likely to be drawn by spam sooner or later. Once they become a victim of a spam campaign, especially those that cause them to lose money, they will have a negative perception towards the services offered on the Internet. They will have hard time trusting even legal and well-known services. This will decrease their level of confidence towards the quality and credibility of the available information on the Internet, especially for the services that require their email addresses.

Employee concentration

Moreover, email spam causes employees to lose their concentration. Spammers craft interesting and ‘too good to be true’ advertisements in their email content in order to gain attention of the readers. Employees who receive such mails will spend at least some time reading them. On the top of that, if they are interested in the products advertised in the mail, they may try to make a comparison with other products sold by other companies. And, if they fall for the spammer’s trick, they will continue spending time on the spammer’s website following the link provided in the mail. Hence, email spam causes employees to lose their concentration and waste time during working hours.

Technical impact

On the technical side, botnets, used as spamming agents by spammers to run their campaign swiftly, requires extra storage and anti-spam filtering software to combat it.

Botnets as spamming agent

Spammers use botnet to send enormous amounts of email spam to the Internet users. A botnets or a zombie network is a collection of compromised computers infected by a malicious program called bots (PC Magazine 2013). This collection of infected computers is used to help spammers in their spamming activities by proliferating the bot malware. These malicious programs allow the spammers to control the infected machines remotely, without the legitimate user’s knowledge. These botnets can also be used to perform spam-related cybercrime, such as identity theft, phishing, click fraud and other security attacks (PC Magazine 2013; MessageLabs Intelligence 2010; Hayati, Chai, et al. 2010; Hayati, Potdar, Chai, et al. 2010). As a result, there will be an increase in the number of compromised machines due to the profits obtained from these cybercrime activities. These botnets could be bought or advanced spammers could try to steal an existing botnet. According to Symantec, 88.2% of email spam messages in 2010 were sent with the help of botnets, when the average spam rate was 89.1% of the emails sent worldwide being spam (MessageLabs Intelligence 2010).

Need for extra storage and anti-spam filtering software

Another technical impact of spam manifests itself in the need for organizations to allocate more storage on their web servers for message filtering, specifically for those implementing flagging or quarantine methods. In order to combat spam, anti-spam filtering software needs to be installed and

updated regularly, which also unnecessarily consumes storage. At a certain point, the extra storage could cause a drastic increase storage cost, particularly for those who rent a hosting package.

Environmental impact

It is interesting to observe that spam contributes even to climate change and Green House Gas (GHG) emissions. An interesting study by McAfee shows that, from the environmental point of view, the average GHG emission associated with one email spam is 0.3 grams of CO₂ as depicted in Figure 2.10. This amount of CO₂ is comparable to driving a vehicle for 3 feet (1 meter). In 2008, McAfee estimated the number of email spams sent worldwide at 62 trillion, which, multiplied by the amount CO₂ per email spam given above, results in a CO₂ emission equivalent to driving around the earth 1.6 million times (McAfee 2009). This environmental impact worsened in 2010 when 95 trillion spam emails were sent, causing an equivalent emission of driving around the earth 2 million times (WebpageFX Team 2011).

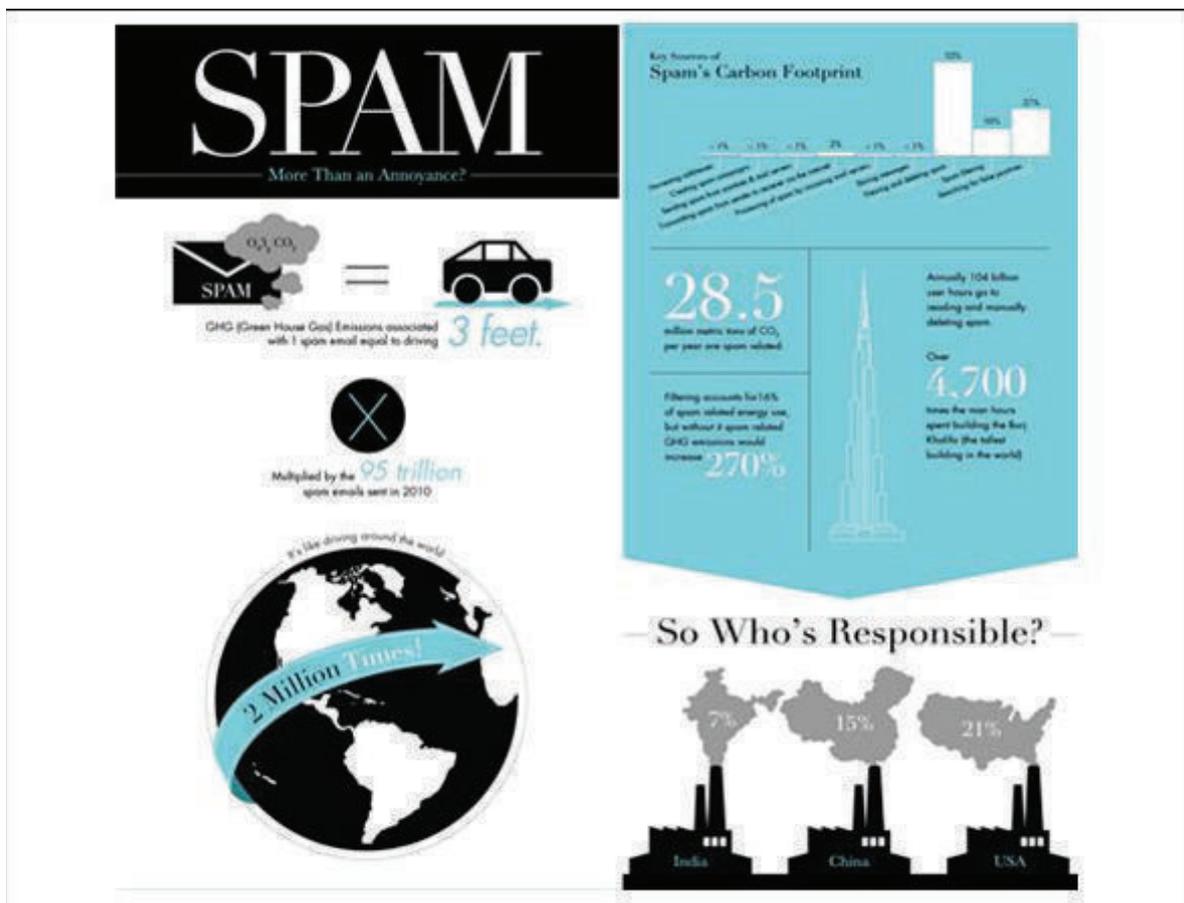


Figure 2.10: Email spam carbon footprint [Source: WebpageFX Team (2011)]

It is also reported that spam campaigns use a total of 33 billion kilowatt-hours(KWh) energy yearly, which is equivalent to the electricity used in 2.4 million homes in the United States (McAfee 2009). It is estimated that 52% of the carbon footprint of spam comes from the activity of viewing and deleting

spam while another 27% originates from the activity of searching for false positives. In addition, spam filtering contributes to 16% of the carbon footprint of spam.

Thus, the massive problems caused by spam are undeniable. Therefore, it is important to identify the costs involved in combating spam and produce email spam cost models, not only to come out with a better solution but also to increase awareness among the Internet community.

2.2.1.4 Email spam cost models methodology

In the previous section, the thesis showed that email spam is a serious global problem. Several researchers have tried to quantify the cost of email spam in order to bring out the full monetary impact of email spam. This section describes the generic methodology used for developing cost models for email spam using two different approaches for data collection – survey and empirical studies. This section describes both in detail.

Email spam cost model development methodology

Figure 2.11 depicts the generic process involved in building a cost model for email spam. It comprises six stages, namely (1) determining cost stakeholders, (2) determining cost category and its cost parameters, (3) data collection, (4) analysis, (5) cost calculation and cost modelling, and (6) cost validation.

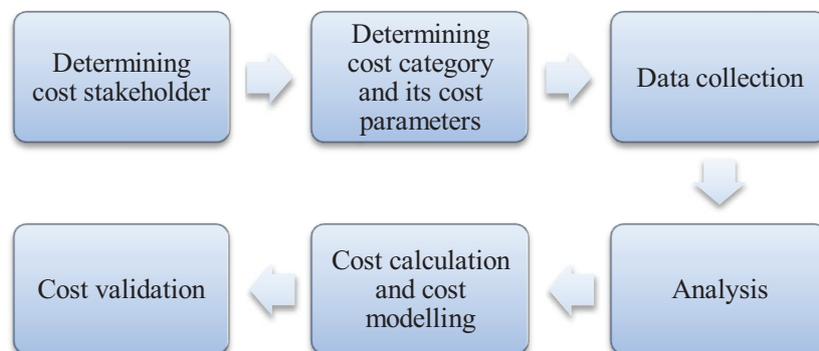


Figure 2.11: Generic cost model development methodology

These stages have been identified on the basis of an extensive review of the existing literature. It is notable that not all these works in the literature have presented their cost model as final; instead, they have just given cost calculations. Yet, they have been taken into account as they could still provide decent input for our cost model.

Stage 1: Determining cost stakeholders – This is the first step, where the main objective is to decide which stakeholders are involved. From the literature review, it was found out that all the stakeholders listed as below are involved:

- Individual (employee/Internet user/spammer)

- Organisation
- Country
- Worldwide

The output of this cost model will be beneficial for the stakeholders in as much as it would obtain a value that could show how far the related stakeholders are affected by the spamming activities. Therefore, in this process, the output of the cost model to be focused on is to be decided from individuals (either users or spammers) or larger parties such as organizations, countries or the worldwide. It is also possible to focus on two stakeholders in a research. Such a research was conducted by Ferris Research (2005, 2009) which generated costs for the US and the worldwide.

Stage 2: Determining cost category and its cost parameters – This is the second step in building a cost model for email spam. In this process, the cost category used for the cost calculation or cost model development is identified. Next, the cost parameters required to calculate the final value of each cost category are identified. It is important to define the related cost parameters that are measurable in terms of time and money. This is the main the reason why the other associated intangibles costs resulting from spamming activities are usually not included in the cost calculation. Hence, the final output could be in either monetary or time value. The common practice is to convert the time into monetary value based on hourly rate.

Stage 3: Data collection – This is the third stage in building a cost model for email spam. There are basically two approaches to obtain data, either to use (1) survey/interview, or (2) through laboratory experiment. If the data are collected through questionnaire/interview, there is a need for a detailed process to construct questions which could produce a value for the cost parameter. If the data are collected empirically, there is a need for a setup process. The types of data can be collected from external or internal sources. The empirical method from internal sources is viable for a company using anti-spam software. However, the researchers who do not have the existing data need to collect them from an external source. Another option is to get and use data from an existing spam repository, e.g. as done by Omar and Samman (2011). Data are collected for a certain duration or until they reach a certain quantity.

Stage 4: Analysis – Once data collection is done, there is a need for data analysis. This is the fourth stage in the methodology and the most critical one in which the data are put to basic statistical analysis to produce the value for each cost parameter. At this stage, all data are used to generate the value for cost parameters specified at the earlier stages. The cost parameter value acts as a requirement and this process produces results for each requirement.

Stage 5: Cost calculation and cost modelling – This is the fifth stage in building the cost model. Here, all the cost parameters defined earlier are gathered and combined using the associated cost functions. The output of this stage is either a cost value or a cost model. The cost calculation and cost modelling is presented in the form of an academic or a commercial report. Nevertheless, there are cost calculation formulations presented in the form of cost calculators. Several studies surveyed depended on the output from the statistics provided by external sources.

Stage 6: Cost validation – Cost validation is the final stage in the methodology. Finally, after the process of creating the cost model, the cost formulation needs to be validated. However, in most cases, the validation method is not revealed to the public. Furthermore, there are only a few studies that have generalized their cost model, and in these cases, they have depended on estimation and statistics provided by the authorities. This stage is required only if the study intends to make a generalization. In most cases, the cost models have been presented only for the particular case study of the research. However, cost validation can also be done by experts in the field or by testing the same model in a case study, e.g. as done by Sonnenreich, Albanese, and Stout (2006). Simulation can be done to prove that the cost model works, e.g. as done by Ilger et al. (2006). Further modifications will be required to achieve the final cost model.

2.2.1.5 Literature review

The previous section described the generic methodology used in building a cost model for email spam using two ways of doing the data collection. This section further establishes the related information and compares the different components in cost models used in the literature. At the end, the section presents a few works of literature that have introduced cost models relevant to this research. Most of these works have proposed new cost models without providing the method for data collection. The section first presents those that have specified their methodology, followed by those that have not.

Literature review on email spam cost model

The thesis now presents the survey studies that have done further work on developing cost models for email spam. Table 2.1 contains a summary of the literature on email spam that specifies the data collection method. The data collection methodology has been categorized based on Galliers' (Galliers 1992) and Minger's (Mingers 2003) work.

Unspecified: literature review on related cost model

Table 2.2 provides a collection of related cost models that have been proposed by researchers. However, they have provided no proofs of the work or examples related to the spam given. In addition, the thesis presents online calculators from several companies to show the parameters used in their spam cost calculations. These online calculators have been included in the literature review in

order to understand the current real world environment and to help in identifying the costs that are taken into account while calculating the total.

Table 2.1: Literature review on email spam cost models

Literature	Method	Stakeholder	Cost category	Cost parameters	Cost model
1	Nucleus Research (2003)	Individuals	Loss of productivity, <i>LoP</i>	<ul style="list-style-type: none"> Average spam messages received per day per employee, <i>ade</i> average time spent per person managing spam, <i>ats</i> total working hours in minutes/day, <i>twhd</i> = 480 total working hours in a year, <i>twhy</i> = 2080 working salary per hour, <i>wsh</i> = \$30 	LoP employee daily = $\frac{ats * ade}{twhd}$Eq. 2.1 LoP employee annually = LoP employee daily * $twhy$ * wsh Eq. 2.2
2	Nucleus Research (2004)	Individuals	Loss of productivity, <i>LoP</i>	<ul style="list-style-type: none"> Total active work email users, <i>taw</i> Number of profiles, <i>k</i> Percentage of US workforce in each profile, <i>pwf</i> Total employees in US in each profile, <i>n</i> Estimated % with active work email accounts, <i>eaw</i> 	In addition to the three cost functions above, the additional cost functions are as follows: $taw = \sum_{i=0}^k pwfi * n * eawi$Eq. 2.3 LoP US annually = $taw * LoP$ employee annuallyEq. 2.4
3	Nucleus Research (2007)	US	Loss of productivity, <i>LoP</i>	<ul style="list-style-type: none"> Total active work email users, <i>taw</i> Number of profiles, <i>k</i> Percentage of US workforce in each profile, <i>pwf</i> Total employees in US in each profile, <i>n</i> Estimated % with active work email accounts, <i>eaw</i> 	In addition to the three cost functions above, the additional cost functions are as follows: $taw = \sum_{i=0}^k pwfi * n * eawi$Eq. 2.3 LoP US annually = $taw * LoP$ employee annuallyEq. 2.4

Literature	Method	Stakeholder	Cost category	Cost parameters	Cost model
4 Rockbridge Associates Inc. (2005)	Telephonic interview (sub sample of 500 adult respondents out of 1000)	US	<ul style="list-style-type: none"> Loss of productivity, LoP Spammer's revenue, SR 	<ul style="list-style-type: none"> Average minutes per week per respondent, $atsw$ Frequency of deleting spam, (days per week), fsw Time spent (minutes) deleting spam in a typical day, fts Estimated time spent for spam in US, ets Number of online adults in US, n = number of adults as per the US census * 77% online users in US = 169 million Average weekly wages, asw Weekly wages per respondent, ws Total respondents, k Amount of spam per year, tsd Average spam messages received per day per respondent, ade Total days in a week, tdw Week in a year, wy Respondents' percentage that received spam, prs 	$atsw = ftsw * fts$ Eq. 2.5 $ets \text{ (minutes)} = atsw * n$ Eq. 2.6 $LoP \text{ in working weeks} = \frac{ets \text{ (week)}}{40}$ Eq. 2.7 $asw = \sum wsw * k$ Eq. 2.8 $LoP \text{ (monetary weekly)} = LoP \text{ in working weeks} * asw$ Eq. 2.9 $LoP \text{ (monetary annually)} = LoP \text{ (monetary weekly)} * 52$ Eq. 2.10 $tsd = ade * tdw * wy$ Eq. 2.11 $pmp = \frac{mp * 100}{prs}$ Eq. 2.12 $SR = pmp * srs * tsd$ Eq. 2.13
5 Rockbridge Associates Inc. (2009)					

	Literature	Method	Stakeholder	Cost category	Cost parameters	Cost model
6	Commtouch (2010)	Laboratory experiment	Organisations		<ul style="list-style-type: none"> Amount of spam detected, tsd Amount of email processed, aep 	Spam percentage, r $r = \frac{tsd}{aep} * 100\%$Eq. 2.14
7	Symantec Intelligence (2012b)					
8	Omar and Samman (2011)	Laboratory experiment (External source – 3 years data from Yale University's Information Technology Services (ITS) department)SS	Organisations	Future loss of productivity	<ul style="list-style-type: none"> Amount of spam detected Amount of email processed Cost of lost productivity per email (low estimation) = \$0.01 Cost of lost productivity per email (high estimation) = \$0.04 	Forecast : 2 year moving average method

	Literature	Method	Stakeholder	Cost category	Cost parameters	Cost model
9	Ferris Research (2005)	Laboratory experiment & Survey	US	<ul style="list-style-type: none"> • Productivity Cost • IT costs • Help desk costs 	<ul style="list-style-type: none"> • Number of business email users • Hourly labour costs • Weeks worked per year • Average number of spam messages received per user • Proportion of users who filter spam manually/on the desktop/on the server • Average spam filter effectiveness and false positive rate • Average effort involved in a help-desk call • Purchase cost of anti-spam product and services • Effort involved in manually sifting spam from legitimate email. 	Not specified
10	Kim et al. (2006)	Survey of 1000 respondents in Seoul	Korea	<ul style="list-style-type: none"> • Inconvenience cost of spam 	<ul style="list-style-type: none"> • Number of spam messages • Use of anti spam program • Type of spam received • Email storage capacity • Email service price 	Conjoint Analysis Method

Table 2.2 Unspecified method literature review on email spam cost models

Literature	Method	Stakeholder	Cost category	Cost parameters	Cost model
11 Judge, Alperovitch, and Yang (2005)	None specified	Spammers	Spammers' profit	<ul style="list-style-type: none"> Spammer Profit, P Probability of getting caught sending spam, p Number of sent messages that are delivered to the intended recipient, N_d where $N_d = N_s * (AS_d) * (AS_b)$ Anti-spam deployment rate, AS_d Anti-spam block rate, AS_b Response rate, R_r Profit per item, P_i Number of sent messages, N_s Cost of acquiring each address amortized over the useful life of that address, C_a Cost to send each message, C_s Cost of punishment, C_p 	<p>Spammer Profit = [(Chance of Getting Away with Crime) * (Number of Delivered Messages) * (Response Rate) * (Profit per Item)] - [(Number of Sent Messages) * (Cost of Address Amortized over Use + Cost to Send)] - [(Risk of Getting Caught) * (Cost of Punishment)]</p> <p>This formula can be further generalized into: $P = [(1-p) * (N_d) * (R_r) * (P_i)] - [(N_s) * (C_a + C_s)] - [(p) * (C_p)]$Eq. 2.15</p>
12 Sonnenreich, Albanese, and Stout (2006)	Suggestion Survey	Organisations	Return On Security Investment, $ROSI$	<ul style="list-style-type: none"> Risk exposure, RE Risk mitigated, RM Solution cost, SC Annual loss exposure, ALE Single loss exposure, SLE Annual rate of occurrence, ARO 	<p>$RE = ALE = SLE * ARO$Eq. 2.16</p> <p>$ROSI = \frac{(RE * RM) - SC}{SC}$Eq. 2.17</p>

	Literature	Method	Stakeholder	Cost category	Cost parameters	Cost model
13	Ilger et al. (2006)	None specified	Spammers	<ul style="list-style-type: none"> • Hardware cost, H • Software cost, S • Operating Cost, O • Labour Cost, L 	<ul style="list-style-type: none"> • Cost for a computer, C • Cost for a monitor, M • Cost for peripheral devices, P • Cost for operating system, OS • Cost for remailers, R • Cost for mail address harvesters, MAH • Cost for web hosting, WH • Sum of internet service cost, I • Electricity cost for running the system, E • Address collection cost, A • Open proxy cost, OP • Cost of installation, IN • Cost for maintenance, MT • Cost for mail production, MP • Cost for acquiring customers, AC 	Total cost $c = H + S + O + L$ where $H = C + M + P$ $S = OS + R + MAH + WH$ $O = I + E + A + OP$ $L = IN + MT + MP + AC$Eq. 2.18
14	Kshetri (2006)	None specified	Cybercrimes (in our case, spammers)	Spammer's probability to commit crime	<ul style="list-style-type: none"> • Monetary benefits of committing the crime; Mb • Psychological benefits of committing the crime; Pb • Monetary opportunity costs of conviction; Ocm • Psychological costs of committing the crime; Ocp • Probability of arrest; Pa • Probability of conviction. Pc • The expected penalty effect, $OcmPaPc$ 	A cybercrime occurs if $Mb + Pb > Ocp + OcmPaPc$.Eq. 2.19

	Literature	Method	Stakeholder	Cost category	Cost parameters	Cost model
15	Online Calculator 1 (SecureMX Mail Scrubbing)	None specified	<ul style="list-style-type: none"> • Individuals • Organisations 	<ul style="list-style-type: none"> • Productivity cost • Productivity cost (to reply to spam) 	<ul style="list-style-type: none"> • No of employees using email • Average employee salary • Emails per employee per day 	Not specified
16	Online Calculator 2 (SpamEater.com)	None specified	<ul style="list-style-type: none"> • Individuals • Organisations 	<ul style="list-style-type: none"> • Wasted time • Lost salary 	<ul style="list-style-type: none"> • No. of people receiving email • Average annual salary • Average no. of total emails per day per person • Percentage of emails which are spam • Average no. of seconds wasted on each spam email 	Not specified
17	Online Calculator 3 (Alt-N Technologies)	None specified	<ul style="list-style-type: none"> • Individuals • Organisations 	<ul style="list-style-type: none"> • Storage cost • Loss of productivity • Return on investment 	<ul style="list-style-type: none"> • No. of employees • Average annual salary • Average daily email per recipient • Average % of spam from total email • Time to delete = 5 seconds • Cost per recipient per year • Time wasted per response = 5 mins • Response rate = 1% • Cost of 1MB archive storage = USD0.60 • Average size of spam message = 16Kb; • Cost of anti spam solution • Annual interest rate 	Not specified

	Literature	Method	Stakeholder	Cost category	Cost parameters	Cost model
18	Online Calculator 4 (iPermitMail)	None specified	<ul style="list-style-type: none"> Organisations 	<ul style="list-style-type: none"> Loss productivity Labour Cost 	<ul style="list-style-type: none"> No. of employees Average spam received per day Average time to remove each spam message (seconds) Average annual salary per employee per year Annual system administrator salary per year 	Not specified
19	Online Calculator 5 (NetworkWorld.com)	None specified	<ul style="list-style-type: none"> Organisations 	<ul style="list-style-type: none"> Productivity Cost Connectivity Cost Storage Cost Support Cost 	<ul style="list-style-type: none"> No. of employees Average fully loaded salary per employee Average number of working days per person per year Average number of received messages per user per day Average percentage of those messages that are spam Average time to handle a spam Cost of office Internet connectivity per month Cost of remote connectivity per month Bandwidth used by e-mail Average message size Storage cost per month per GB inc. management Average cost of user support per user per year Percentage of support time for spam 	Not specified
20	Online Calculator 6 (Spamfighter)	None specified	<ul style="list-style-type: none"> Organisations 	<ul style="list-style-type: none"> Loss of productivity 	<ul style="list-style-type: none"> No. of mailboxes in the company Hourly salary for each employee Average amount of spam every day for each mailbox Time taken to delete each spam (seconds) 	Not specified

	Literature	Method	Stakeholder	Cost category	Cost parameters	Cost model
21	Online Calculator 7 (Modest Software)	None specified	<ul style="list-style-type: none"> Organisations 	<ul style="list-style-type: none"> Loss of productivity 	<ul style="list-style-type: none"> No. of affected employee mailboxes in the company Average (yearly) salary per employee OR <ul style="list-style-type: none"> Average hourly wages per employee Typical no. of spam email messages received per employee mailbox per day Average time, in seconds, spent resolving a spam email messages. 	Not specified
22	Online Calculator 8 (VicomSoft Ltd)	None specified	<ul style="list-style-type: none"> Organisations 	<ul style="list-style-type: none"> Return on investment 	<ul style="list-style-type: none"> No. of Employees Average cost per hour per employee Average no. of hours an employee spends filtering spam email per day 	Not specified
23	Online Calculator 9 (Computer Mail Services)	None specified	<ul style="list-style-type: none"> Individuals Organisations 	<ul style="list-style-type: none"> Lost salary Loss of productivity 	<ul style="list-style-type: none"> No. of Employees with Email No. of workdays per year per employee Average hourly salary per employee Average no. of spam emails per day per employee No. of seconds wasted with each spam message 	Not specified

2.2.1.6 Evaluation

Section 0 provided a comprehensive review of the literature on the cost models of email spam. This section presents a critical evaluation of the imperative information gathered from this literature. This chapter covers four key elements for cost model studies, which are:

1. Evaluation of data collection method
2. Evaluation of stakeholders
3. Evaluation of cost categories
4. Evaluation of cost parameters

Evaluation of data collection method

In selecting the data collection method, it is important to consider the cost model targeted to be built by the researchers. Several cost models have been developed based on survey and interview instruments; hence their values are usually taken as an average user's opinion. The benefit of this value is to give insight into the opinions of related respondents; however, this value is based on their estimation.

On the contrary, the data collected using laboratory experiment are more natural. The obvious disadvantage of collecting through this method is that it usually needs a setup process. Yet, this method works successfully for organizations that run anti-spam filtering software business because they would originally have done it earlier. The disadvantage of this method is that the value they generate is calculated based on their spam filter's result, and thus cannot be generalized.

On a positive note, there is no certain best method to collect data. It depends more on the researchers' decision and the cost models they intend to develop. Regardless of the data collection method, however, the dataset will provide a comprehensive basis to decide how much confidence can be placed on the results. A model based on a very small scale study cannot be generalized because the results will not represent the whole population adequately.

Evaluation of stakeholders

It can be seen that the stakeholders included in the previous literature involve similar parties. These stakeholders can be categorised into two, spam and anti-spam. The frequency of the stakeholders as the focus of the research is summarised in Table 2.3. Most of the research done so far focuses on the anti-spam stakeholders. The common practice shows that, if the researchers focus on spammers, the cost models takes only spammers as the stakeholders, except in case of Rockbridge Associates Inc. (2005, 2009). However, if the researchers focus on the anti-spam side, they could produce the cost calculation or cost model for smaller entities, such as employees. Additionally, they could generalise

the cost model to be applicable to bigger entities, such as organisations, countries, or the worldwide. Based on Table 2.3, it is obvious that most of the researches are focused on organisations as their stakeholders, followed by individuals. This is because the basic parameters required for cost calculation are taken on per employee or per individual basis. Furthermore, it is also noticeable that all online calculators focus on organisations as their stakeholders, as the data can only be provided by a particular organisation.

Nonetheless, it is also interesting to note that there are several works that focus on spammers. Rockbridge Associates' estimate of the spammers' revenue costs was generated on the basis of their successful campaigns among users and did not include the cost involved for spammers (Rockbridge Associates Inc. 2005, 2009). On the other hand, another two researchers have included the cost for spammers in their work (Judge, Alperovitch, and Yang 2005; Ilger et al. 2006). The present work mainly proposes to change the economic model of the spammers in an attempt to make the revenue of the spammers less than their cost.

Table 2.3 Frequency analysis of stakeholders

No.	Literature	Stakeholders													
		Spam					Anti-spam								
		Spammers' Revenue	Spammers' cost	Other spammers' related research	Individual	Organisation	Country	Worldwide	Spammers' Revenue	Spammers' cost	Other spammers' related research	Individual	Organisation	Country	Worldwide
1	Nucleus Research (2003)				√										
2	Nucleus Research (2004)				√										
3	Nucleus Research (2007)				√				√						
4	Rockbridge Associates Inc. (2005)	√			√						√				
5	Rockbridge Associates Inc. (2009)	√			√						√				
6	Commtouch (2010)									√					
7	Symantec Intelligence (2012b)									√					
8	Omar and Samman (2011)									√					
9	Ferris Research (2005)													√	
10	Kim et al. (2006)													√	
11	Judge, Alperovitch, and Yang (2005)	√													
12	Sonnenreich, Albanese, and Stout (2006)										√				
13	Ilger et al. (2006)									√					

No.	Literature	Stakeholders										
		Spam			Anti-spam			Other spammers' related research	Individual	Organisation	Country	Worldwide
		Spammers' Revenue	Spammers' cost									
14	Kshetri (2006)				√							
15	Online Calculator 1 (SecureMX Mail Scrubbing)						√		√			
16	Online Calculator 2 (SpamEater.com)							√	√			
17	Online Calculator 3 (Alt-N Technologies)							√	√			
18	Online Calculator 4 (iPermitMail)								√			
19	Online Calculator 5 (NetworkWorld.com)								√			
20	Online Calculator 6 (Spamfighter)								√			
21	Online Calculator 7 (Modest Software)								√			
22	Online Calculator 8 (VicomSoft Ltd)								√			
23	Online Calculator 9 (Computer Mail Services)							√				
Total Frequency		3	2		1		9	13	5	1		

Evaluation of cost categories

In the summary given in Table 2.4, it is interesting to observe that 14 out of 23 works in the literature focus on loss of productivity cost. It is probably because this is the cost that usually takes the highest value among all categories. A look at the statistics, where almost all online calculators in Table 2.0 include this cost in their calculators, indicates that it could also be one of the easiest costs that could be calculated. These costs can be calculated for an individual or can be generalised to be applicable to larger stakeholders, using statistics from external sources.

In all, six out of 23 works in the literature are focused on the IT costs that include storage, hardware, software and connectivity costs. These costs can be calculated for all stakeholders. Apparently, these costs require more inside knowledge of the organisations, and hence it is not easy to estimate the value of parameters related to this cost.

Remarkably, there are three out of 15 studies that focus on the return on investment costs and spammers' profits. The thesis will focus on the differences among these studies in the next subsection.

A total of two out of 15 studies focus on help-desk or support cost. Other works provide unique independent cost studies that introduce and calculate new types of costs.

There is also some literature that has not been included in the summary since its output does not qualify to be categorised under these cost categories. Instead of just using the output, higher levels of calculations are required for these, before they could be categorised under one of these cost categories.

Table 2.4: Frequency analysis for cost categories

	Literature	Loss of productivity	IT cost (storage, hardware, software, connectivity)	Help desk/ Support cost	Inconvenience cost	Spammers' Profit	Return on Investment
1	Nucleus Research (2003)	√					
2	Nucleus Research (2004)	√					
3	Nucleus Research (2007)	√					
4	Rockbridge Associates Inc. (2005)	√				√	
5	Rockbridge Associates Inc. (2009)	√				√	
6	Commtouch (2010)						
7	Symantec Intelligence (2012b)						
8	Omar and Samman (2011)	√					
9	Ferris Research (2005)	√	√	√			
10	Kim et al. (2006)				√		
11	Judge, Alperovitch, and Yang (2005)					√	
12	Sonnenreich, Albanese, and Stout (2006)						√
13	Ilger et al. (2006)		√				
14	Kshetri (2006)						
15	Online Calculator 1 (SecureMX Mail Scrubbing)	√	√				
16	Online Calculator 2 (SpamEater.com)						
17	Online Calculator 3 (Alt-N Technologies)	√	√				√
18	Online Calculator 4 (iPermitMail)	√	√				
19	Online Calculator 5 (NetworkWorld.com)	√	√	√			

Evaluation of cost parameters

The cost parameters have been evaluated according to the related cost category. This subsection presents the differences and similarities between all the cost parameters listed from different cost models.

Evaluation of cost parameters for loss of productivity

According to Ferris Research, loss of productivity cost is *the amount of lost time while users deal with spam and filter spam* (Ferris Research 2005). All the parameters found in the literature are listed below:

- i. Average spam messages received per day per employee, *ade*
- ii. Average time spent per person in managing spam, *ats*
- iii. Total working hours in minutes per day, *twhd* = 480
- iv. Total working hours in a year, *twhy* = 2080
- v. Working salary per hour, *wsh* = \$30
- vi. Total active work email users, *taw*
- vii. Number of profiles, *k*
- viii. Percentage of US workforce in each profile, *pwf*
- ix. Total employees in US in each profile, *n*
- x. Estimated % with active work email accounts, *eaw*
- xi. Average minutes per week per respondent, *atsw*
- xii. Frequency of deleting spam, (days per week). *ftsw*
- xiii. Time spent (minutes) deleting spam in a typical day, *fts*
- xiv. Estimated time spent for spam in US, *et s*
- xv. Number of online adults in US, *n* = number of adults as per the US census * 77% online users in US = 169 million
- xvi. Average weekly wages, *asw*
- xvii. Weekly wages per respondent, *wsw*

- xviii. Total respondents, k
- xix. Amount of spam detected, tsd
- xx. Number of emails processed, aep
- xxi. Cost of lost productivity per email (low estimation) = \$0.01
- xxii. Cost of lost productivity per email (high estimation) = \$0.04
- xxiii. Number of business email users
- xxiv. Hourly labour costs
- xxv. Weeks worked per year
- xxvi. Average number of spam messages received per user
- xxvii. Effort in manually sifting spam from legitimate email.
- xxviii. Number of employees using email
- xxix. Average employee salary
- xxx. Emails per employee per day
- xxxi. Number of employees
- xxxii. Average annual salary
- xxxiii. Average daily emails per recipient
- xxxiv. Average % of spam in total emails
- xxxv. Time to delete = 5 seconds
- xxxvi. Time wasted per response = 5 mins
- xxxvii. Response rate = 1%
- xxxviii. Cost per recipient per year
- xxxix. Number of employees
- xl. Average spam received per day
- xli. Average time to remove each spam message (seconds)

- xlii. Average annual salary per employee per year
- xliii. Number of employees
- xliv. Average fully loaded salary per employee
- xlv. Average number of working days per person per year
- xlvi. Average number of received messages per user per day
- xlvii. Average percentage of those messages that are spam
- xlviii. Average time to handle a spam
- xlix. Number of mailboxes in the company
 - l. Hourly salary for each employee
 - li. Average amount of spam every day for each mailbox
 - lii. Time taken to delete each spam (seconds)
 - liii. Number of affected employee mailboxes in the company
 - liv. Average (yearly) salary per employee

OR

Average hourly wage per employee

- lv. Typical number of spam email messages received per employee mailbox per day
- lvi. Average time, in seconds, spent resolving a spam email message.
- lvii. Number of employees with email
- lviii. Number of workdays per year per employee
- lix. Average hourly salary per employee
- lx. Average number of spam emails per day per employee
- lxi. Number of seconds wasted with each spam message

As can be seen, there is a lot of repetition of similar parameters. Hence, these repetitions have been eliminated and all similar parameters have been listed next to them as below:

i. Number of affected employee mailboxes in the company = Number of employees with email
= Number of mailboxes in the company = Number of employees = Number of business email
users = Total respondent

ii. Cost per recipient per year

OR

Average hourly wages per employee = Average hourly salary per employee = (Hourly salary
for each employee working salary per hour = \$30)

OR

Average weekly wages = weekly wages per respondent * total respondent

OR

Average (yearly)/annual salary per employee = Average hourly salary per employee *
Number of workdays per year per employee = Average fully loaded salary per employee *
Average number of working days per person per year = Hourly labour costs * Weeks worked
per year

With total working hours in minutes/day = 480, total working hours in a year = 2080

iii. Average time, in seconds, spent resolving a spam email message = Number of seconds wasted
with each spam message = Time taken to delete each spam (seconds) = Average time to
remove each spam message (seconds) = (Time to delete = 5 seconds)

OR

Average time to handle a spam = effort in manually sifting spam from legitimate email =
average time spent per person managing spam, *ats*

OR

Average minutes per week per respondent = Frequency of deleting spam (days per week). *
Time spent (minutes) deleting spam in a typical day

Time wasted per response = 5 mins with Response rate = 1%

- iv. Typical number of spam email messages received per employee mailbox per day = Average number of spam emails per day per employee = Average amount of spam every day for each mailbox = Average spam received per day = Average spam messages received per day per employee = Average number of spam messages received per user

OR

Average number of received messages per user per day * Average percentage of those messages that are spam = Average daily email per recipient * Average % of spam from total email

- v. Number of online adults in US = 169 million
- vi. Total active work email users
- vii. Number of profiles
- viii. Percentage of US workforce in each profile
- ix. Total employees in US in each profile
- x. Estimated % with active work email accounts
- xi. Average spam rate = Amount of spam detected / Amount of email processed * 100
- xii. Cost of lost productivity per email (low estimation) = \$0.01
- xiii. Cost of lost productivity per email (high estimation) = \$0.04

In order to calculate the loss of productivity cost, four basic parameters have been identified as listed from i-ix above. The first parameter is the number of affected employee mailboxes in the company (Modest Software). Similar parameters can be found from other researches, such as number of employees with email (Computer Mail Services; SecureMX Mail Scrubbing), number of mailboxes in the company (Spamfighter), number of employees (iPermitMail; Alt-N Technologies; NetworkWorld.com) and number of business email users (Ferris Research 2005).

Second parameter needed to calculate the loss of productivity cost is the cost per recipient per year (Alt-N Technologies). This cost parameter can be identified by defining employees' salary using a different time unit, whether hourly, weekly or annually. Similar parameters for calculating the second parameter are average employee salary (SecureMX Mail Scrubbing), average hourly wages per employee (Modest Software), average hourly salary per employee (Computer Mail Services), hourly

salary for each employee (Spamfighter), working salary per hour (Rockbridge Associates Inc. 2005, 2009), average weekly wages (Rockbridge Associates Inc. 2005, 2009), average yearly/annual salary per employee (Modest Software; iPermitMail; Alt-N Technologies), average fully loaded salary per employee (NetworkWorld.com) and hourly labour costs (Ferris Research 2005). Further multiplication of the parameters with the number of working days or number of working weeks might be needed to convert it to a higher level of time unit (e.g., week and year). Similar parameters to define this parameter are number of workdays per year per employee (Computer Mail Services), average number of working days per person per year (NetworkWorld.com) and weeks worked per year (Ferris Research 2005).

The third cost parameter needed is the average amount of time used to handle a spam message (NetworkWorld.com) or average time spent per person managing spam (Nucleus Research 2003, 2004, 2007). This also includes the effort in manually sifting spam from legitimate emails as defined in Ferris Research (2005). Similar to the second parameter, various time units such as seconds and minutes have been used. Common practice shows that most of the studies have used second, except in case of Rockbridge Associates Inc. (2005, 2009). Similar cost parameters used to define the amount of time used to handle a spam message in seconds are average time (in seconds) spent resolving a spam email message (Modest Software), number of seconds wasted with each spam message (Computer Mail Services), time taken to delete each spam (seconds) (Alt-N Technologies; Spamfighter) and average time to remove each spam message (seconds) (iPermitMail). The cost parameters that have used “minute” as the time unit are average minutes per week per respondent calculated by multiplying the frequency of deleting spam (days per week) and time spent (minutes) deleting spam in a typical day (Rockbridge Associates Inc. 2005, 2009). In addition, Online Calculator 3 also defined this cost by calculating the cost of responding to spam email using cost parameters such as time wasted per response and response rate (Alt-N Technologies).

The fourth parameter is the typical number of spam messages received per employee mailbox per day (Modest Software) or similarly defined as the average number of spam emails per day per employee (Computer Mail Services), average amount of spam every day for each mailbox (Spamfighter), average amount spam received per day (iPermitMail), average number of spam messages received per day per employee (Nucleus Research 2003, 2004, 2007) and the average number of spam messages received per user (Ferris Research 2005). Another way of defining this parameter is to multiply an average percentage of spam from the total number of emails with the average daily emails per recipient. This method is also used in using cost parameters such as the average number of received messages per user per day and the average percentage of the spam messages (NetworkWorld.com).

The first four cost parameters are basic cost parameters required to calculate the loss of productivity cost at the level of the company. However, in order to generalise the cost to the level of the country,

more parameters are required such as the number of online adults in the country (Rockbridge Associates Inc. 2005, 2009), total number of active work email users, number of profiles, percentage of the country's workforce in each profile, total number of employees in the country in each profile, and estimated percentage of active work email accounts (Nucleus Research 2003, 2004, 2007). A comparison of the researches of Rockbridge and Nucleus Research shows a big difference in their cost parameter definition. The cost parameters defined by Nucleus Research are fit to calculate the cost based on the industry profile while it also fits the data collection for the employees at the companies (Nucleus Research 2003, 2004, 2007). On the other hand, Rockbridge Associates' cost parameters are more general as they collect their data from public respondents and use the number of online adults as their cost parameters (Rockbridge Associates Inc. 2005, 2009). While the Nucleus Research cost model seems to be more focused on loss of productivity of US employees (Nucleus Research 2003, 2004, 2007), Rockbridge Associates' cost model not only calculates loss of productivity by the general users but also manages to calculate the spammers' revenue (Rockbridge Associates Inc. 2005, 2009).

The cost parameters discussed so far have focused on calculating the current or past loss of productivity cost. However, the research by Omar and Samman focused on calculating and forecasting the future loss of productivity cost based on the past data (Omar and Samman 2011). The cost parameters used in this research are amount of spam detected, number of emails processed and cost of lost productivity. However, they have calculated the range of this cost using the low estimation cost per email at \$0.01 and high estimation cost per email at \$0.04. These fixed values are obtained from the output of other researches.

Loss of productivity is usually vaguely defined. This cost category could have mistakenly be identified as the time taken to identify and delete spam only (such as in Online Calculator 3 and 6 (Alt-N Technologies; Spamfighter)). However, loss of productivity cost should include the time taken to cater to these spam messages too. Hence, in this research, the loss of productivity cost is defined as the amount of time lost due to any activity related to dealing with email spam, including not just the amount of time used to handle/manage spam, but also to read, identify, decide, delete/remove, report, flag, or even to handle false positive email.

Although different units are used in defining the cost parameters, most of them refer to similar definitions. Hence, the only difference is the focus of the cost, either to be used by that particular company or to generalize it for a country. The higher the levels of stakeholders to estimate the cost, the more the parameters required in the formulation. However, it also means that the more average values used while focusing on a certain company, the more detailed and accurate the cost value will be.

It is also noticeable that some studies are using predefined values, such as hourly salary. For each employee, working salary per hour is stated as \$30 (Rockbridge Associates Inc. 2005, 2009). Some of the predefined values use input from other researches, such as the number of online adults in the US at 169 million (Rockbridge Associates Inc. 2005, 2009), or the low and high estimation cost per email at \$0.01 and 0.04 respectively (Omar and Samman 2011).

Evaluation of cost parameters for spammer's profit

There are only three works relating to spammer's profits. Even among these, two are for the same company (Rockbridge Associates Inc. 2005, 2009) using similar parameters, and hence only one of these and the third need to be considered. In all, there are 20 parameters defined for calculating spammer's profits in the researches by Rockbridge Associates and Judge, Alperovitch and Yang as listed below:

- i. Spammer profit
- ii. Probability of getting caught sending spam
- iii. Number of sent messages that are delivered to the intended recipient
- iv. Anti-spam deployment rate
- v. Anti-spam block rate
- vi. Response rate
- vii. Profit per item
- viii. Number of sent messages
- ix. Cost of acquiring each address amortized over the useful life of that address
- x. Cost to send each message
- xi. Cost of punishment
- xii. Spammer's revenue
- xiii. Amount of spam per year
- xiv. Average spam messages received per day per respondent
- xv. Total days in a week
- xvi. Week in a year

- xvii. Respondent's percentage that received spam
- xviii. Respondent's that made a purchase from spamming activities
- xix. Respondent's percentage that made a purchase from spamming activities
- xx. Spammer's revenue for 1 spam = 1 cent

Omitting the parameters that refer to similar values other parameters are as listed below:

- i. Spammer profit = Spammer's revenue
- ii. Number of sent messages that are delivered to the intended recipient = Average spam messages received per day per respondent
- iii. Response rate = Respondent's percentage that made a purchase from spamming activities
- iv. Profit per item = spammer's revenue for 1 spam
- v. Number of sent messages = Amount of spam per year

Both of these cost models have been generated from different perspectives. Rockbridge Associates has calculated the actual profit made by the spammers taking into account the actual purchases made by the respondents (Rockbridge Associates Inc. 2009, 2005). On the other hand, Judge, Alperovitch and Yang have estimated this value by considering the probability of not getting caught (Judge, Alperovitch, and Yang 2005).

Nonetheless, Rockbridge Associates has used generalization in its dataset whereby it calculated the percentage of respondents who made purchases as a result of spamming activities (Rockbridge Associates Inc. 2009, 2005). Furthermore, pre-defined values, 1 spam equal to 1 cent, have been used for spammers' revenues.

Although there were no predefined values in the research by Judge, Alperovitch and Yang, their method to obtain these cost parameters is unclear. It is also unclear how they have generated the value of each parameter. In contrast, Rockbridge Associates has provided a clear method for obtaining the data, i.e., by using telephonic interviews (Rockbridge Associates Inc. 2009). However, Judge, Alperovitch and Yang's research provides an insight to further consider other parameters, such as cost to send each message, and cost of punishment (Judge, Alperovitch, and Yang 2005). In fact, it could be seen that Rockbridge Associates' way of estimating the spammers' profit costs actually consists of the first part of the cost model used by Judge, Alperovitch and Yang.

Evaluation of cost parameters for IT costs

For simplicity, the IT cost is defined as a cost category that includes the labour cost, hardware cost, software cost, operating cost, storage cost and connectivity cost. Further detailed evaluation of cost parameters will be done accordingly.

Labour cost

Below is the list of parameters for the labour cost:

- i. Hourly labour cost
- ii. Weeks worked per year
- iii. Cost of installation
- iv. Cost of maintenance
- v. Cost of mail production
- vi. Cost of acquiring customers
- vii. System administrator's salary per year

Hourly labour cost and weeks worked per year are the parameters needed to calculate the labour cost taken from Ferris Research (Ferris Research 2005). Cost of installation, cost of maintenance, mail production and acquiring customers are parameters defined in calculating the labour cost for spammers (Ilger et al. 2006). System administrator's salary per year was the parameter taken from Online Calculator 4 (iPermitMail).

This type of cost can be calculated for both spam and anti-spam. For the spammers, any labour work including installation, maintenance, mail production, and acquiring customers can be totalled up for the labour cost. While for the anti-spam side, any management and system administration work can be calculated to estimate it. Expenses on employees specifically given the task of dealing with spam can also be calculated in the labour cost. However, the labour cost could overlap with the loss of productivity and support cost.

Hardware cost

Listed below are the parameters from a research on cost for spammers (Ilger et al. 2006).

- i. Cost of a computer
- ii. Cost of a monitor
- iii. Cost of peripheral devices

These costs are considered essential for the spammers to run their spamming activities. However, they are usually not included when calculating the cost for anti-spammers.

Software cost

Below is the list of parameters for the software cost category:

- i. Purchase cost of anti-spam product and services.
- ii. Average spam filter effectiveness and false positive rate.
- iii. Cost of the operating system
- iv. Cost of the remailers
- v. Cost of the mail address harvesters
- vi. Cost of web hosting

The first two parameters from the Online Calculator have been used to calculate the email cost for organizations (NetworkWorld.com). The next four parameters have been taken from a research on the total cost for spammers (Ilger et al. 2006). Ilger et al. have identified two types of software requirement, either basic software which every computer needs, or software for special activities like spamming. These parameters represent different stakeholders. Hence, they are also not comparable. However, it is obvious that any software cost should be categorized under this category.

Operating cost

Operating cost is considered as any cost involved in running the task. The parameters for operating cost are listed below:

- i. Sum of internet service cost
- ii. Electricity cost for running the system
- iii. Address collection cost (bought or self-collected)
- iv. Open proxy cost

The thesis follows the exact cost category stated by the research (Ilger et al. 2006); however there might be an overlap with other costs, such as labour cost or connectivity cost. In this case, the cost of address collection could be referred to as manual labour work, and hence categorized under labour cost. In addition, the sum of Internet service costs could also be categorized under a specific category called connectivity cost.

Storage cost

The parameters for storage cost are listed below:

- i. Average size of a spam message
- ii. Cost of 1MB archive storage
- iii. Storage cost per month per GB including management

In order to calculate the storage cost, researchers need to gather basic information, such as the number of users with mailbox, average spam messages received per day, average size of a spam message, and the cost of 1MB archive storage (Alt-N Technologies). These average values are usually calculated on the basis of a survey, or a random value is taken on the basis of individual experience. In this case, the cost of 1MB archive storage is valued at USD0.60 and average size of a spam message is taken as 16kB.

On the other hand, the third parameter found in the literature considers storage cost per month per GB which already includes the management cost (NetworkWorld.com). Nowadays, there are plenty of storage capacity packages that include management of the storage on the market. However, if handled separately, management of the storage could also go under labour cost.

Connectivity cost

Listed below are the parameters for connectivity cost (NetworkWorld.com):

- i. Cost of office Internet connectivity per month
- ii. Cost of remote connectivity per month
- iii. Percentage of bandwidth used by an email

The main idea behind this is to calculate the cost of connectivity specifically used for email activities. Compared to the parameter sum of internet service cost defined earlier in Operating Cost (Ilger et al. 2006), these three parameters listed above managed to define a more accurate information needed to calculate connectivity cost.

Evaluation on cost parameters for help desk/support

Help-desk cost as defined by Ferris Research is “users call center help desk from time to time, to seek help dealing with spam issue which covers the cost of providing the help-desk service and the cost of user’s time using the service” (Ferris Research 2005). Listed below are the parameters for help-desk or support cost from several researchers:

- i. Average effort involved in a help-desk call

- ii. Average cost of user support per user per year
- iii. Percentage of support time for spam

Ferris Research defined the first parameter to calculate help-desk costs (Ferris Research 2005). The rest of the parameters were taken from Online Calculator 5 It is common to define parameters to calculate the cost of help-desk or support based on an average value. For the second parameter where the average cost of user support per user per year is defined, there is a need to obtain more information about the number of employees. However, for the first parameter, Ferris Research included the cost of users' time using the service in this cost while they also calculated the cost of loss of productivity (Ferris Research 2005). In our opinion, the cost of users' time using the service is also included as the time used in managing spam; thus, it should have been included in the loss of productivity cost. Nonetheless, each individual could have given different definitions over similar cost categories.

Evaluation on cost parameters for return on investment

In order to evaluate the cost parameters for ROI, the thesis lists all the parameters found from research below:

- i. Risk exposure
- ii. Risk mitigated
- iii. Solution cost
- iv. Annual loss exposure
- v. Single loss exposure
- vi. Annual rate of occurrence
- vii. Total cost of spam
- viii. Cost of anti-spam solution
- ix. Annual interest rate
- x. Number of employees
- xi. Average cost per hour per employee
- xii. Average number of hours an employee spends filtering spam emails per day

Risk exposure, risk mitigated, solution cost, annual loss exposure, single loss exposure and annual rate of occurrence are parameters taken from the researches to calculate the cost of return on security investment (Sonnenreich, Albanese, and Stout 2006). On the other hand, the total cost of spam, cost of anti-spam solution and annual interest rate are the parameters defined in Online Calculator 3 (Alt-N Technologies). The other parameters were taken from Online Calculator 8 (VicomSoft Ltd).

ROI cost for both calculators were calculated in days. However, for Online Calculator 8, the cost of solution was taken based on their company's solution. Basically, for both online calculators, their purpose is to calculate when the customer's anti-spam solution investment will be worth it. Hence, the parameters involved in the calculations are total cost investment including the cost of anti-spam solution including the employee's effort in filtering spam. Aiming for a smaller ROI value, this ROI calculator is to identify whether the anti-spam solution that they have invested will be worth it or not. However, it also depends on the amount of email spam received in their system.

While research by Sonnenreich et. al provided a more general way of calculating the ROI for all security problems (Sonnenreich, Albanese, and Stout 2006), it does not show the detailed solution or cost involved for the spam email case. Nonetheless, since this cost model is based on the survey and scoring method, it is not easy to obtain these values and this solution is suitable if it is valued at the organization level.

However, this cost category was omitted from further research mainly because this cost category is important only for the management and decision level.

Evaluation on cost parameters for inconvenience cost

Inconvenience cost is stated as “*user's valuation of the negative effects of spam including time spent, loss of useful email, intangible psychological distress, decreased labour productivity and inconvenience of having to avoid using email*” (Kim et al. 2006). Listed below are the parameters from this study:

- i. Number of spam messages
- ii. Use of anti spam program
- iii. Type of spam received
- iv. Email storage capacity
- v. Email service price

Data for this research were collected based on 1,000 residents of Seoul, South Korea. They divided the types of spam into commercial and obscene. This research provides a unique new way of

estimating the intangible negative impact of email spam. Since this is the only research that focuses on estimating the intangible cost, it is impossible to compare parameters involved in this cost with other research.

This section provides an in-depth study of email spam, its impact and related email spam cost models using different methodologies which is also one of the main contributions of this thesis. The outcome of this subsection will also help us in deciding the best method to conduct our research. A detailed study on each cost model will help in evaluating the cost categories and cost parameters used in email spam which can be further used for our Spam 2.0 cost model. Other cost models such as data centre and web server are investigated as it is the closest model for the purpose of contrast. Justification is provided as to why the new cost model of Spam 2.0 is needed.

2.2.2 Other Related Cost Models

This section provides studies on cost models that are presented quite similar to the email cost model. Section 2.2.2.1 provides a summary of the related costs models such as the cost model focus, cost category, cost parameters and the formulation itself. A critical analysis to compare these studies is presented in Section 2.2.2. The thesis further discusses similarities and differences between these cost models. A detailed analysis on why these cost models are not fitted with Spam 2.0 is also included in this subsection.

2.2.2.1 Literature review on related cost models

This section presents the studies on cost models that contain similarities with the email cost models and also contain possible features similar to Spam 2.0. For clarity, cost models that are presented in this section are summarized in Table 2.5. The evaluation for each cost model shows the similarities and differences of each cost model. Cost parameters that will be beneficial for our cost models are highlighted in the next subsection.

Table 2.5: Literature review on related cost models

Literature	Cost model focus	Cost category	Cost parameters	Cost model
1 Joshi and Joshi (2012)	Web applications and operating cost of data centre	Human labour, H Components procurement, P Components maintenance, M Software licence, SL Software support, SS Energy utilization, E Space usage, L		Operating cost of data centre at time instant t , $OP(t) = H_{cost}^t + P_{cost}^t + M_{cost}^t + SL_{cost}^t + SS_{cost}^t + E_{cost}^t + L_{cost}^t$Eq. 2.20

Literature	Cost model focus	Cost category	Cost parameters	Cost model
2 Mihut and Tomai (2010)	IT department	Cost of space Cost of power Cost of cooling Cost of operation	<ul style="list-style-type: none"> • Total cost of space, a • Area of space occupied with equipment, b • Area of space occupied with server, c • Total cost of power, d • Total power of IT equipment, e • Total power consumed by server, f • Period of time, g • Total cost of cooling, h • Total weight of IT equipments, i • Total of weight of server, j 	$Cost_{total} \text{ for output Administration of IT server} =$ $Cost_{space} \text{ occupied with the server} +$ $Cost_{power} \text{ consumed by server} + Cost_{cooling} \text{ for server} +$ $Cost_{operation} \text{ for Administration of IT server} \dots \dots \dots \text{Eq. 2.21}$ <ul style="list-style-type: none"> • $Cost_{space}$ occupied with the server = $a * c / b$..Eq. 2.22 • $Cost_{power}$ consumed by server = $d * (f * g) / (e * g)$..Eq. 2.23 • $Cost_{cooling}$ for server = $h * j / i$..Eq. 2.24 • $Cost_{operation}$ for Administration of IT server {salary, purchase new products, courses for personnel, purchase of spare parts of equipments, programming} = {setup and configuring + service + develop new programs + support for users} ..Eq. 2.25

	Literature	Cost model focus	Cost category	Cost parameters	Cost model
3	Li et al. (2009)	Cloud computing	<p>Server cost</p> <p>Software cost</p> <p>Facilities cost</p> <p>Support and maintenance cost</p> <p>Network cost</p> <p>Power cost</p> <p>Cooling cost</p> <p>Space cost</p>	<p>Amortization period unit, Arp</p> <p>Hours consumed, $time$</p> <p>Number of physical servers in resource pool, N_s</p> <p>Cost per physical server of same configuration, VI_{ps}</p> <p>Unit price of type II software, VI_s</p> <p>Unit price of type I software, VI_o</p> <p>Unit price of type III software, VI_m</p> <p>Subscription factor – percentage of unit price that yield annual fee, S_s, S_o, S_m</p> <p>Number of type II software licence, $N_{servtic}$</p> <p>Number of type I software licence, N_{oslic}</p> <p>Number of type III software licence, $N_{monitoric}$</p> <p>Number of new network switches per year. N_{switch}</p> <p>Number of network interface controllers (NICs) per virtualized server, S_{NIC}</p> <p>Number of ports per NIC, P_{NIC}</p> <p>Price per switch, P_s</p> <p>Port number of a network switch, N_{port}</p> <p>Number of administrators responsible for support and maintenance, N_{Labour}</p> <p>Average time spent on unit system under utilization, T_{use}</p>	<p>$Arp = (1 + 0.05) * time / (30 * 24 * A_p)$Eq. 2.26</p> <p>$Cost_{servers} = VI_{ps} * N_s * Arp(time)$Eq. 2.27</p> <p>$Cost_{software} = [S_s * VI_s * N_{servtic} + S_o * VI_o * N_{oslic} + S_m * VI_m * N_{monitoric}] * Arp(time)$Eq. 2.28</p> <p>$Cost_{networking} = P_s * N_{switch} * Arp(time)$Eq. 2.29</p> <p>$N_{switch} = S_{NIC} * P_{NIC} * N_s / N_{port}$Eq. 2.30</p> <p>$Cost_{support\ and\ maintenance} = N_{Labour} (T_{use} * N_s + T_{idle}) R_{salary}$Eq. 2.31</p> <p>$Cost_{power} = L_s * E_s * S_{rp} * Arp(time)$Eq. 2.32</p> <p>$Cost_{cooling} = L * (1 + P) * Powercost(time) / H...$ Eq. 2.33</p> <p>$Cost_{facilities} = N_{rack} * VP_{fp} * Arp(time)$Eq. 2.34</p> <p>$Cost_{space} = A_p * S_{pace} * Arp$Eq. 2.35</p> <p>$S_{pace} = (R_{SF} * N_{rack}) / R_{SPACE}$Eq. 2.36</p> <p>$(N_s * W_{Server} + N_{rack} * W_{rack}) / S_{pace} \leq C_{pressure}$...Eq. 2.37</p> <p>$TCO = Cost_{servers} + Cost_{software} + Cost_{networking} + Cost_{support\ and\ maintenance} + Cost_{power} + Cost_{cooling} + Cost_{facilities} + Cost_{space}$</p>

Time spent on all the idle systems, T_{idle}
 Rating number of salary averages, R_{salary}
 Sum of the power rating of working servers, S_{rp}
 Price per hour of 1kW of electricity, E_s
 Steady-state constant, L_s
 Number of racks in working, N_{rack}
 Cooling load factor, L
 Airflow redundancy constant, P
 Inefficiency constant, H
 Number of racks, N_{rack}
 Price of facilities per rack, VP/p
 Hours consumed, $time$
 Cost per square foot to build cloud, A_p
 Square feet per rack, R_{SF}
 Percent of space taken by racks in all, R_{SPACE}
 Weight of a physical server, W_{Server}
 Weight of a rack, W_{rack}
 Constant pressure confronted by unit floor,
 $C_{pressure}$
 Annual percentage rate, R_a

	Literature	Cost model focus	Cost category	Cost parameters	Cost model
4	Karidis, Moreira, and Moreno (2009)	Data centre	<p>Construction of data building centre</p> <p>Power and cooling infrastructure</p> <p>Acquisition cost of the servers</p> <p>Cost of electricity to power and cool the servers</p> <p>Labour cost of operating and managing servers</p>	<p>Target power density of data centre, D</p> <p>Infrastructure cost per watt of capacity for desired tier level, T</p> <p>Cost per square foot of raised floor, F</p> <p>Normalized cost (dollar per watt of capacity) of data centre, C</p> <p>Cost of the mechanical and electrical infrastructure for 1kWh of computing, I</p> <p>Floor space cost per kWh of computing, F</p> <p>Server cost, S</p> <p>Average utilization of the servers in data centre, u</p> <p>Electricity cost, $E(u)$</p> <p>Total physical cost per hour for each kW of server installed, $C(u)$</p> <p>Server cost per kWh for m-socket server, $S(m)$</p> <p>Labour cost per kWh for m-socket server, $L(m)$</p>	<p>$C = T + F/D$Eq. 2.38</p> <p>$C(u) = I + F + S + E(u)$Eq. 2.39</p> <p>Total (physical plus labour) cost of computing for an m-socket server operating at utilization u,</p> $\frac{C(u)}{u} = \frac{I+F+S(m)+L(m)+E(u)}{u} / kWh$Eq. 2.40

Literature	Shah and Patel (2005)	Cost model focus	Cost category	Cost parameters	Cost model
5	Shah and Patel (2005)	Data Centre	Cost space Cost power + hardware Cost cooling Cost operation	Net operating income (in \$), calculated as difference between gross operating income and actual operating costs, NOI Area of the computer room, $A_{data\ centre}$ Capitalization rate, Cap Cost of grid power per month, $U_{\$grid}$ Power consumed by the hardware, networking or cooling equipment, $P_{consumed\ hardware}$ Amortization and maintenance costs of either the power delivery or cooling equipment, $U_{\$,A\&M}$ Cooling load factor, L_1 Capacity utilization factor, J_1 Burdened power delivery factor, K_1 Burdened cooling cost factor, K_2 Ratio of the total no of information technology personnel servicing data centre to the number of racks, M_1 Ratio of the total number of facilities personnel servicing the data centre to the number of racks in the data centres to the number of racks, M_2	$Cost_{total} = Cost_{space} + Cost_{power\ hardware} + Cost_{cooling} + Cost_{operation}$ <p>.....Eq. 2.41</p> $Cost_{space} = \frac{(NOI/ft^2)(A_{data\ centre})(\%Occupancy)}{Cap\ Rate}$ <p>.....Eq. 2.42</p> $Cost_{power} = U_{\$,grid} P_{consumed\ hardware} + U_{\$,A\&M} power P_{consumed\ hardware}$ <p>.....Eq. 2.43</p> $Cost_{cooling} =$ <p>.....Eq. 2.44</p> $U_{\$,grid} L_1 P_{consumed\ hardware} + U_{\$,A\&M} cooling J_1 L_1 P_{consumed\ hardware}$ <p>.....Eq. 2.44</p> $Cost_{space\ power\ cooling} =$ <p>.....Eq. 2.45</p> $\left(\frac{\$}{ft^2}\right) (A_{critical}, ft^2) + (1 + K_1 + L_1 + K_2 L_1) U_{\$,grid} P_{consumed\ hardware}$ <p>.....Eq. 2.46</p> $Cost_{personnel\ per\ rack} = (M_1 + M_2 + M_3) S_{avg} = M_{total} S_{avg}$ <p>.....Eq. 2.46</p> $Cost_{depreciation\ per\ rack} = IT_{dep} = \frac{Rack\ Purchase\ Cost}{L_{lifetime\ of\ rack\ (usu.3\ years)}}$ <p>.....Eq. 2.47</p> $Cost_{software\ per\ rack} = \sigma_1 = \frac{Total\ licensing\ cost}{R}$ <p>.....Eq. 2.48</p> $IT\ Operation\ Cost_{total} = R(M_{total} S_{avg} + IT_{dep} + \sigma_1)$ <p>.....Eq. 2.49</p>

				<p>Ratio of the total number of administrative personnel servicing the data centre to the number of racks in the data centre, M_3</p> <p>Ratio of the total number of all personnel servicing the data centre to the number of racks in the data centres, M_{total}</p> <p>Average salary of data centre IT, facilities, administrative person per rack, S_{avg}</p> <p>Straight line monthly depreciation of IT equipment, IT_{dep}</p> <p>number of racks utilized in a data centre, R</p> <p>Software and licensing costs per rack per month, σ_1</p>	<p>$Cost_{total} =$</p> $\left(\frac{\$}{f^2}\right) (A_{critical}, f t^2) + (1 + K_1 + L_1 + K_2 L_1) U_{\$grid} P_{consumed hardware} + R(M_{total} S_{avg} + IT_{dep} + \sigma_1) \dots \dots \dots \text{Eq. 2.50}$
--	--	--	--	---	--

2.2.2.2 Evaluation

Section 0 presented a summary of the literature review on other related cost models. In this section, the critical evaluations of vital information gathered from these studies are presented, which are:

1. Evaluation on the cost models
2. Evaluation on the cost categories

The differences and similarities existing (if there is any) in each cost model and cost category are going to be highlighted in these subsections.

Evaluation on the cost models

It is important to consider that cost models presented in Table 2.5 have a different focus. For example, data centre cost models were developed by both Patel and Shah (2005) and Karidis, Moreira and Moreno (2009). Nonetheless, Karidis, Moreira and Moreno in their paper also provided an extended research on the cost/performance model for data centres. In 2009, Li et. al developed a cost model for cloud computing. This cost model presents a lot of similarities to data centre cost models, except that it focuses on the need of dynamic scalability as cloud computing adopts the architecture to continuously adapt to the users' changing requirements automatically (Li et al. 2009). On the other hand, the cost model developed by Mihut and Tomai in 2010 is focused on IT department. This cost model is a combination of the cost model developed by Patel and Shah (2005) with activity-based costing (ABC) method to calculate the cost of operating an IT department. Joshi and Joshi (2012) developed a cost model focusing on data centres with extra emphasis given catering to the problem of the new web applications development (Joshi and Joshi 2012).

Components used in developing data centres, cloud computing and IT departments possess similarities with our cost models where the expenses incurred fall into similar cost categories. When comparing with the email spam cost model, several categories matched up with cost parameters in IT costs (refer to 'Both of these cost models have been generated from different perspectives. Rockbridge Associates has calculated the actual profit made by the spammers taking into account the actual purchases made by the respondents (Rockbridge Associates Inc. 2009, 2005). On the other hand, Judge, Alperovitch and Yang have estimated this value by considering the probability of not getting caught (Judge, Alperovitch, and Yang 2005).

Nonetheless, Rockbridge Associates has used generalization in its dataset whereby it calculated the percentage of respondents who made purchases as a result of spamming activities (Rockbridge Associates Inc. 2009, 2005). Furthermore, pre-defined values, 1 spam equal to 1 cent, have been used for spammers' revenues.

Although there were no predefined values in the research by Judge, Alperovitch and Yang, their method to obtain these cost parameters is unclear. It is also unclear how they have generated the value of each parameter. In contrast, Rockbridge Associates has provided a clear method for obtaining the data, i.e., by using telephonic interviews (Rockbridge Associates Inc. 2009). However, Judge, Alperovitch and Yang's research provides an insight to further consider other parameters, such as cost to send each message, and cost of punishment (Judge, Alperovitch, and Yang 2005). In fact, it could be seen that Rockbridge Associates' way of estimating the spammers' profit costs actually consists of the first part of the cost model used by Judge, Alperovitch and Yang.

Evaluation of cost parameters for IT costs' on page 63) such as labour, software and hardware costs. Thus, this information could be used as an extra valuable input to develop an accurate cost model for Spam 2.0, especially for the company-owned server and other hardware themselves. Nonetheless, the difference is clear in the naming of cost categories. Since data centres and cloud computing emphasize on power and cooling, there is a need to define the cost of power and cooling. The email spam cost model on the other hand highlights the cost of storage for keeping the email spam content; hence, there is a need to define the cost of storage. Furthermore, when comparing the cost model of data centres, cloud computing and IT departments with email spam and Spam 2.0, the costs of space, power, cooling, electricity, operation and labour are used as a whole. Nonetheless, these costs are not fully utilized in the case of email spam and Spam 2.0.

Thus, it is obvious that the cost categories included in the cost model will depend on the purpose of developing the cost model and the emphasis given when operating the type of cost model i.e. email spam, Spam 2.0, data centre, cloud computing and IT department.

Evaluation on the cost categories

The thesis first provides a frequency analysis of cost categories as shown in Table 2.6, to observe which cost categories are included and defined in the related cost models. The thesis enlists all the cost categories that are clearly defined in these cost models and evaluates its frequency.

Table 2.6: Frequency analysis for the cost categories

	Literature cost categories	Joshi and Joshi (2012)	Mihut and Tomai (2010)	Li et al. (2009)	Karidis, Moreira, and Moreno (2009)	Patel and Shah (2005)	Frequency
1	Space cost	√	√	√	√	√	5
2	Power cost		√	√	√	√	4
3	Cooling cost		√	√	√	√	4
4	Energy utilization	√					1
5	Components procurement	√					1
6	Operation cost		√			√	2
7	• Labour cost	√			√		2
8	• Support cost			√			1
9	• Maintenance cost	√		√			2
10	• Facilities cost			√	√		2
11	• Hardware					√	1
12	• Server cost			√			2
13	• Software cost			√			1
	• Software licence	√					1
	• Software support	√					1
14	• Network cost			√			1

As shown in Table 2.6, we have identified 16 cost categories from all five related cost models. Space cost was defined in all five cost models. Space cost in Karidis, Moreira and Moreno (2009) however was defined as the construction of a data-building centre. Power cost and cooling cost were defined separately in three out of five cost models. Both power cost and cooling cost however were combined in the research by Karidis, Moreira and Moreno (2009).

Operation cost is clearly defined in two out of five cost models. It is observed that operation cost could contain several other cost categories. The operation cost that was defined in Mihut and Tomai (2010) included labour cost, hardware, facilities, software and support and maintenance cost. Instead, Patel and Shah (2005) only included labour cost, software licence and software support. In this research, hardware cost is combined with power in another cost categories (Patel and Shah 2005) . Operation cost was not defined clearly in the research of Joshi and Joshi (2012), Li et al. (2009) and Karidis, Moreira, and Moreno (2009). However, these researches defined other cost categories explicitly. For example, labour, maintenance, facilities and server costs were all distinctly defined in two out of five cost models. Karidis, Moreira, and Moreno (2009) in their research defined power and infrastructure cost which is also categorized as facilities cost.

In addition, support cost, hardware cost, software cost and network cost were also separately defined once in these studies. Software cost defined in Li et al. (2009) included both price and licence cost for software and defined support and maintenance cost in another cost category. By contrast, software cost in Joshi and Joshi (2012) is defined individually as two cost categories which are software licence and software support.

Other cost categories that were defined in one out of five cost models are energy utilization and components procurement. Energy utilization that was clearly defined in Joshi and Joshi (2012) is considered as power and cooling cost. Components procurement that was defined in Joshi and Joshi (2012) could also be categorized as hardware and facilities cost.

Nonetheless, there are some repetition and overlapping between all these 16 categories. Although some of the cost categories are not clearly defined in some of the cost models, it does not mean that it was not included. Some of them are included in other cost categories resulting in different cost models with different number of cost categories. For example, Joshi and Joshi (2012) introduced seven cost categories, Mihut and Tomai (2010) with four cost categories, Li et al (2009) with eight cost categories, and Karidis, Moreira and Moreno (2009) and Patel and Shah (2005) with five and four cost categories respectively.

As the cost categories presented in this subsection are comparable to cost parameters provided in the email spam cost model (refer to Section 2.2.1.4), we also need to evaluate attributes involved in the cost categories in this subsection which are space cost, power cost, cooling cost and operation cost.

Operation cost also included other subcategories such as labour cost, support cost, maintenance cost, facilities cost, hardware cost, server cost, software cost and network cost.

Space cost

Space cost was included as space usage in Joshi and Joshi (2012), cost of space in Mihut and Tomai (2010), space cost in Li et al. (2009), construction of data building center in Karidis, Moreira and Moreno (2009) and cost space in Patel and Shah (2005).

The cost of space in Mihut and Tomai (2010) considered attributes such as the total cost of space, area of space occupied with equipment and area of space occupied with the server. Meanwhile, space cost in Li et al. (2009) included attributes such as cost per square foot, square feet per rack, percent of space taken by racks in all and number of racks involved in defining their space cost. Karidis, Moreira and Moreno (2009) defined attributes such as floor space cost per kWh of computing to estimate the cost of construction of data-building centre. Cost space as defined in Patel and Shah (2005) included attributes such as net operating income per square foot, area of the computer room, occupancy percentage and capitalization rate. Space usage in Joshi and Joshi (2012) however was not mentioned clearly as the cost models were created to address changes for new additional web applications and operating cost.

There were several differences in the attributes defined to estimate the cost of space. Nonetheless, the focus of this cost category in all cost models was similar which was to estimate the cost of the space used by the server in the set-up. In addition, the cost of space was not a major concern when estimating the cost of email spam which is why this cost category was not included in any of the email spam cost models.

Power cost

Power cost was clearly defined as cost of power in Mihut and Tomai (2010), power cost in Li et al. (2009) and power and cooling infrastructure in Karidis, Moreira and Moreno (2009). This cost category was not clearly defined in Joshi and Joshi (2012) but it was considered under the category of energy utilization. In the research of Patel and Shah (2005), power cost was calculated combined with hardware.

In order to calculate the power cost consumed by the server, Mihut and Tomai (2010) have defined attributes such as the total cost of power, total power consumed by the server, period of time and total power of IT equipment. Li et al. (2009) defined attributes such as the sum of the power rating of working servers, price per hour of 1kW of electricity and steady-state constant to calculate power cost. Power cost in Karidis, Moreira and Moreno (2009) was combined with cooling cost to calculate the cost of electricity and it considered attributes such as the average utilization of the servers in the data centre and electricity cost. This cost also included all the power distribution equipment (cables,

transformer and panels) and power backup equipment (Karidis, Moreira and Moreno 2009). Attributes that were defined to calculate power cost in Patel and Shah's research were the cost of grid power per month, power consumed by the hardware and amortization and maintenance costs of power delivery equipment. The maintenance cost of power delivery equipment could also be categorized under maintenance cost.

It is observed that power cost was defined to calculate the cost of power involved in the activity and mainly focused on the server. Similar to space cost, it was observed that this cost category was also not included in any email spam cost models.

Cooling cost

Cooling cost was visibly defined in the research of Mihut and Tomai (2010) as cost of cooling, cooling cost in Li et al. (2009) and cost cooling in Patel and Shah (2005). Cooling cost in Karidis, Moreira and Moreno (2000) was combined under the cost category called power and cooling infrastructure. Similar to power cost, cooling cost was also not clearly defined in Joshi and Joshi (2012) but it was considered under the category of energy utilization.

Mihut and Tomai (2010) defined cost of cooling by focusing on the cost of cooling for the server. Cost of cooling was calculated based on attributes such as the total cost of cooling, total weight of IT equipment and total of the weight of the server. In the research of Li et al. (2009), cooling cost was calculated based on the power consumed by equipment which is then converted to heat. Hence, cooling cost is calculated using several attributes such as the cooling load factor, airflow redundancy constant and inefficiency (humidification) constant. Cooling cost in the research of Karidis, Moreira and Moreno (2009) was combined with power cost to produce power and cooling infrastructure. This power and cooling infrastructure was focused on calculating the cost based on server infrastructure with different tiers. Some of the cooling equipment included in the calculation were water chillers, heat exchangers and computer room air conditioners (Karidis, Moreira, and Moreno 2009). Cost cooling defined in the research of Patel and Shah (2005) included attributes such as cost grid per month, cooling load factor, power consumed by the hardware, capacity utilization factor and maintenance cost of cooling equipment.

It is observed that cooling cost would focus only on working servers. The percentage of cooling used only for working servers is estimated based on the weight of the server out of the total equipment (Mihut and Tomai 2010). In this case, the information on total cooling cost needs to be obtained. Researches by Li et al. (2009), Karidis, Moreira and Moreno (2009) and Patel and Shah (2005) provided detailed information on calculating the total cooling cost. They estimated cooling cost based on the heat produced by working servers combined with other factors which could affect the heating and cooling environment. Nonetheless, this cost category was also not included in the email spam cost model as it was not a major cost in the email spam environment.

Operation cost

Operation cost was clearly defined in Mihut and Tomai (2010) and Patel and Shah (2005) as cost operation. Operation cost in this subsection was comparable to operating cost that we have defined in the evaluation of email spam cost models (refer to page 65).

The operation cost that was defined in Mihut and Tomai included salary, purchase of new products (both hardware and software), courses for personnel, purchase of spare parts for equipment and programming but there was no detailed information given on the attributes involved (Mihut and Tomai 2010). Furthermore, Mihut and Tomai's cost model focused on combining traditional costing with activity-based costing (Mihut and Tomai 2010); hence, it produced a total cost in different unit compared to other cost models.

For IT operation cost in their research, Patel and Shah defined attributes such as number of racks utilized in a data centre, ratio of the total number of all personnel servicing the data centre to the number of racks in the data centres (including IT, administrative and facilities personnel), software licensing costs per rack per month, straight-line monthly depreciation of IT equipment and average salary of data centre IT, facilities and administrative per rack (Patel and Shah 2005). IT operation cost in this research basically covers labour, support, maintenance, software and hardware.

Nonetheless, as mentioned earlier, operation cost was defined as any cost involved in running the task. Other research did not clearly define operation cost but they included other cost categories such as labour cost, support cost, maintenance cost, facilities cost, hardware cost, server cost, software cost and network cost. For research that has explicitly defined these costs, the thesis further explains them below.

Labour cost

Labour cost was clearly defined in the cost model by Joshi and Joshi (2012) as human labour and Karidis, Moreira and Moreno (2009) defined it as the labour cost of operating and managing servers. Joshi and Joshi (2012) however did not provide the attributes involved in the cost model. Karidis, Moreira and Moreno (2009) defined labour cost per kWh for an m -socket server to calculate the labour cost of operating and managing servers.

Although labour cost was not mentioned clearly, labour cost is also included in other cost categories such as the cost of operation for administration of IT server (Mihut and Tomai 2010), support and maintenance cost (Li et al. 2009) and operation cost (Patel and Shah 2005). Labour cost in Mihut and Tomai (2010) however was not addressed explicitly but was combined with other aspects such as setup and configuring, servicing, developing new programs and giving support for users, salary and courses for personnel and purchasing new products. Labour cost in Li et al. (2009) was addressed as support and maintenance cost that considered a few parameters such as the number of administrators

involved in support and maintenance, their salary, number of server, average time spent for support and maintenance and time spent on all idle systems. The labour cost that was included in Patel and Shah (2005) defined several attributes such as average salary and ratio of the total number of IT personnel, facilities personnel and administrative personnel servicing a data centre to the number of racks in a data centre. It was defined under operation cost that considered other cost categories such as software cost and server cost.

This type of cost is comparable to labour cost in the email spam cost model, but in the email cost model, the thesis specifies labour cost as one of the IT costs. Labour cost in email spam is usually defined clearly. However, in this subsection, labour cost can be defined clearly or it is included in other cost categories. The labour cost in email spam cost models defined earlier included all labour work including installation, maintenance or any other labour work involved in the activities except for support cost under one cost category. In this subsection, there is no specific practice observed in defining labour cost. Labour cost in the research of Karidis, Moreira and Moreno (2009) was calculated under the server's administration and operation. Mihut and Tomai (2010) did not define labour cost specifically but considered labour cost for set-up, operation and support. In addition, Li et al. (2009) included labour cost as one of the attributes in calculating cost for support and maintenance. Patel and Shah (2005) included cost personnel per rack under operation cost and considered costs for IT personnel, facilities personnel and administrative personnel.

Nonetheless, similar to email spam cost model, labour cost is usually defined as the cost that covers the work involved in the main activities. For example, considering for the spammer, labour cost will include the cost for labour in installation, maintenance, main production and acquiring customers, while calculating the labour cost for anti-spam will involve the management, support and system administration work in dealing with spam. Compared to the email spam cost model which usually defined labour cost as total hours spent specifically for employees in charge of spam or anti-spam work, the labour cost that was defined in this cost model depends on the number of working servers either per socket (Karidis, Moreira, and Moreno 2009) or per rack (Patel and Shah 2005).

Support cost

Support cost was defined in the cost model of Joshi and Joshi (2012) as software support while Li et al. (2009) defined it as support and maintenance. Support cost was also included in one of the attributes to calculate the cost of operation for the administration of an IT server in the research of Mihut and Tomai (2010). Nonetheless, no detailed attributes were included in both Joshi and Joshi (2012) and Mihut and Tomai (2010).

Li et al (2009) in their research combined both support and maintenance costs and listed attributes such as the number of administrators responsible for support and maintenance, average time spent on unit system under utilization, time spent on all the idle systems and rating number of salary averages.

Compared to the email spam cost model, support cost was used to define for helpdesk service or support services dealing with spam. However, it was observed that this type of cost category in this subsection can also be included in labour cost.

Maintenance cost

Maintenance cost was defined in Joshi and Joshi (2012) as components maintenance whilst Li et al. (2009) defined it as support and maintenance cost. Joshi and Joshi (2012) did not provide the details of attributes involved, while support and maintenance cost in the research of Li et. al (2009) was already covered in the previous subsection.

This type of cost category was not included in the email spam cost model. Still, this cost is important to be included if the maintenance involved an extra cost that needs to be paid to a third party. Otherwise, this cost can also be included in labour cost.

Facilities cost

Facilities cost was defined in the research of Li et al. (2009) while Karidis, Moreira and Moreno (2009) defined it as facilities cost and construction of data-building data centre. The attributes defined to calculate facilities cost were the number of racks and price of facilities per rack (Li et al. 2009). Karidis, Moreira and Moreno (2009) defined attributes such as infrastructure cost per watt of capacity for desired tier level to calculate facilities cost. This was another way of calculating the total cost of server and relevant hardware. Although facilities cost was not defined clearly in other research, it is similar to components procurement (Joshi and Joshi 2012) and power and hardware cost (Patel and Shah 2005).

This type of cost category is suitable to be considered as a setup cost or cost for relevant facilities needed to run the main activities. Hardware cost and server cost can also be included in this cost category.

Hardware cost

Hardware cost was defined as cost of power and hardware (Patel and Shah 2005). However, in this subsection, hardware cost is usually included under other cost categories such as components procurement (Joshi and Joshi 2012) and facilities cost (Li et al. 2009).

Hardware cost defined in the email spam cost model listed the basic hardware needed to combat or send email spam such as cost for a computer, cost for a monitor and cost for peripheral devices.

Server cost

Server cost was defined in Li et al. (2009) as server cost whilst Karidis, Moreira and Moreno (2009) defined it as acquisition cost of the servers. Server cost could be calculated using attributes such as cost per physical server of the same configuration and number of physical servers in the resource pool

(Li et al. 2009). Server cost in the research of Karidis, Moreira and Moreno (2009) was defined as server cost per kWh for an m -socket server.

This type of cost category was also not included in the email spam cost model. However, it was included in hardware cost or translated to storage cost. In this subsection, server cost is sometimes not defined clearly, however it can also be calculated under other cost categories such as components procurement (Joshi and Joshi 2012) and IT operation cost (Patel and Shah 2005). IT operation cost for example includes cost depreciation per rack which includes attributes such as rack purchase cost and lifetime of rack (Patel and Shah 2005).

Software Cost

Software cost was clearly defined in the research by Li et al. (2009), while in the research of Joshi and Joshi (2012), software cost was divided into two cost categories called software licence and software support. Attributes that were used to calculate software costs were subscription factor (percentage of unit price that yield annual fee), unit price of software which depends on the type and number of software licence which depends on the type (Li et al. 2009). Software cost was also indirectly included in the operation cost for the administration of an IT server (Mihut and Tomai 2010). Attributes that were used to define the cost of software per rack under operation cost for the administration of an IT server are total licensing cost and number of racks utilized in a data centre.

Software cost was also defined in email spam cost models. Email spam cost models listed the details of software types such as anti-spam product and services, operating system, remailers and mail address harvesters. Nonetheless, software cost in this subsection was calculated thoroughly as it had included licence costs and cost for software support.

Network cost

Network cost was defined only in the research of Li et al. (2009) as cost of networking. The attributes involved in calculating this cost was price per switch, number of physical servers of same configuration, number of NICs per virtualized server, number of ports per NIC and port number of a network switch.

This cost category was also not defined in the email spam cost models; however, it is comparable to connectivity cost. Nonetheless, the attributes defined in this subsection are more detailed and suitable if a company sets up a big network system.

2.2.3 Summary of Literature Review on Cost Models

In Section 2.2, a detailed explanation of the relevant literature on cost models was given. Figure 2.12 presents a graphical summary of cost models included in this subsection.

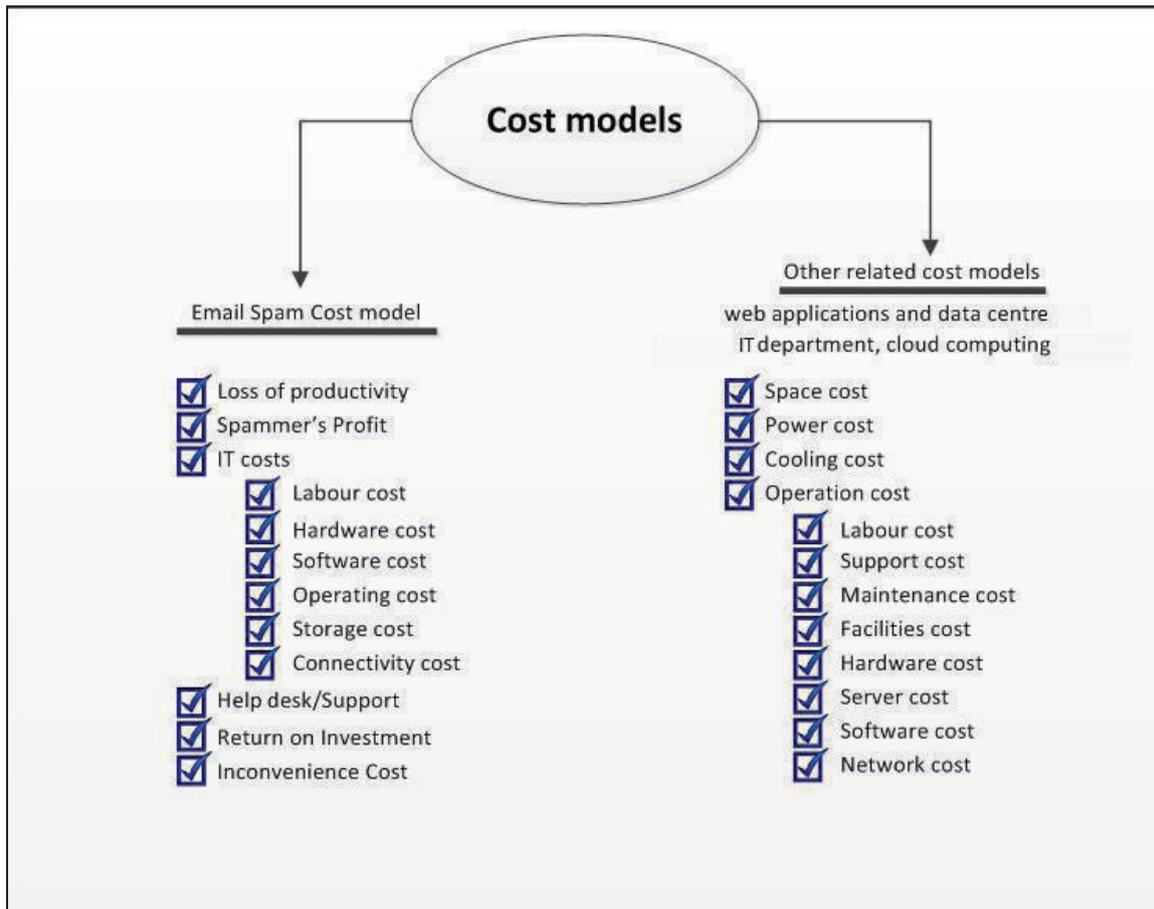


Figure 2.12: Graphical summary of literature review on cost models

Reviews on email spam cost models and other related cost models are done such as relevant cost categories, cost parameters and attributes to provide a better input towards the cost model of Spam 2.0. As mentioned earlier, the research focus is twofold, a cost model of Spam 2.0 and also a survey on public awareness, knowledge and perception of Spam 2.0. Hence, Section 2.3 further provides an overview of studies done on public awareness, knowledge and perception.

2.3 Public Awareness, Knowledge and Perception of Spam 2.0

This section discusses previous related works on public awareness, knowledge and perception. To the best of the authors' knowledge, currently, there is no study which focuses on awareness, knowledge and perception issues in the context of Spam 2.0. Thus, related studies in the area of computer security are considered and, where possible, the thesis also includes related studies in public health and general crime. The explanation in this section is divided into three subsections:

- Public awareness
- Knowledge

- Perception

In each section, the definitions of related terms are given, followed by detailed purpose, focused issues and related questions used to assess particulars and participants involved for each study. This section is the basis of the stages for questionnaire development in this research.

2.3.1 Public Awareness

Public awareness or generally termed as social awareness is defined as “*naming the problem, speaking out, consciousness raising and researching*” (Greene and Kamimura 2003). From this definition, it is clear that raising awareness could further lead to raising consciousness towards an issue. Hence, it is anticipated that the public should be made aware of the issues of Spam 2.0.

From the author’s view, there is only minimal information on public awareness particularly on Spam 2.0. Perhaps the concept of awareness itself was not seen as common research in traditional engineering and hard computer science; thus, this issue is typically taken lightly by researchers in this field (Siponen 2001). However, without a certain level of adequacy in the awareness of Spam 2.0, the Spam 2.0 campaign will easily be spread out on the Internet. Furthermore, the main target of Spam 2.0 is the public user; hence, if the users are unaware of Spam 2.0, then they will be easily exposed and become a victim of the Spam 2.0 campaign.

The most common awareness issue that was highlighted in computer security-related reports was training awareness. Such studies also investigated the cost or budget, standards, policies and procedures implemented in those organizations (Quinn 2006; Stander, Dunnet, and Rizzo 2009). This type of study usually focused on organizations with the survey targeted at security practitioners in a particular country. The objective of this type of research was to assess whether awareness training is effective or not in the organization. However, research that focused on public level awareness was commonly carried out to merely investigate whether the public were aware of the specific issues being studied. From this point, some studies also extended their focus to include testing hypotheses and finding relationships between defined variables. The thesis now provides other awareness-related research in the computer field.

An awareness of information security was highlighted in a quantitative study on a sample size of 1,483 Japanese Internet users (Takemura and Umino 2009). Eight hypotheses were initially defined to study the relationship between awareness to information security with job, gender, attitude towards risk, age, Internet use term, habitation, information security educating situation and experience of encountering information security incidents (Takemura and Umino 2009). In this research, Internet users’ awareness to information security was defined by five indexes which are recognition concerning individual information, recognition concerning illegal copies, recognition concerning

counter measures and awareness to moral and recognition concerning the Internet (Takemura and Umino 2009). From the analysis of variance based on the non-parametric method, it was reported that Japanese Internet users' awareness to information security is different in terms of individual attributes.

In another awareness-related research, self-awareness before and after joining social networks was studied and its impact on information security vulnerability was simultaneously investigated (Hasan and Hussin 2010). In this research, users' awareness of using social networks was evaluated based on their behaviour during account registration such as whether they read the terms and conditions provided by the social networking applications during sign-up and whether they are using the same password for their email account and social network account. Respondents were also asked if they had revealed their personal information on the social network. The respondents were gathered from a close-ended survey involving 119 Malaysian students (Hasan and Hussin 2010). Based on the data collection and case studies in Pakistan and Bangladesh mentioned in this paper, the authors concluded that most users are unaware when using social networks, and having a lack of awareness caused them to disclose their personal information which leads to information security vulnerability (Hasan and Hussin 2010).

In an earlier study, there was also a research by Lang et al. (2009) on awareness of the potential risks of social networking sites while focusing on attitudes towards data security issues. However, their respondents were university students in the age group of 18–24 years living in Ireland. Data analysis in this research was collected from a web-based survey from 351 respondents and through a meta-analysis of 120 profiles on social networking sites including Facebook (FB) and Bebo. Some of the analysis of web survey responses was reported based on the profile of the respondents, password security, attitudes towards the risk of data loss, awareness of viruses and similar threats and security of personal data on social networking sites. On the other hand, analyses of social network experimental data are more focused on whether strangers' invitations are accepted or not and on the sharing of sensitive data. In order to assess the awareness of viruses and similar threats, researchers simply asked direct questions to enable them to evaluate respondents' awareness such as "*Have they ever experienced a virus, worm or other intruder on their computer?*" (Lang et al. 2009). To this question, 22% of respondents stated that they were unsure (Lang et al. 2009). Accordingly, 39%, 44% and 56% of respondents stated that they are totally unaware of Trojan, worms and malware, respectively (Lang et al. 2009). Respondents were also asked if they know that Bluetooth devices, CD, DVD and USB flash drives could carry viruses. This research stated that most respondents have a casual attitude towards data backup and password protection (Lang et al. 2009). In addition, most respondents were incompetently unaware about virus threats and shared sensitive information on their social networking sites (Lang et al. 2009).

Also focusing on social networking applications, Acquisti and Gross in their research studied the issues of awareness, information sharing and privacy on FB (Acquisti and Gross 2006). Data were collected from both 40-question survey data from 294 respondents in a US academic institution and 7,000 profiles mined from FB (Acquisti and Gross 2006). Analyses were conducted covering privacy concerns, FB usage, awareness of FB rules and profile visibility and attitudes towards FB. This research analysed the impact of privacy concerns on members' behaviour. Based on the result, it was found out that even individuals with privacy concerns joined the sites and shared excessive volumes of personal information (Acquisti and Gross 2006). Privacy issues were also managed by placing trust on the tools and options provided by FB. However, the research stated that they found the proof of misconceptions about the online community's actual size and the visibility of members' profiles. Focusing on the awareness of one's ability to control who can see one's profile, the question asked was "*Who can actually read your complete profile on the Facebook?*" with the options to choose "*Do not know,*" "*No control,*" "*Some control*" and "*Complete control.*" To this question, 22% stated that they did not know what the FB privacy settings were or did not remember if they had ever altered them (Acquisti and Gross 2006). About a quarter of the sample did not know where the location of the settings was (Acquisti and Gross 2006). On the other hand, to assess awareness of true visibility of their profile, respondents were asked if anybody at their institution could search their profile. To this question, 24% mistakenly believed that their profile could not be searched by anybody (Acquisti and Gross 2006). In addition, when asked "*How many people could search for their profile on Facebook?*" with options such as "*a few hundred,*" "*a few thousands,*" "*tens of thousands,*" "*hundreds of thousands*" and "*millions,*" more than half underestimated the size of this online community (Acquisti and Gross 2006).

A survey consisting of 53 questions was conducted on 175 respondents with the objective to assess public attitude and awareness of computer crime and abuse (Dowland et al. 1999). In this research, the researchers specifically assessed the awareness of relevant legislations on computer crime and abuse (Dowland et al. 1999). Moreover, they studied the influence of media towards public perception on this particular issue. In order to assess the awareness of relevant legislations, questions such as "*Have they every ever heard of certain acts in UK?*" were asked (Dowland et al. 1999). As a result, it was found out that the media were successful in giving information and raising awareness about the existence of computer crimes (Dowland et al. 1999). However, it was revealed that the media has failed in educating the public on the possible corrective actions (Dowland et al. 1999).

2.3.2 Knowledge

The thesis presents existing studies related to 'knowledge' in this section. It is expected that individual knowledge could assist the particular person in the recognition, management or prevention

of Spam 2.0. An individual who is knowledgeable about an issue will use their knowledge in the best possible way to address problems arising from Spam 2.0 and will act differently from those who are not knowledgeable. Hence, knowledge study is very important to see the level of understanding that the public has on a particular issue. For example, it is expected that an individual who is knowledgeable on Spam 2.0 will have the ability to recognize and differentiate between spam and non-spam, know the causes of clicking on spam, know how to seek help if Spam 2.0 existed in his personal space, know how to report Spam 2.0 and know how to further prevent Spam 2.0 from being disseminated on the Internet.

Knowledge can be divided into two categories, which are perceived knowledge and actual knowledge. Perceived knowledge or subjective knowledge is defined as “*what we think we know*,” while actual knowledge or objective knowledge is defined as “*what we actually know*” (Flynn and Goldsmith 1999). As mentioned in Flynn and Goldsmith (1999), objective knowledge according to Brucks (1985) is an actual knowledge construct as measured by some sort of test. In this thesis, the actual knowledge of Spam 2.0 is defined as the extent to which an individual is able to recognize general factual information about Spam 2.0. On the other hand, the perceived knowledge of Spam 2.0 is defined as an individual’s subjective assessment of their knowledge on Spam 2.0.

2.3.2.1 Perceived knowledge

In a survey assessing perceived knowledge, it is common to directly ask respondents to rate their level of knowledge on a particular issue. In a research by Levine and Donitsa-Schmidt (1998), a survey was carried out on 309 students to examine the relationship between computer experience, computer-related attitudes, computer-related confidence and perceived computer-based knowledge. In this research, perceived computer knowledge was assessed through 11 questions (Levine and Donitsa-Schmidt 1998). In all these 11 questions, respondents were asked to rate their level of knowledge based on a 5-point Likert scale ranging from “*none*” to “*very high*.” Some questions that were asked to be rated were “*understanding computer terminology*,” “*using utilities software*” and “*using Internet and e-mail*.”

A survey with a sample of 330 older adults was conducted on perceived knowledge, in a research study by Ellis and Allaire (1999) to assess the relationship between age, education, computer knowledge and computer anxiety with computer interest among older adults. In this research, computer knowledge was assessed by two 5-point Likert-type scale items ranging from no knowledge to expert level (Ellis and Allaire 1999). The first question asked to the participants was to “*select the level of computer knowledge that would best describe them*.” For the second question, respondents were asked to rate their level of computer knowledge compared to their age-group peers.

2.3.2.2 Actual knowledge

Although it is not mentioned clearly, actual knowledge is referred to when assessing public knowledge. To the best of the author's knowledge, there is no study in computer security that was carried out specifically to assess the public's knowledge. Nonetheless, Spam 2.0 could be considered as a threat to the online community. Similarly, any health disease would be considered as a threat or problem for the community. For this reason, this section presents studies on actual knowledge from the health field. In public health research, knowledge studies are commonly carried out based on a certain existing knowledge that contains symptoms, causes, suitable solutions or related information about that particular issue to evaluate the actual knowledge.

The issues of knowledge about acquired immunodeficiency syndrome (AIDS) and its modes of transmission were highlighted in a survey on Latino adults in northern California (Urizar Jr and Winkleby 2003). The sample size used was 461 women and 356 men from the community and another 188 men from a labour camp. Knowledge about AIDS and its modes of transmission was assessed through eight questions with 1 point given for each right answer resulting in scores varying from 0 to 8. These questions focused on the causes of AIDS including biomedical transmission and casual transmission. Eight (8) statements were given with responses using a 4-point Likert-type scale such as *definitely true – probably true – probably false – definitely false* and *very much at risk – somewhat at risk – very much at risk – not at risk*.

A survey was carried out on the issue of awareness and knowledge on methicillin-resistant *Staphylococcus aureus* (MRSA) with 545 respondents (McLaughlin et al. 2008). Further knowledge questions were then only focused on those who had heard of MRSA, leaving 521 respondents including 345 members of the public and 176 hospital visitors. Twenty-four out of 345 public members as well as 2 hospital visitors had a personal history of MRSA. Hence, this survey grouped the respondents into three groups consisting of 321 public, 174 visitors and 26 with a history of MRSA. In this research, the respondent's knowledge on MRSA was assessed through a basic open-ended question, "*What do you understand by MRSA?*" In addition, knowledge on the treatment of MRSA was asked based on three questions. Responses were rated on a 5-point Likert-type scale ranging from "*strongly disagree*" to "*strongly agree*." The result for this section was compared between the three groups that were identified earlier.

A pilot study on the awareness and knowledge of andropause was carried out in Hong Kong (Yan 2009). Data from 500 Chinese males aged more than 40 were gathered via face-to-face interviews. In this research, knowledge of andropause was assessed through questions on symptoms and treatment. The researchers identified those who were enlightened about the andropause issue by asking if they had heard of andropause. To the andropause-enlightened respondents, 12 symptoms of andropause are

listed, to which they had to answer either “yes” or “no.” The total individual knowledge score could range from 0 to 12. These symptoms were gathered from several existing studies. Moreover, the participants’ knowledge was assessed based on four questions on the treatment of andropause with “yes” and “no” answers as their options.

The issue of awareness was also studied in the research of awareness, knowledge and self-reported test rates on hepatitis B (Veen et al. 2010). Data were collected based on a survey involving 355 Turks in the Netherlands (Turkish-Dutch). The questionnaire used in this research was first developed through group focus discussions. Knowledge of hepatitis B in this research was evaluated using 10 statements. Respondents were allowed to choose either “true,” “not true” or “don’t know.” Knowledge was scored 1 for each correct answer and 0 for each incorrect answer; hence, the maximum score was 10 representing the highest knowledge score and 0 representing the lowest knowledge score. Six of the questions on knowledge were derived from an existing study which relates to the transmission and consequences of hepatitis B. Two questions on prevention were developed through the focus group discussion, while two other questions were on factual knowledge regarding confusion between hepatitis A and hepatitis B. The overall knowledge was also analysed based on two groups, those with low knowledge having an individual score of 0–5 and high knowledge for those who obtained an individual score of more than 5.

2.3.2.3 Perceived knowledge and actual knowledge

There are several studies which include or compare perceived and actual knowledge. In one of these studies, 16,677 students in grades 7–12 in US secondary schools were interviewed about their knowledge on correct condom use (Crosby 2001). In this study, three statements related to correct condom use were asked to participants with the options to answer either “true,” “false” or “don’t know” to assess actual knowledge. Participants were also asked to rate their perceived knowledge by the questions “*You are quite knowledgeable about how to use a condom correctly*” using a 5-point Likert scale ranging from “*strongly agree*” to “*strongly disagree*.” Results gathered from these two variables were then compared to see if there was any misconception about the participants’ knowledge.

In another research focusing on prostate cancer, a 31-item survey questionnaire was carried out on 108 African Americans (Agho 2001). The survey’s objectives were to measure both actual and perceived knowledge of prostate cancer, use of prostate screening service and demographic characteristics. In this particular survey, the perceived knowledge of prostate cancer was assessed based on statements such as “*I am very knowledgeable about prostate cancer*” and “*I am generally aware of the symptoms of prostate cancer*” to which they had to answer either “true” or “false.” To assess the actual knowledge, 21 statements were given to which respondents had to choose either

“true” or “false.” These statements included questions on causes, factors, preventive measures and treatment of prostate cancer and other related facts. The correlation between actual and perceived knowledge of prostate cancer was also examined in this research..

In a survey on teachers’ knowledge about epilepsy and attitudes towards students with epilepsy involving 512 elementary and middle school teachers in the United States, seven questions on knowledge and four questions on both knowledge and attitude were taken from an existing scale. These questions were actual knowledge and scored in the range of -3 (*I disagree very much*) to 3 (*I agree very much*). In addition, the researchers also added another 12 attitude and knowledge questions with similar scoring systems. Perceived knowledge was assessed in this survey with regard to the teachers’ general knowledge of the conditions and life circumstances of persons with epilepsy. Respondents were asked to rate on a 6-point scale ranging from “no knowledge” to “extensive knowledge”.

2.3.3 Perception

This section explains three aspects of perception which are the perception towards crime, perception towards crime’s punishment and fear of crime with the crime in this context specifically referring to Spam 2.0. The detailed elucidation was given in the book by Wood and Viki (2004); however, the examples given in that book were only for general crimes. Still, these aspects were also being studied in several studies in the computer security field such as research work by Dowland et. al (1999) and Al-Alawi and Abdelgadir (2006), although they were not mentioned clearly. The thesis presents related research work which will help build similar questions focusing on Spam 2.0 in order to determine how seriously Spam 2.0 is being seen by the public.

2.3.3.1 Perception towards crime

The perception towards crime could indicate how serious a crime is being viewed by the public. In addition, the relationship between perception towards crime and other attributes such as socio-demographic, victimization and fear of crime has been often evaluated in existing general crime-related research (Wood and Viki 2004), whilst, for crime involving computers or the Internet, the perception towards crime has been explored in several studies such as Dowland et al. (1999) and Harris (2000).

In a survey consisting of 53 main questions on 175 respondents in the UK to assess public attitude and awareness, the respondents’ perception on computer crime and abuse was also being evaluated (Dowland et al. 1999). One of the questions that were asked in relation to the respondent’s perception on computer crime and abuse was if the respondent felt that it was a problem. To this question, over 80% felt that computer crime and abuse was a problem. Respondents’ perception was also asked in

detail to the question where they need to assess the seriousness of potential abuse scenarios based on a scale of 5, from very serious to no crime. Scenarios that were listed in this research were viruses, viewing someone else's data, altering someone else's data, theft of computer equipment, unauthorized copying of software, unauthorized copying of data, computer fraud and sabotage. According to the authors, "*computer hackers represent the most 'hyped' forms of abuse in the mass media*"; hence, respondents' perception towards hackers was also asked. A further question asked was whether respondents consider hacking as acceptable, with 29% of respondents feeling that hacking is tolerable. Moreover, motivations for hacking such as out of curiosity, to make money, for the thrill of it, to beat the system and for malicious reasons were also asked with "yes," "no" and "don't know" answers as their options.

Based on the Dowland et al. (2000) work, another research was carried out by Al-Alawi and Abdelgadir (2006) to compare attitudes and opinion on computer crimes between UK and the Kingdom of Bahrain. The objective of this study is to test the hypothesis if the perceived level of safety is a factor in the willingness of the public to conduct online transactions. The questionnaire was developed similar to that in the research by Dowland et al. (2000) except the focus of computer crimes in this research is the crime of unauthorized copying of software or software piracy. For this reason, they also included copyright laws in the Kingdom of Bahrain. Based on the data collected from 500 respondents, it was found that there is a positive relationship between the perceived level of safety and the willingness of the public to conduct online transactions.

In another survey conducted by Harris (2000) to compare information security ethics between college students, data were collected from 712 college students based on 16 crime scenarios. In this survey, instead of directly asking respondents' perception about certain crimes, respondents need to choose their answer based on 5-point scales which are ethical, acceptable, questionable, unethical and computer crime. Some of the crime scenarios that were given in this survey were selling of shareware by the individual, changing of data that others used, changing of data to avoid payment of dollars, failure of reporting an error in a program, copying software for backup only and giving an old version of a program to someone else when the person has received the new version. The purpose of this survey was to test three hypotheses involving between sensitivity of ethics and academic levels, gender and subject as a part scenario. Two main results reported were as follows: (1) there is a difference in attitudes as students mature through the educational process in 12 out of 20 individual situations and (2) there is a difference in attitudes between genders in 8 out of the 20 individual situations.

2.3.3.2 Perception towards crime's punishment

Perception towards crime's punishment similarly could indicate how serious a criminal is being viewed by the public and how retributive people are towards a criminal for doing a certain crime. It is common to find out the relationship between the perception towards crime's punishment and other variables such as socio-demographic factors including gender and racial differences in general crimes (Wood and Viki 2004). Such research works however yielded both negative and positive results (Wood and Viki 2004). Nevertheless, in the case of crimes involving computers, there were not many research works that focused on this issue.

Focusing on general crime, in the research on the public perception of sentencing in Perth, Western Australia, 554 residents were interviewed from a sample frame of 800 (Indermaur 1987). The study objectives are twofold which are to measure public perception of the incidence of crime and punishment and public attitudes towards sentencing and simultaneously comparing sentences nominated after two crime presentations. To assess public perception towards crime's punishment, respondents were asked the question, "*Would you say the sentences handed down by the courts are too severe, about right or not severe enough?*" Respondents were asked to choose the answers based on a 5-point Likert scale: either "*too severe,*" "*some are too severe,*" "*some not enough,*" "*not severe enough*" and "*don't know.*"

For crimes involving computers, Dowland et al. (2000) in their survey involving 175 respondents in the UK that was carried out to assess public attitudes and awareness on computer crime and abuse have included several questions to assess public perception towards computer crime and abuse, specifically hacking. Respondents were asked if confessed or convicted hackers should be allowed to work in the computing field. The result was 59% stating that they should, and only 25% responding that they should not. Respondents were also asked if hackers should be allowed to have a computer at home with 59% stating that they had no problem with it and only 23% stating that they were against it.

2.3.3.3 Fear of crime

In relation to general crime, it was shown that a fear of crime causes a negative impact on individual behaviour and quality of life (Wood and Viki 2004). Other variables investigated in relation to fear of crime included attitudes to crime and punishment. While a certain level of trust is needed for all the users to participate in the Internet community, the fear of crime involving the Internet could cause them to distrust the services provided. Hence, it might affect how users interact or do business with each other on the Internet.

In one of the studies on general crime, 554 Perth residents were interviewed from a sample frame of 800 to measure the public perception of the incidence of crime and punishment and public attitudes towards sentencing (Indermaur 1987). In this research, the issue of fear of crime was assessed by asking two questions which were “*How safe do you feel walking alone at night in your neighbourhood?*” and “*How safe do you feel walking alone at night in Perth City?*” Responses were rated on a 4-point Likert-type scale containing the options “*very safe,*” “*safe,*” “*unsafe*” and “*very unsafe.*”

In the research on awareness, information sharing and privacy on FB, Acquisti and Gross (2006) carried out a comparison study between data collected from a survey involving 294 respondents and data from 7,000 profiles mined from FB. This research also included questions indicating the fear of crime relating to their studies. Thus, questions asked to the respondents were “*Specifically, how worried would you be if a [certain scenario took place]*” to indirectly assess respondents’ concerns on the issues being studied. Scenarios given in these questions included the state of the economy, threats to personal privacy, the threat of terrorism, the risk of climate change and global warming. Respondents were asked to choose their answer based on a 7-point Likert scale of how worried they were. The highest concern results were recorded for the statements that referred to threats to personal privacy which is “*A stranger knew where you live and the location and schedule of the classes you take*” followed by the statement “*Five years from now, complete strangers would be able to find out easily your sexual orientation, the name of your current partner, and you current political view.*”

Taking into consideration the fear of dangerous or harmful issues, the research by Al-Alawi and Abdelgadir (2006) did not assess the fear of a particular issue directly. Instead, they evaluated if the perceived level of safety affects the respondents’ willingness to conduct online transactions (Al-Alawi and Abdelgadir 2006). The perceived level of safety in this research was determined by questions that were associated to giving their personal information over the Internet. Thus, in this case, the dangerous or harmful issue that was the focus of this research was to give out information over the Internet. It was reported in this research that the perceived level of safety is indeed a factor of willingness to conduct online transactions.

Fear of dangerous or harmful issues was also commonly assessed in the health field to evaluate how hazardous a disease is as seen by the public. Similarly, fear of a certain disease is viewed to how vulnerable and at what risk the respondents think they are towards a certain issue. Hence, in health research, variables such as perceived risk and perceived vulnerability are commonly being evaluated for this matter such as research on human immunodeficiency virus (HIV) (Gerrard, Gibbons, and Bushman 1996), AIDS (Urizar Jr and Winkleby 2003) and alcohol-related harm (Wild et al. 2001).

A literature review was compiled by Gerrard et al. (1996) to find the relationship between perceived vulnerability to HIV and precautionary sexual behaviour. The methods used in this research were both

quantitative and qualitative. A detailed analysis on both strength and flaws of the concept and methodology used in the existing studies was reported in this paper. As stated in this paper, all related studies used similar questions such as “*What is the likelihood that you will contract HIV*” or “*What is the likelihood that you will develop AIDS?*” to assess perceived vulnerability to HIV, and these questions were commonly rated by 5-point scales (Gerrard, Gibbons, and Bushman 1996).

A survey on knowledge about AIDS and its modes of transmission focusing on Latino adults from northern California was carried out in 2000 (Urizar Jr and Winkleby 2003). The sample size comes from 461 women and 356 men from the community and 188 men from a labour camp. The results showed that most of the respondents see AIDS as a serious community problem. Nonetheless, the perceived extent of the AIDS problem was assessed through one question which is “*How much is AIDS a problem for Latinos in your community?*” Participants answered this question using a 4-point Likert-type scale of “*A lot,*” “*Some,*” “*A little*” and “*Not at all.*”

In the research on alcohol-related harm in 2001, the authors assessed the relationship between perceived vulnerability and alcohol-related harm via a survey administered to 286 university students (Wild et al. 2001). It was shown that there is a positive relationship between drinking problem and perceived risk of experiencing harm (Wild et al. 2001). This research used two questions asking respondents on perceived vulnerability to alcohol-related harms which were (1) “*To what extent do you believe that you would be personally at risk of getting hurt or getting sick because of your drinking?*” and (2) “*To what extent do you believe that some other person your age who drinks the way you do would be at risk of getting hurt or getting sick?*” The responses were rated on a 7-point Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree).

2.4 Open Issues

The outcome of this section will enable us to focus on the unsolved and arising issues related to the cost of spam. From the studies presented on cost models in Section 2.2, the open issues to be solved in the issues for cost models for spam are listed below:

- 1) Overlapping of the definition on cost categories.
- 2) Unspecified cost models and unavailability of cost models publicly.
- 3) Dependency of the cost parameters on external data, whereas in certain cases, external data might not portray the latest value.
- 4) Inability to measure storage cost without having an internal system. Total dependency on internal data and these data are not publicly available.

- 5) There were no real/empirical data; hence, these cost models are using either average-based or predefined values.
- 6) Some of the cost models provide quantification; however, there was no further work to develop a cost model based on those quantifications.
- 7) Some of the cost models do not cover certain cost categories.
- 8) Lack of information on how to generate values for certain cost parameters in some cost models.
- 9) Some of the cost models could not be fitted to estimate the cost of Spam 2.0.
- 10) Inexistence/lack of studies on Spam 2.0 cost. Inability to identify the cost of Spam 2.0. Existing cost models' focuses are not to estimate the cost of spam. Unlike the email spam where the cost of spam is borne by the organization, for Spam 2.0, the cost of Spam 2.0 is unknown while the cost borne by stakeholders is still unknown.
- 11) Lack of focus of cost models on certain stakeholder.
- 12) Company-based studies.
- 13) Existing cost models specifically focused on two countries.
- 14) Due to the growth of Spam 2.0, the inexistence of studies on the cost of Spam 2.0 makes it hard to quantify the exact cost of Spam 2.0.
- 15) Inability to measure the actual time wasted on Spam 2.0.

While carrying out the studies on the cost of Spam 2.0, it was also found out that the cost of Spam 2.0 could be easily propagated on the Internet if the users lack awareness and knowledge of Spam 2.0. Henceforth, through studies presented in Section 2.3, the thesis presents the open issues on the awareness, knowledge and perception of Spam 2.0 as listed below:

- 1) Lack of awareness, knowledge and perception studies on security-related issues.
- 2) Inexistence of studies on public knowledge on Spam 2.0/the current Spam 2.0 public awareness situation is unknown.
- 3) Inexistence of studies on public knowledge on Spam 2.0/the current Spam 2.0 public knowledge situation is unknown.
- 4) Inexistence of studies on perception on Spam 2.0/the current Spam 2.0 public perception situation is unknown.

2.5 Chapter Summary

The focus of this chapter was to provide the background on cost model, public awareness, knowledge and perception of Spam 2.0. The first part of the chapter discusses the existing cost model study. A detailed overview and evaluation on email spam, web server and data centre cost models is presented. A comprehensive survey of cost models that have been examined in this chapter is one of the main contributions of this study to highlight the unresolved issues in the domain of email spam cost models and to simultaneously capture the knowledge of cost model development. In the second part of the chapter, a literature review on the public awareness, knowledge and perception of Spam 2.0 is provided. A few potential studies underlying the theoretical part of awareness, knowledge and perception are then outlined. Finally, all unresolved open issues regarding cost, public awareness, knowledge and perception of Spam 2.0 are listed.

Chapter 3

Problem Definition

This chapter covers:

- ▶ A formal definition for cost, awareness, knowledge and perception of Spam 2.0;
- ▶ Problems associated with Spam 2.0 cost, awareness, knowledge and perception;
- ▶ The research issues that need to be addressed; and
- ▶ The research methodology that is adopted in this research to systematically address the identified research issues.

3.1 Introduction

Chapter 1 presented a thorough overview of Spam 2.0, cost models and its function and other several relevant issues: specifically, public awareness, knowledge and perception to establish the prominence of this thesis. A detailed summary of all relevant topics were presented in Chapter 2. Chapter 2 also presents the previous literature on public awareness, knowledge and perception on various topics. Finally, a list of open issues on cost, public awareness, knowledge and perception of Spam 2.0 is presented in Section 2.4.

Based on the open issues listed previously, Chapter 3 clearly outlines the main problems to be solved and identifies the formal description of research issues as the basis for the development of subsequent chapters. This chapter then continues to provide research methodologies used in our research. The science and engineering approach is adopted to study the cost of Spam 2.0 which is explained in Section 3.4. On the other hand, in order to study public awareness, knowledge and perception on Spam 2.0, a quantitative approach is adopted which is further explained in Section 3.5. The conclusion for this chapter is then presented in Section 3.6.

3.2 Problem Definition

The thesis defines the problems identified from a comprehensive review on existing studies in Chapter 2 according to two categories, which are (1) cost and (2) awareness, knowledge and perception.

3.2.1 Cost

This thesis mainly focused on the Spam 2.0 cost model. The need to develop a cost model is growing with the increasing rate of Spam 2.0 propagation. To the best of our knowledge, there has been no research that produces a clear amount of the costs for Spam 2.0. Having a cost model will help the parties involved to measure the consequences caused by Spam 2.0 and further assist them to propose better solutions.

As discussed in Chapter 2, the existing cost models are mostly focused on email spam while most research on Spam 2.0 is focused on the method of prevention or detection. Hence, there was almost no research that specifically caters to Spam 2.0 cost model. Obtaining information on cost categories and cost parameters that are related to Spam 2.0 can only be done through extensive research on both topics and the author's knowledge.

At the earliest stage of the research, it was also found out through extensive review on current literature that Spam 2.0 data are unavailable since it is a new area of research. Although there are several publicly available spam data, they were only focused on email spam. As discussed in Chapter 1, there were a lot of different characteristics between email spam and Spam 2.0 that may have resulted in different values. Hence, obtaining the Spam 2.0 data will be crucial in ensuring that the research will be able to produce a cost model fit for Spam 2.0.

Through a detailed review on methodologies used in the email spam cost model in Chapter 1, to develop a cost model, there is a need to obtain data through surveys or interviews, or through laboratory experiments, or the existing spam repository. There was no existing Spam 2.0 repository and it is our intention not to determine the cost model through people's opinion. Hence, it is important to develop an internal system in order to measure related cost categories defined for the cost model.

Through detailed definition of cost categories and cost parameters that are involved in the Spam 2.0 cost model (which are included in Chapter 5), it was also found that a survey is usually used to estimate the actual time wasted on email spam. Thus, the resulting value is an average of the user's estimation. In order to estimate the cost of Spam 2.0, the error of the user's estimation has to be reduced; hence, it is important to develop a way to measure the actual time wasted on Spam 2.0 automatically.

It was also found out that there is no specific way to evaluate the existing cost models since most of the studies only produced cost models and cost values based on their data sets. While there are plenty of email spam cost models that can be compared to each other, Spam 2.0 cost models on the other hand is a relatively new research. Moreover, since there was no existing spam repository found on

Spam 2.0, the evaluation of the cost model based on external input was impossible. Hence, the evaluation stage of the Spam 2.0 cost model will be a crucial process.

As discussed in Chapter 2, it was also rare to find survey studies on security-related issues. Most of the research on public awareness, knowledge and perception are more focused on social science. Thus, it will be a difficult challenge to adapt existing studies to the topic of Spam 2.0.

3.2.2 Awareness, Knowledge and Perception

Apart from focusing on the cost of Spam 2.0, this thesis also plans to investigate on the awareness, knowledge and perception of Spam 2.0. The thesis carried out an in-depth review on related literature on existing studies focused on awareness, knowledge and perception. There were an extensive amount of researches that were focused on awareness, knowledge and perception. For these studies, it was found out that knowledge is usually categorized into two categories, which are perceived knowledge and actual knowledge, while when discussing about perception, there are a few types of perceptions typically considered in those studies such as perception towards crime, perception towards crime's punishment and fear of crime. Unfortunately, such categories came from different fields. There were only a few that were related to computer security.

Existing research on Spam 2.0 were mostly focused on the method to combat Spam 2.0 and to stop Spam 2.0 propagation. Nonetheless, other aspects of Spam 2.0 such as public awareness, knowledge and perception that could possibly impact Spam 2.0 proliferation were not the focus in the existing research. Therefore, there were no validated research question items found on Spam 2.0 specifically. Hence, it will be crucial to prepare Spam 2.0-related questions on awareness, knowledge and perception.

3.3 Research Issues

Spam 2.0 is the latest new type of spam attacking web users. Instead of focusing on the spam-filtering problems that are usually studied by most researchers, this thesis focuses on Spam 2.0 cost. Moreover, public awareness, knowledge and perception on Spam 2.0 are also investigated. It is expected that awareness, knowledge and perception of Internet users can influence how they manage and combat Spam 2.0 which would then reduce the cost of Spam 2.0 management.

Although there were numerous studies covering some of the cost categories, through problems presented in Section 2.4, it was found that no single cost model is fitted for Spam 2.0 costs. Hence, the thesis selects and describes two research issues highlighted for this research.

3.3.1 Research Issue I: Developing Spam 2.0 Cost Model to Identify Related Costs.

There were several cost models created on email spam. However, there was none that is fitted for Spam 2.0 cost models. Furthermore, the increasing rate of Spam 2.0 ensures that there is a need for a cost model. The main problems in developing a cost model for Spam 2.0 are listed as follows:

- Unavailability of information on cost categories and cost parameters related to the Spam 2.0 cost model.
- Inexistence of Spam 2.0 data/unavailability of data to be used on developing the Spam 2.0 cost model.
- Inability to measure certain cost categories for Spam 2.0 without having an internal system (such as spam-filtering facilities, etc.) and survey.
- Inability to measure actual time wasted on Spam 2.0 automatically.
- No specific evaluation found on existing cost models.
- Inexistence of external data on Spam 2.0 to be used as cost model input.

The technical problems outlined above formed a number of research questions to be addressed in developing a cost model for Spam 2.0. Such research questions are as follows:

- RQ1: Can the research develop an internal system that can define all relevant cost categories and cost parameters for the Spam 2.0 cost model?
- RQ2: Can the research develop an internal system to produce enough Spam 2.0 data to be used for estimating the cost model?
- RQ3: Can the research measure the actual time wasted on Spam 2.0 without users having to estimate the value themselves?
- RQ4: How does the research evaluate the cost model?

All these questions are answered in Solution I written in Chapter 4 as presented in the theoretical framework.

3.3.2 Research Issue II: Insufficient Information and Exploration on Public Awareness, Knowledge and Perception Regarding the Topic of Spam 2.0.

Most of the current work on Spam 2.0 focuses on the technical part, which is the method of detection and prevention of Spam 2.0 (Stringhini, Kruegel, and Vigna 2010; Sureka 2011; Liu et al. 2008; Shin, Gupta, and Myers 2011; Markines, Cattuto, and Menczer 2009; Chu et al. 2010; Hayati, Potdar, Smyth, et al. 2010; Hayati, Potdar, Talevski, et al. 2010; Hayati, Potdar, Chai, et al. 2010). Filtering that implementing these detection and prevention techniques will impede Spam 2.0 from being proliferated to the user's system. However, there is no guarantee that these filters are flawless. If it manages to bypass the system, then this might pose a dangerous situation for unaware and unknowledgeable users. Furthermore, if the users who were attacked have a light perception on the effects of Spam 2.0, then users might fall for the trick and become one of the victims. It is observed that Spam 2.0 propagates because of users' lack of awareness, lack of knowledge and erroneous perceptions that influence how they handle Spam 2.0. Nonetheless, it is undeniable that the public's awareness, knowledge and perception could help to reduce Spam 2.0 propagation.

Still, there was insufficient information and exploration done on the awareness, knowledge and perception of public users on Spam 2.0. The thesis listed the difficulties in doing the research as follows:

- Inexistence of exact similar studies on Spam 2.0 to be adapted into the research.
- Inexistence of validated Spam 2.0-related questions on awareness, knowledge and perception.

Thus, the thesis defines research questions to be answered by the survey for awareness, knowledge and perception survey as follows:

- RQ6: To what extent are the public users aware of Spam 2.0?
- RQ7: To what extent is the knowledge of public users on Spam 2.0?
- RQ8: What is the perception of public users towards Spam 2.0?

All these questions are answered in Solution II written in Chapter 4. Based on the research issues and detailed research questions that were clearly stated in this subsection, the thesis now presents the methodologies that are adapted in this research. A methodology is the research process or philosophy used to interpret data and reach a conclusion. A multimethodological approach is needed to solve our problems. Two methodologies chosen to be adapted in this research are:

- 1) Design Science Methodology
- 2) Quantitative Methodology

Research Methodology I is further explained in Section 3.4 while Research Methodology II is further described in Section 3.5. Research Methodology I is associated to Solution I while Research Methodology II is associated to Solution II. Both Solution I and II are explained in Chapter 4.

3.4 Research Methodology I: Design Science

In order to conduct research, there is a need to identify the appropriate combination of processes, methods and tools that can be used (Nunamaker and Chen 1990) to be able to interpret data and formulate a conclusion. To address research issue I, the research adopts the design science methodology. A design science-based approach is commonly associated to the research that leads to the development of new techniques, architecture, methodologies, devices or a set of concepts, which can be combined to form a new theoretical framework. Common research processes involved in this methodology are problem definition and developing conceptual solution, implementation, experimentation, testing and validation of prototype using the appropriate criteria (March and Smith 1995; Hevner et al. 2004).

A design science based methodology comprises of three main stages:

- Problem definition
- Conceptual solution
- Implementation, testing and evaluation

3.4.1 Problem Definition

In this stage, real problems are clearly explained to highlight the significance of conducting the research. This stage involves the process of analysing, interpreting, discussing, and evaluating current problems based on specific measures and perspectives. Through the exploration of the specific domain and relevant literatures, this stage has been carried out and explained in Section 0

3.4.2 Conceptual Solution

The conceptual solution stage emphasizes on producing and applying knowledge in order to create effective technology-oriented solutions for the selected problems. In this stage, researchers are constructing ways of performing goal-directed activities which are done through the design and building of tools, an environment or system through implementation. Such a design involves the study

and an in-depth understanding of the domain, the applications of the issue's knowledge and experience to solve the problem and the creation and the evaluation of a proposed solution. Thus, a conceptual framework which is an abstract model of the practical solution is designed and functioned as a road map for the implementation process. A detailed framework for this research is provided in Section 4.2.

3.4.3 Implementation, Test and Evaluation

In this stage, the implementation and experimentation of the proposed solution are carried out. The processes involved will show how well the proposed solution performs. The feasibility, usability and functionality of the working system will be tested and validated hence providing both benefits and drawbacks of the whole solution. The analysis of the results provides an insight for the evaluation of the research outcomes.

3.5 Research Methodology II: Quantitative

In order to solve the identified problems related to public awareness, knowledge and perception as addressed in Section 3.2, this research adopts the quantitative research method by carrying out a survey. The first step in choosing the suitable research method is to reflect on the research questions. It may be recalled that the main idea is to describe the extent of awareness, knowledge and perception of Spam 2.0 and such questions using the keyword "extent" relate to quantity. As suggested by Davidsson (2004), "*questions that are inherently quantitative in nature are best answered by quantitative investigations.*" Thus, in order to examine the extent of awareness, knowledge and perception of Spam 2.0, it is best to use quantitative methodologies (Davidsson 2004).

This methodology consists of four stages:

- 1) Problem definition
- 2) Survey design
- 3) Data collection and distribution
- 4) Analysis and assessment

The thesis now explains in detail each of these components in the following subsections.

3.5.1 Problem Definition

Similar to the problem definition stage involved in design science methodology (refer to Section 3.4.1), problems are defined to show the importance of conducting a survey which are already stated in Section 3.2.

3.5.2 Survey Design

Based on the problems defined in the earlier stage, survey objectives are established followed by determining the sample and respondents targeted depending on the survey aims. In this stage, it is also important to choose the correct way of carrying out the survey. Finally, this stage involves the process of designing the survey question which will be able to provide information that is needed and further answer the survey objectives. The survey questions then will be run for pre-test to ensure that the questions asked are relevant and accurate. Finally, after the appropriate changes, the survey then can be conducted.

3.5.3 Data Collection and Distribution

In this stage, the survey can be distributed to the targeted respondents. This process will be carried out until it reaches the targeted sample size. In order to increase the response rate to reach the sample size, several practices have been suggested such as to give incentives, reminders or thank-you letters to the respondents.

3.5.4 Data Analysis and Assessment

In this stage, data that were collected from the survey are analysed and assessed. Based on the data, the output from this stage will provide the answer for the research objectives and henceforth, a conclusion can be drawn.

3.6 Conclusion

To link and simplify the methodologies used in this research, the thesis summarizes a multimethodological approach to our research as presented in Burstein and Gregor (1999) adapted from Nunamaker and Chen (1990) in Figure 3.1. In this figure, only the related approach that was taken to the research was included.

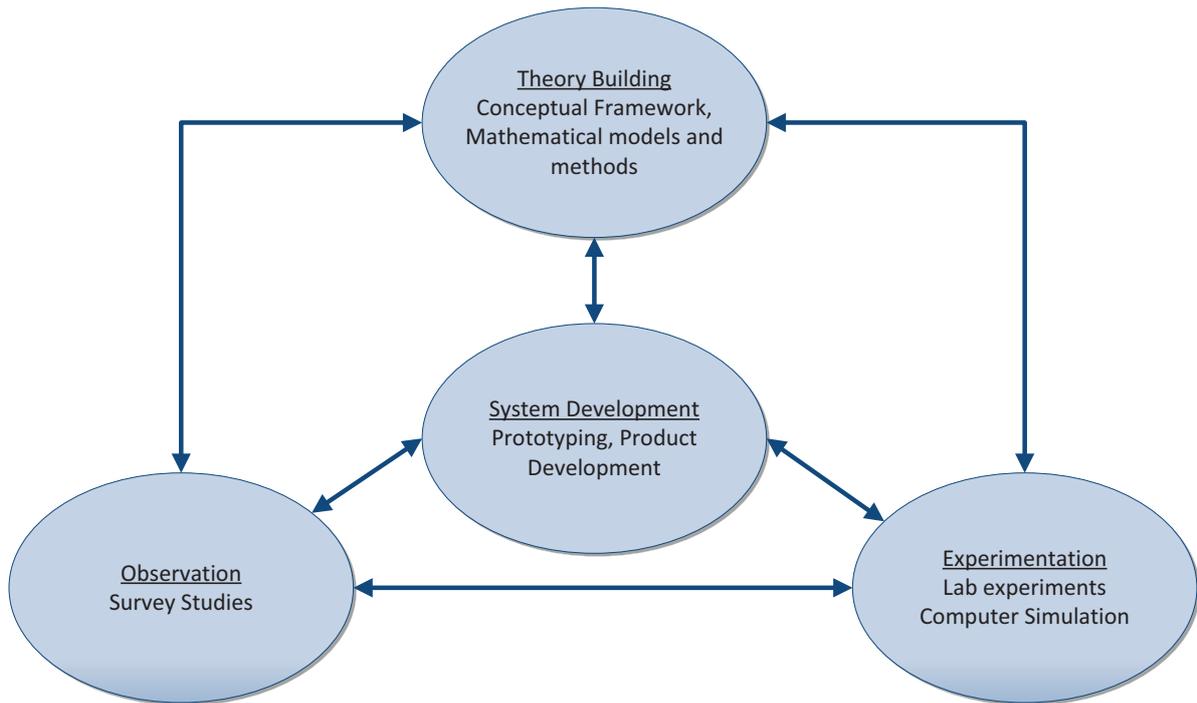


Figure 3.1: Multimethodological approach [Burstein and Gregor (1999) adapted from Nunamaker and Chen (1990)].

This chapter first defines the problems to be solved based on the literature review presented in Chapter 2. This includes the problems of the Spam 2.0 cost model and issues regarding public awareness, knowledge and perception on Spam 2.0. There were two research issues with eight research questions drawn from the problem definition which can be solved by adopting the design science and quantitative research methodology. In the next chapter, the outline of the proposed/designed solution in the conceptual framework is presented.

Chapter 4

Conceptual Solution

This chapter presents

- ▶ An outline of the proposed solutions related to the problems that are solved in this thesis;
- ▶ Conceptual framework of the proposed solutions; and
- ▶ Detailed conceptual processes adopted in the development of the proposed solutions.

4.1 Introduction

As outlined in Chapter 2, plenty of studies have been done on email spam cost models. However, it is evident that the area of the Spam 2.0 cost model is very new and up to date, since no cost models have been created for Spam 2.0. Continuing from a detailed review on current literature, Chapter 3 has defined problems and presented two research issues with eight research questions to be addressed in solving the problems. In the following sections, the thesis provides an overview of the solutions to each of the research issues discussed in Section 3.3. Section 4.3 provides detailed descriptions for each solution. Section 4.4 presents a summary of this chapter followed by a conclusion in Section 4.5.

4.2 Overview of the Solution

The problems in developing the Spam 2.0 cost model and how they relate to our survey study have been clearly defined in Chapter 3. The thesis has listed two research issues as follows:

- Developing the Spam 2.0 cost model
- Insufficient information and exploration of the public awareness, knowledge and perception on the topic of Spam 2.0

Section 3.3.1 identifies the questions that need to be addressed in order to solve Research Issue I which is to develop the Spam 2.0 cost model. Thus, there is a need to develop a solution which has the following features:

- Ability to define a solution that will be able to define all relevant cost categories and cost parameters for the Spam 2.0 cost model.
- Ability to collect related Spam 2.0 data to be analysed in developing the cost model.

- Ability to develop an internal system that can measure targeted attributes for Spam 2.0.
- Ability to develop a solution to measure the actual time wasted on Spam 2.0 automatically.
- Ability to evaluate the cost model.

All of these features are presented in Solution I. The idea is to develop a workable internal system that will be able to portray the real-world situation of Spam 2.0. This internal system will be able to provide input for the related cost parameters and cost categories. Meanwhile, for Research Issue II, the solution will be a survey. This survey will have to provide these features:

- Ability to provide questions that will be able to measure public users' awareness on Spam 2.0.
- Ability to provide questions that will be able to measure public users' knowledge on Spam 2.0.
- Ability to provide questions that will be able to measure public users' perception on Spam 2.0.

All of these features are explained in Solution II. A detailed explanation of the processes involved in each solution is presented in Section 4.3. The thesis first provides an overview of the conceptual solution used in this research as in Figure 4.1:

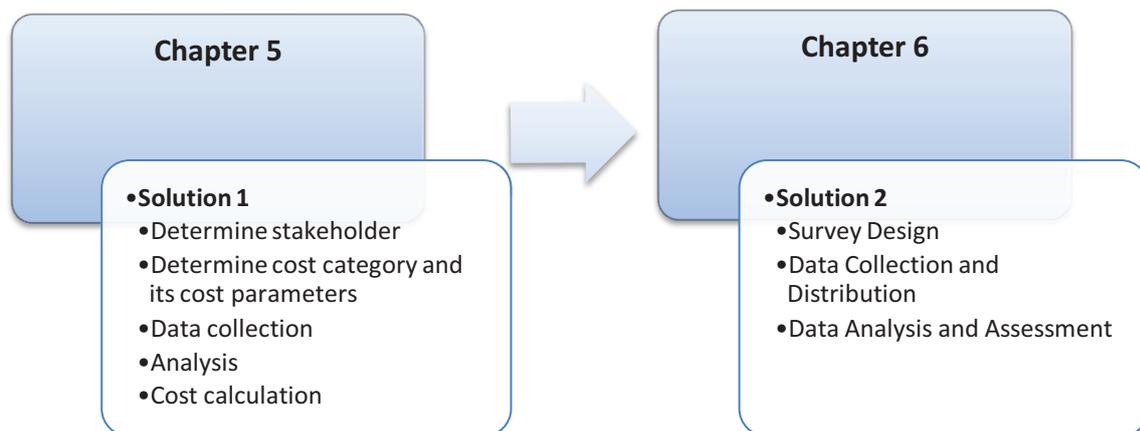


Figure 4.1: Overview of the conceptual solution

Figure 4.1: shows how the dissertation is organized in this chapter. The conceptual solution is proposed to address the two main problems outlined earlier. In the next section, the details of each proposed solution are discussed.

4.3 Solution Description

Earlier sections have presented an overview of the related solutions which are constructed to address the issues defined earlier. The thesis now explains the solutions used in this research. To strengthen the description of the conceptual solution, examples and implemented components are clearly explained.

4.3.1 Solution I

This research aims to develop a Spam 2.0 cost model. This section explains the processes taken to fulfil the objective. The thesis lists the processes involved adopted from the email cost model methodology as discussed in Section 0:

- 1) Determine stakeholder
- 2) Determine cost category and its cost parameters
- 3) Data collection
- 4) Analysis
- 5) Cost calculation

4.3.1.1 Determine stakeholder

As defined in Section 2.1, the stakeholder can be either an individual, organization, country or worldwide. A stakeholder that is involved can also be divided into spammer and non-spammer. As explained in the ‘Evaluation on stakeholders’ in Section 2.2.1.6, it is common practice to take an employee in an organization as the basis and work the calculation out to a higher level, which is a country.

4.3.1.2 Determine cost category and its cost parameters

Based on the evaluation on email spam cost models in and other related spam cost models, the thesis now defines cost categories and cost parameters associated with stakeholders for Spam 2.0 as summarized in Table 4.1.

Table 4.1: Summary of proposed Spam 2.0 cost model.

	Stakeholder	Cost category	Cost parameters
1	Non-Spammer	Storage cost, <i>SC</i>	<ul style="list-style-type: none"> • Average size Spam 2.0 post • Total number of spam • Cost of 1MB archive storage • Size of storage for all spam units in a month • Number of months • Storage cost per month per GB • Related additional cost per month (space cost, power cost, cooling cost and labour for server maintenance)
		Loss of Productivity, <i>LoP</i>	<ul style="list-style-type: none"> • Average annual salary per individual • Time wasted on Spam 2.0 in seconds (identify Spam 2.0, resolve Spam 2.0 problem and delete/remove) • Average Spam 2.0 received everyday • Total working hours • Number of months
		Labour Cost, <i>LC</i>	<ul style="list-style-type: none"> • Percentage support time specifically used for Spam 2.0 work (installing spam-related filtering software, effort in manually handling any problems arising from Spam 2.0) • Web administrator and maintenance work • Hourly labour cost • Weeks worked per year • Annual system administrator salary per year
		Connectivity cost, <i>CS</i>	<ul style="list-style-type: none"> • Cost of Internet connectivity per month • Percentage of bandwidth used by Spam 2.0 activities • Connectivity set-up cost
		Software cost, <i>SoC</i>	<ul style="list-style-type: none"> • Purchase of anti-spam product's licence and support

	Stakeholder	Cost category	Cost parameters
2	Spammer	Spammer's profit, <i>SP</i>	<ul style="list-style-type: none"> • Response rate • Profit per item • Number of Spam 2.0 posted every year
		Operating cost, <i>OC</i>	<ul style="list-style-type: none"> • Connectivity/bandwidth - sum of Internet service cost • Percentage of bandwidth used for spamming activities • Electricity cost • Address collection cost (bought) • Open proxy cost • Address collection cost (self-collected)
		Hardware cost, <i>HC</i>	<ul style="list-style-type: none"> • Cost for a computer • Cost for a monitor
		Software cost, <i>SoC</i>	<ul style="list-style-type: none"> • Cost of operating system • Cost for web hosting • Cost for spamming-related software

Thus, two types of total cost for spammer and non spammer are defined below:

- (i) $Total\ cost_{spammer} = Operating\ cost + Hardware\ cost + Software\ cost$
- (ii) $Total\ cost_{non\ spammer} = Storage\ Cost + Loss\ of\ productivity + Labour\ cost + Connectivity\ cost + Software\ cost$

Nonetheless, due to time constraints and limited resources, this research only focuses on the total cost for non-spammers as it will be the cost of Spam 2.0 towards the non-beneficial stakeholder. As mentioned earlier, identifying the total cost for non-spammers is a challenging task in the context of Spam 2.0 and this is the first work done in this area. Therefore, although the equations look vague and indeterminate, they were necessary to use in the light of available data to compute the costs.

4.3.1.3 Data collection

In order to generate values for cost parameters defined above, there is a need to collect data from the real world. Data collected in this solution come from two sources: HoneySpam 2.0 and a survey.

HoneySpam 2.0

In order to resolve Research Issue I, the research needs to set up an experiment which will imitate the real environment. Instead of downloading a real forum and classifying the content whether they are spam or non-spam, it was decided to construct a data set by creating a honeypot. This honeypot is called HoneySpam 2.0 following the name of Spam 2.0 (Hayati et al. 2009).

HoneySpam 2.0 was a set-up in an online discussion forum using Simple Machines Forum (SMF), an open source discussion forum. This forum was then advertised by listing the URLs in Pligg sites. HoneySpam 2.0 is designed based on the idea of a honeypot where a vulnerable server can be attacked by spammers (web spambots or human spammers). This honeypot was initially designed to capture Spam 2.0 content and to study spambot behaviour (Hayati et al. 2009). In this research, the content posted by these spammers in the forum is used as an input for the cost model.

The forum was set up similar to a normal forum where the users need to first register and login. The first phase of the data collection was from June 2010 to November 2010 and the second data phase was from February 2011 to June 2011. Since April 2011 onwards, CAPTCHA was implemented for forum registration as the first layer of security. The forum goes online for the whole duration and continues collecting spam data except when server maintenance is done.

During the whole duration of data collection, a total of 62,798 spam profiles were created, 141 spam personal messages were sent between the forum users and 450,772 spam post messages were posted and there were no polls created in the forum.

Survey

This survey refers to Part C that is embedded in Solution 2. The whole component of the survey will be explained in detail in the next section (refer to Section 4.3.2 on page 125). The tool used for survey data collection is Qualtrics. The main idea behind this data collection is to allow us to assess the actual time taken to identify Spam 2.0 without asking the respondents directly, thus resolving Research Issue I for Research Question 3. Therefore, 10 questions were designed consisting of five spam and five non-spam examples, which will allow us to measure the time taken to identify Spam 2.0. These 10 questions are included in the third part of the survey. Users are asked to identify whether the Spam 2.0 examples presented in the questions are considered as spam or not. While answering the questions, the tool used in this survey measures the time taken for all respondents. Users then will be asked to provide a justification for their answer, although this is optional. Related components of the survey that help to resolve the issue are presented in this subsection, which are:

- 1) Timing function
- 2) Design of questions interface
- 3) Sources of Spam 2.0 examples
- 4) Spam 2.0 examples and justifications

Timing function

As the aim of this research is to estimate the time used in identifying spam, the timing function available in Qualtrics was used to record the number of clicks and time for page submission made by participants when answering the questions. It is assumed that participants have made the decision during the time recorded for page submission. Hence, the analysis will only be focused on this attribute and the value of recorded time will represent the attribute called time taken to identify Spam 2.0. The data obtained specifically involved with timing function from this survey are called timing data set.

Design of question interface

The interface of the questions was designed to ensure that the time recorded represent the actual time taken for respondents to identify Spam 2.0. Thus, the research tries to minimize the relative error caused by page loading by embedding a smaller size of the examples' pictures (30–100kB) so that it is quicker and easier to load. On a dial-up connection with slow speed (28–56kbps) (Savage and Waldman 2005), this will take a range of 8–27 seconds to load this (Numion). Therefore, with Internet speed using the current connection asymmetric digital subscriber line (ADSL) (1–8mbps) or ADSL2 (3.5–12mbps), a 100-kB picture will load in less than 1 second (Access Communications Pty Ltd 2013).

In order to minimize the time taken before users will be able to see spam examples provided on the screen, the questions were designed to fit the examples into one page so that the users do not have to scroll down therefore, little navigation and minimal clicks are required. A screenshot of the page is shown in Figure 4.2.

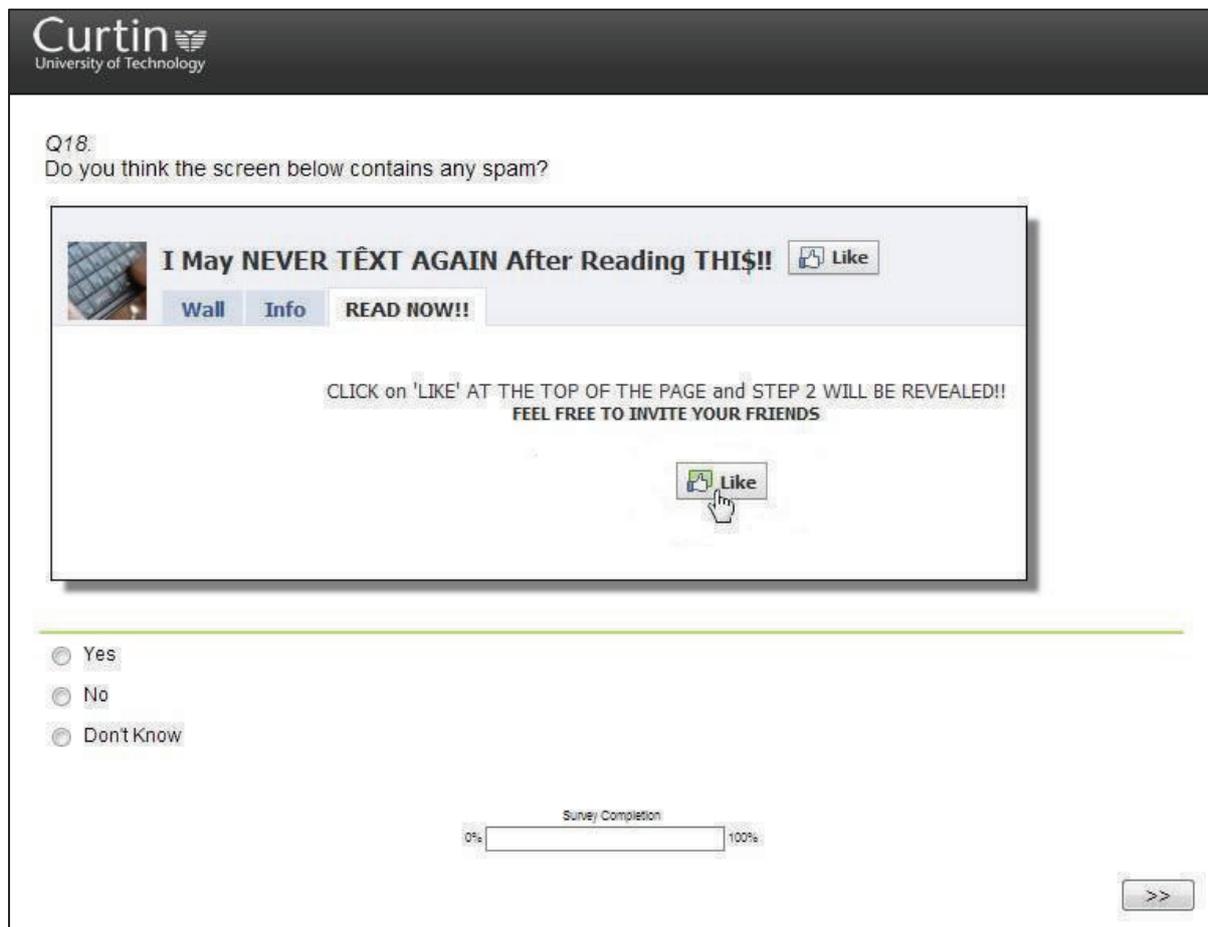


Figure 4.2: Screenshot of survey

Sources of Spam 2.0 examples

Spam 2.0 examples were collected from real-world samples from Yahoo Messenger (YM) which is an instant messaging services and FB, which is a social networking application. For YM examples, screen captures from the message, both online and offline, were used. For FB examples, screen captures from the profile, application and messages pages were used. Table 4.2 shows a summary of sources for these Spam 2.0 examples. Spam examples were collected from YM as it is an earlier web 2.0 application while FB is an example of the latest web 2.0 applications.

Spam 2.0 examples and justifications

The thesis provides the screen capture of Spam 2.0 examples used in this survey in Figure 4.3 to Figure 4.12. Each example is then categorized into either spam or non-spam followed by the

justification of the categorization. Recall again that the first five Spam 2.0 examples are from FB and the following five are from YM.

Table 4.2: Questions' category and source.

Question	Spam/Non-spam	YM/FB
1	Spam	FB
2	Spam	FB
3	Non-spam	FB
4	Non-spam	FB
5	Non-spam	FB
6	Non-spam	YM
7	Spam	YM
8	Spam	YM
9	Spam	YM
10	Non-spam	YM

Example 1

Figure 4.3 shows a screen capture on Example 1 which is a page of an FB profile. This profile was categorized as spam not only because this page promotes the “see who stalks your profile” application but also because of the nature of this application when it is being approved, i.e. an automatic link will be posted on the account owner’s wall. Links that promote the application above stated “*I can’t believe this works*” that were posted repetitively on this page are a strong indication that this application is a spam.



Figure 4.3: Screenshot of Question 1

Example 2

Figure 4.4 shows a screen capture on Example 2 which is an FB application page. Any application that require the account owner to like the page before being able to see the content is considered as

spam (Cluley 2012; Jeffries 2008). In addition, this page contains a dubious title with weird characters embedded in it that could cause suspicion in the users. Hence, it is decided that Example 2 is a spam.



Figure 4.4: Screenshot of Question 2

Example 3

Figure 4.5 shows a screen capture on Example 3 which is a page containing FB messages. This message is only given to users subscribed to this airline company. Since this user subscribed to this airline company, it comes from an authorized source. Moreover, this company is a valid airlines company in Malaysia. If the users stop subscribing to this airline company, then they would not receive such promotions any more. Users will be able to check the link provided if it is a real or fake link. While conducting the survey, the link which runs under a secure connection (https) still works fine. Thus, it is decided that Example 3 is a non-spam.

Example 4

Figure 4.6 shows a screen capture on Example 4 which is a page containing FB messages. This message is received from a personal contact (authorized) and the message is an invitation to a private event. There are two links attached to this message, and when checked, both of them are valid links from FB and the link is related to the text in this message. Hence, it is decided that this message is categorized as non-spam.

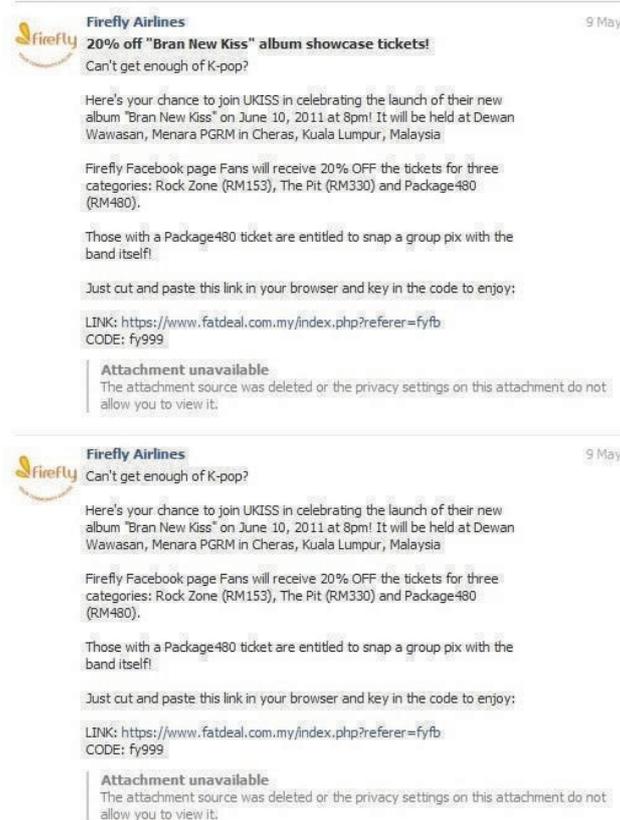


Figure 4.5: Screenshot of Question 3

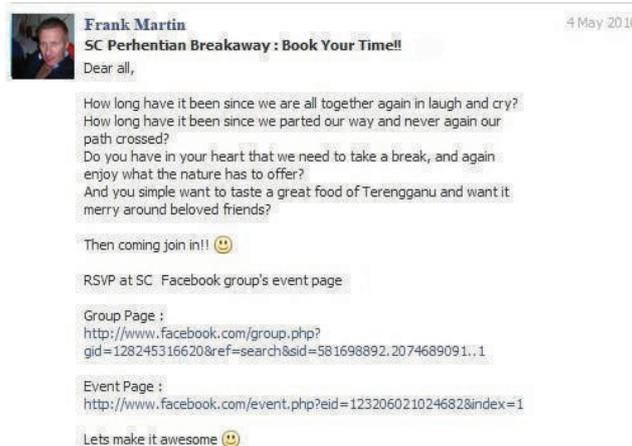


Figure 4.6: Screenshot of Question 4

Example 5

Figure 4.7 shows a screen capture on Example 5 which is a page containing FB messages from a personal contact. There were many links attached to this message. However, all the links were shared on a secure connection (https) and they contained a URL to Picasa, a valid image sharing application. When checking the links, users are brought to the photos in southern Africa, as mentioned in the message. This message was originally from a personal contact who frequently shares his travel

pictures. Since it is a regular activity posted from a personal contact, this example is considered as non-spam.

Example 6

Figure 4.8 shows a screen capture on Example 6 which is a YM online instant message from a personal contact. The screen capture shows a normal message; the status also indicates that it is a normal user and does not show any suspicious activity. It also has been validated that it really comes from a friend. Hence, this example is considered as non-spam.

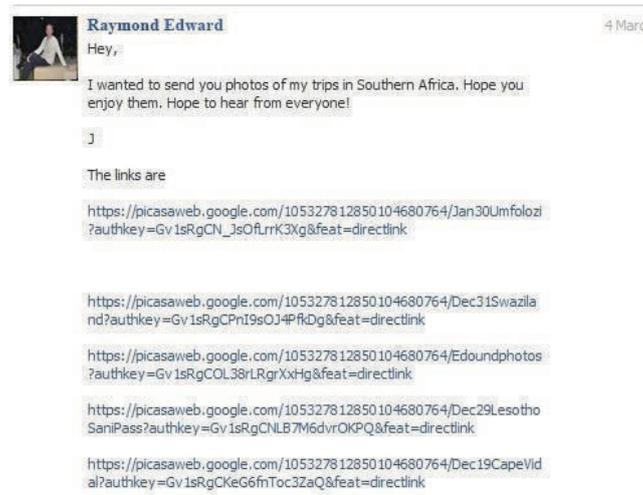


Figure 4.7: Screenshot of Question 5

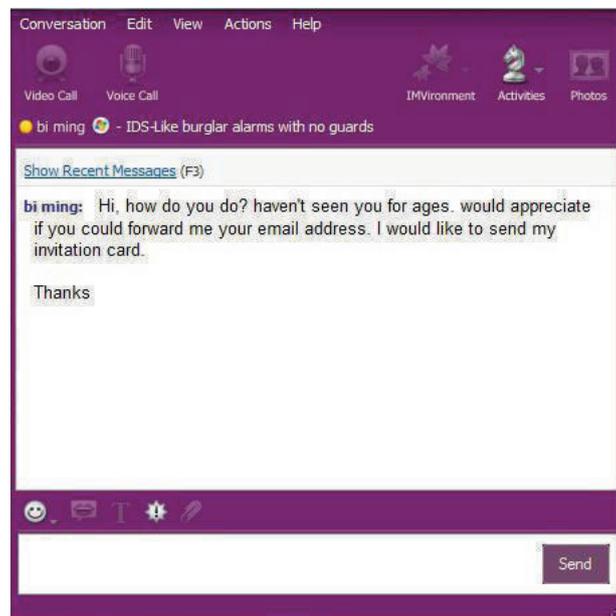


Figure 4.8: Screenshot of Question 6

Example 7

Figure 4.9 shows a screen capture on Example 7 which is a YM offline message sent from a personal contact. The message seems suspicious seeing that it promotes a dubious link. This message was also

sent in a colourful text to attract users to click on the link. To validate whether it was really sent by this contact, the author contacted the sender and the sender confirmed that it was not sent by her. Hence, this message is considered as spam.



Figure 4.9: Screenshot of Question 7

Example 8

Figure 4.10 shows a screen capture on Example 8 which is a YM offline message sent from an unknown contact. The link attached in this message seems dubious. The message content tries to catch the reader's attention as it was sent in a colourful text. The message content itself seems catchy to promote readers to click on the link. The sender's id seems suspicious as it contains the usual spam word. Based on the indications given above, this example is considered as spam.

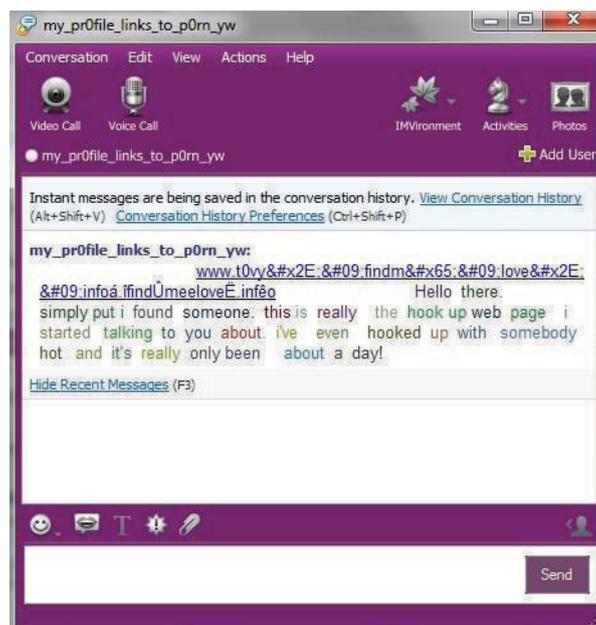


Figure 4.10: Screenshot of Question 8

Example 9

Figure 4.11 shows a screen capture on Example 9 which is a YM offline message. The sender sends a blank message with a space included in it. Although the message does not seem to give any harm to the receiver, it is still best not to reply as the sender might only want to check if the users are active. It also seems very suspicious as it comes from an unknown sender. Thus, this message is considered as spam.

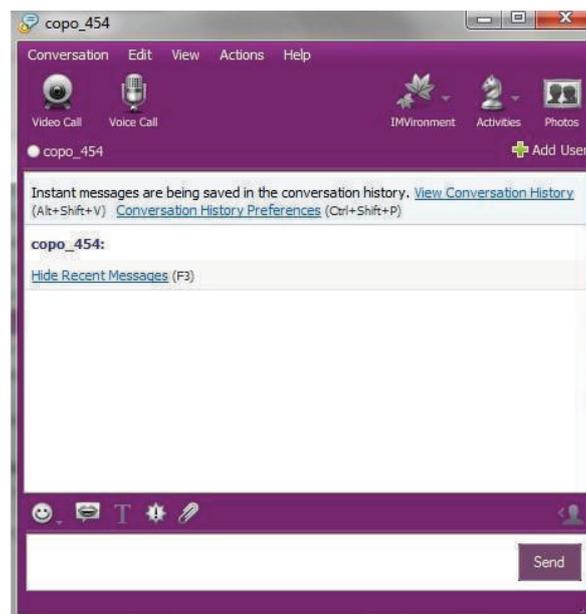


Figure 4.11: Screenshot of Question 9

Example 10

Figure 4.12 shows a screen capture on Example 9 which is a YM offline message. This message comes from a contact that has not been added to the receiver's list. Furthermore, this message was intended to promote alumni or community activity for a university. The message content is also aligned with the link provided. The link is also a genuine link and does not raise any suspicion; thus, this example is considered as non-spam.

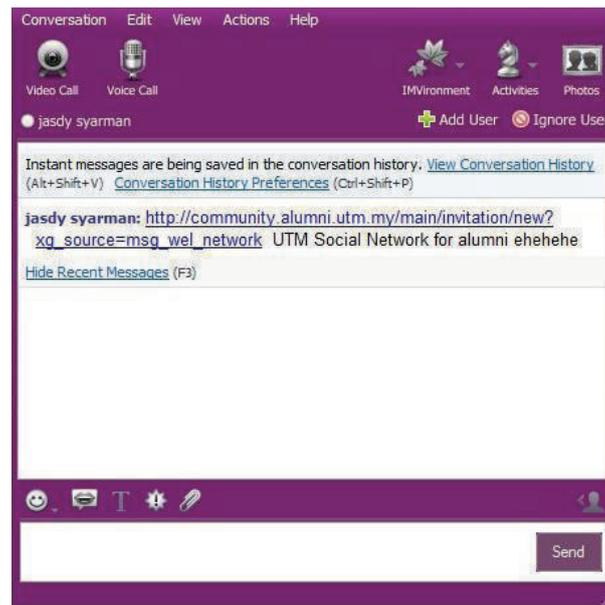


Figure 4.12: Screenshot of Question 10

Generally, in order to categorize whether an example is considered as spam or non-spam, the research follows the heuristics listed below:

- (1) Authorization issue – Does it come from a known source?
- (2) Validation issue – Does it truly comes from that known source?
- (3) Trust issue – Does it raise suspicion or consists of suspicious content?

4.3.1.4 Analysis

The analysis is carried out based on the cost categories and input data sets. The HoneySpam 2.0 data set will be the main input for storage cost. HoneySpam 2.0 includes data set on spam profiles, spam personal messages and spam posts. Thus, the basic analysis of each type of spam units is presented which includes the number of spam received for each type of spam units; highest and lowest spam units received each month and any identified patterns are discussed. In addition, spam units preferred to be manipulated by spammers are identified. Analysis on the spam unit that consumed more storage and the spammer's spamming volume are included in the next chapter.

The timing data set that comes from the survey explained earlier will be the main input to estimate the loss of productivity cost. The basic statistics on the number of responses recorded for each example are analysed. Simple analysis such as total time recorded, average time spent and maximum and minimum time spent for each example is carried out. Related theories with loss of productivity cost are also discussed in Chapter 5.

4.3.1.5 Cost calculation

The cost calculations for each cost category will depend on the formulation that contains cost parameters which are defined in the next section. Each cost parameter is then generated from the data source explained in Section 0 The generated values for each cost parameter are calculated based on the formulations.

In this section, a detailed description involved in Solution 1 is provided. Solution 1, which is based on the design science research methodology, relies on HoneySpam 2.0 and the timing data set from the survey for the data collection stage. The rest of the survey is explained in the next section.

4.3.2 Solution II

This research aims to study and report the public awareness, knowledge and perception of Spam 2.0. This section explains the processes taken to fulfil the objective.

4.3.2.1 Survey Design

To allow us to easily reach the research target group, which are the Internet users, a web survey was chosen to be used as the research instrument. Additionally, using a web survey for data collection has the advantage of lower cost (Weible and Wallace 1998) and faster feedback (Nowack 1997).

Questionnaire development

Although the thesis intends to explore the extent of awareness, knowledge and perception of Spam 2.0, the awareness, knowledge and perception concept has already been widely presented in other studies. Thus, the main idea behind this questionnaire development stage is:

- 1) A study on awareness, knowledge and perception items from computer security and other fields is conducted. Previous validated scales used in existing research studies are followed where possible.
- 2) A study on how to create Spam 2.0-related questions based on current literature review on Spam 2.0 is carried out.
- 3) Accordingly, related parts of collection instrument contained items from the research literature that were modified or developed for the purpose of the study. Thus, questions on the awareness, knowledge and perception of Spam 2.0 were created based on the earlier steps.

The questions were developed in accordance with Dillman's design method for Internet surveys (Dillman 2007). Based on the expert's feedback, minor modifications to the wording of questions in the survey were made. The developed web survey questionnaire consists of 29 questions, divided into

three sections as shown in Figure 4.13. Section A consists of basic demographic questions such as age, gender, education and frequency of Internet usage. Section B comprises of questions on awareness, perception and knowledge. Section C covers the Spam 2.0 identification questions that were designed to assess time taken by respondents to identify Spam 2.0. Section C was developed mainly as a part of the data collection method used in the Spam 2.0 cost model study.

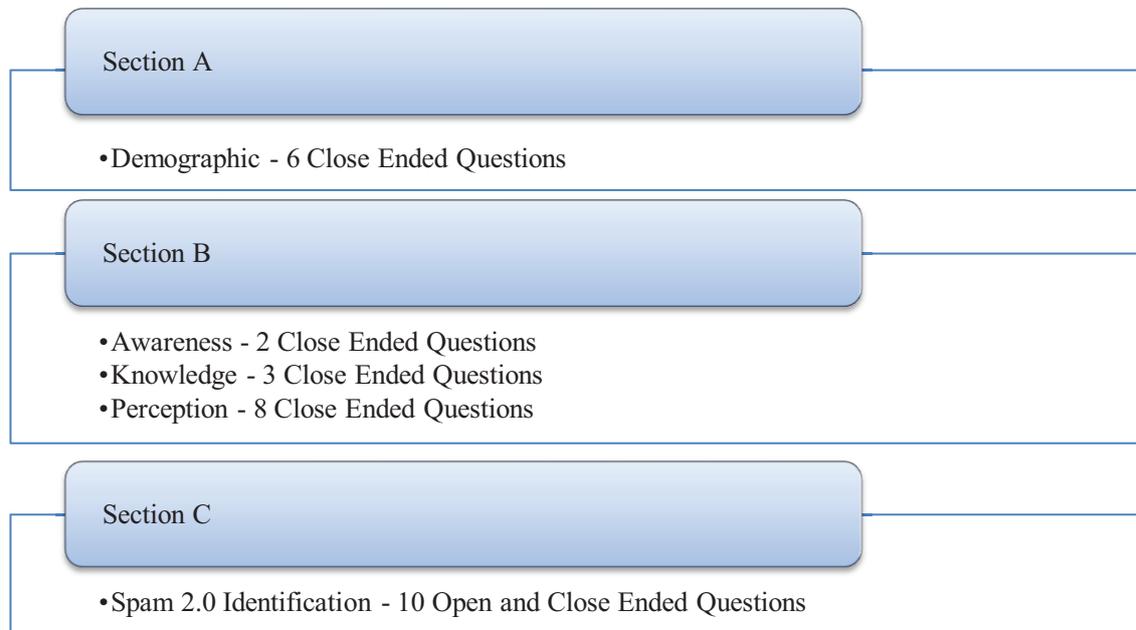


Figure 4.13: Survey design

This survey consists mainly of close-ended questions developed from the existing literature except the 10 questions in Section C. Open-ended questions are used to allow users to provide justification for their chosen answers. Nonetheless, these open-ended questions were optional. The final version of the web survey is included in Appendix

4.3.2.2 Data collection and distribution

For this study, the research used Qualtrics Survey, which is a surveying tool available to Curtin University students and personally administered by the author. The option provided in the software allows researchers to opt for compulsory questions that have to be answered by users, hence decreasing missing data (Stanton 1998).

Sampling method

Data were collected from 368 Internet users in the duration of 7 weeks from 17 February 2012 to April 2012. Only respondents who are at least 18 years old were permitted to participate in the survey. The participants were recruited through link advertising. Personal invitations were given through personal email lists which embedded a personal URL linked to the survey. The participants were

asked to distribute the link to their contacts. Recruitment was also done through link advertising on FB using several personal accounts. In order to achieve more responses and reach a broader response age group, the instrument was distributed and publicized through an invitation linked to the survey on the FB walls of a few community groups. Consequently, the response rate could not be determined. In order to reduce the number of non-responsive potential respondents, the link in FB was published several times and personal email lists were contacted for a reminder to fill up the survey (Wygant et al. 2005; Fan and Yan 2010).

4.3.2.3 Data analysis and assessment

The analysis, assessment of the public awareness, knowledge and perception survey are then presented in Chapter 6 including findings and discussion resulted from the collected data. Analysis of this research is done mainly using SPSS to provide descriptive analysis and simple analysis.

4.4 Summary

Whereas the previous chapters defined the problems, this chapter provides a detailed explanation on their solutions. Section 4.2 provides an overview of Solution I and Solution II. Solution I focuses on unravelling the problems related to the Spam 2.0 cost model (explained in Section 4.3.1) and Solution II focuses on solving the problems related to the public awareness, knowledge and perception of Spam 2.0 (explained in Section 4.3.2). The processes involved in each solution are explained clearly. The processes involved in Solution I include (1) determining the stakeholder, (2) determining the cost category and its cost parameters, (3) data collection, (4) analysis and (5) cost calculations. Solution I uses data from two sources which are HoneySpam 2.0 and timing data sets. The timing data set was one of the smallest parts included in a survey. The overall results of the survey are reported as Solution II. The processes involved in Solution II are (1) survey design, (2) data collection and distribution and (3) data analysis and assessment. These solutions are explained in different chapters accordingly.

4.5 Conclusion

This chapter mainly focuses on the proposed solution for problems that have been defined in earlier chapters. In the first part of the chapter, the overall conceptual processes are linked together to provide a detailed explanation on how problems in this thesis are going to be solved. In the later part of this chapter, two chosen solutions are discussed. Solution 1 focuses on the framework to solve cost-related problems using the design science research methodology while Solution 2 focuses on public awareness, knowledge and perception problems using the quantitative research methodology. The conceptual process used to develop these solutions has also been described in this chapter.

Chapter 5

Spam 2.0 Cost Model

This chapter covers:

- ▶ Introduction to Spam 2.0 cost model;
- ▶ Storage cost calculation; and
- ▶ Loss of productivity cost calculation.

5.1 Introduction

This chapter focuses on Solution I which relates to the Spam 2.0 cost model. As mentioned earlier, the current literature lacks studies on determining the real cost of Spam 2.0. This chapter addresses these problems by providing a detailed solution on estimating the cost of Spam 2.0. The estimation of the Spam 2.0 cost starts by addressing the stakeholder involved in the process. In this research, the focus of the cost revolves around the non-spammer. Calculations of the cost will be worked out based on the individual. This process is followed by determining cost categories and cost parameters involved in the calculations. Thus, Spam 2.0 cost is defined as

Total cost of Spam 2.0

$$\begin{aligned} &= \text{Storage cost} + \text{Loss of productivity} + \text{Labour cost} + \text{Connectivity cost} \\ &+ \text{Software cost} \end{aligned}$$

However, the thesis only focuses on two costs, which are storage cost and loss of productivity cost, for several reasons, such as these costs are expected to be the biggest contributors to our Spam 2.0 cost and time constraint. Connectivity cost however will be included in the storage cost as the package usually comes together. Defining the cost itself requires a comprehensive understanding of email spam cost models and other related cost models. This chapter presents the storage cost and loss of productivity cost of Spam 2.0. Cost parameters involved in each calculation for each cost category are presented in separate subsections accordingly. Although different experiments were carried out and the data were obtained from two different sources, the end result is produced based on a spam unit.

Based on the cost parameters defined in each cost category, detailed cost calculations are generated from the data collected from several sources such as surveys and HoneySpam 2.0. Explanations on

each data collection are provided in Chapter 4. Nonetheless, information obtained from these data collections are analysed and laid out in interrelated subsections in this chapter.

This chapter is organized as follows. Section 5.2 focuses on the formulation and calculation for storage cost of Spam 2.0, while Section 5.3 focuses on the loss of productivity cost. Chapter 5 is concluded in Section 5.4.

5.2 Storage Cost

This section elucidates on the experiments conducted in order to estimate the storage cost of a discussion forum. The thesis first highlights the storage cost formulation defined in Table 4.1 in Section 4.3.1.2 such as follows:

- Size of storage for all spam units in a month, $SS = \sum_{i=1}^3 \sum_{j=1}^{m_i} noc_{ij} * 8B$
- Total accumulated cost of storage, $TCS = \sum_{k=1}^n SS_k * SC_{gm} + RC_{gm}$
- Current average cost of spam per MB, $ACSM = \frac{TCS}{SS}$
- Average total cost of storage for all spam unit in a year, $ATCS_{yearly} = \frac{ACSM}{n} * 12$
- Estimated total average size of a spam unit, $ASC_i = ATCS_{yearly} * ans$

where

- Number of characters, noc
- Spam units, $i = \{\text{spam personal messages, spam profile, spam posts}\}$ the
- Total number of spam for each spam units, m_i
- Storage cost per GB per month, SC_{gm}
- Related additional cost per month, RC_{gm}
- Number of month, n
- Estimated number of spam received in a year, ans

The thesis presents the formulation that can be used for estimating the storage cost of Spam 2.0. The formulation provided here is suitable to be used as there is an exact amount of spam stored internally and each value of the attributes is known clearly. Nonetheless, in the case of the exact amount being unknown, it is a normal practice to estimate the size of storage for all spam units based on an average

size of spam and average number of spam received and using only two of the last equations. Thus, the above formulations are still valid.

From the formulation, it is apparent that we need to identify several cost parameters based on data collected from HoneySpam 2.0 and a storage cost survey. The estimation of the average size of a spam unit depends on the data collected from HoneySpam 2.0 explained in Section 0. Thus, Section 0 first explains the statistics obtained from the data set. Section 0 explains the process in identifying storage cost per GB through a cost survey. Three sources are used to estimate the storage cost of 1MB storage and connectivity of 1MB bandwidth: self-hosted servers, commercial web hosting and cloud hosting. In order to make a comparison, the thesis focuses only on storage and connectivity cost even for self-owned servers and thus drops other costs such as power cost, cooling cost and space cost. Next, Section 0 explains the process of estimating the storage cost based on the formulations. Finally, the related discussion is presented in Section 0

5.2.1 HoneySpam 2.0 Data Set Statistics

Spam units involved in setting up a discussion forum are shown in Figure 5.1. The HoneySpam 2.0 data set contains a total of 62,798 spam profiles, 141 spam personal messages and 450,772 spam post messages collected from June 2010 to June 2011. The time frame for the data set can be divided into two phases. The first phase includes data collected from June 2010 to November 2010 and the second phase includes data collected from February 2011 to April 2011. No polls were created in this duration; thus, the storage cost is calculated based on three spam units: spam profiles, spam personal messages and spam post messages.

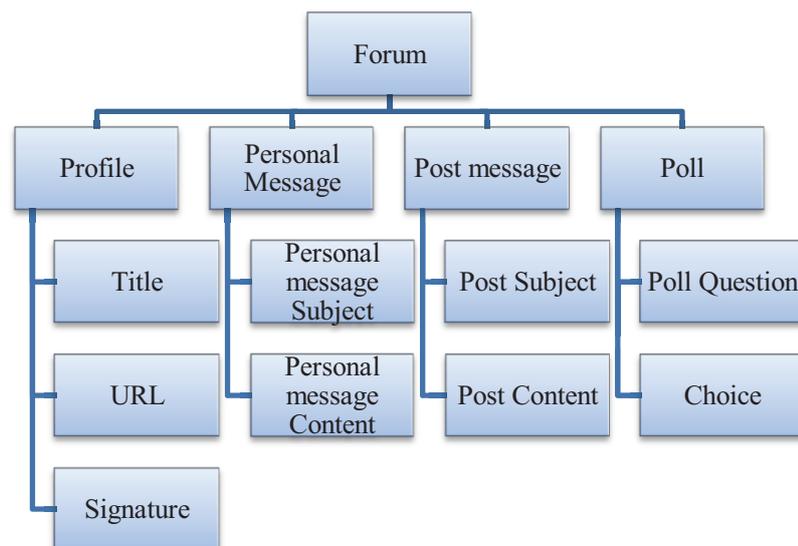


Figure 5.1: Spam units involved in a discussion forum

Overall, in order to estimate the amount of storage used for Spam 2.0 in HoneySpam 2.0, there is a need to consider these formulations below:

- *Profile = title + URL + signature*
- *Poll = poll question + choice*
- *Personal message = personal message subject + personal message content*
- *Post message = post subject + post content*

All content in the forum is stored as HTML. The total size of spam units is calculated based on the number of characters from each spam units. The number of characters for each spam unit needs to be calculated where each ASCII character is equal to 8 bits or 1 byte of size. For each of the formulations above, this process is used to produce the total size used for each spam unit.

5.2.1.1 Spam profile

Table 5.1 presents the details on spam profiles created in the forum from June 2010 to June 2011 including the number of profiles, total number of characters for spam profiles, size for each month's spam profile and accumulated size for all spam profiles. The total number of characters for spam profile is a sum of three forum attributes which are website title, website URL and signature.

Table 5.1: Spam profile

Month-Year	No. of profiles	No. of characters	Size (MB)	Accumulated size (MB)
Jun-10	39	6,341	0.0060	0.0060
Jul-10	1,615	272,715	0.2601	0.2661
Aug-10	6,242	874,888	0.8344	1.1005
Sep-10	7,961	1,338,842	1.2768	2.3773
Oct-10	1,626	307,644	0.2934	2.6707
Nov-10	190	19,815	0.0189	2.6896
Feb-11	8,910	1,094,089	1.0434	3.7330
Mar-11	32,040	4,799,605	4.5773	8.3103
Apr-11	387	20,281	0.0135	8.3238
May-11	374	19,627	0.0187	8.3425
Jun-11	3,414	416,908	0.3976	8.7401
Total	62,798	9,164,648	8.7401	

Only 39 profile spam were created in the forum in the first month of its running. The number continued to increase progressively in the following months. In the first phase, the forum received the highest number of spam profiles created in September 2010 (7,961) followed by August 2010 with 6,242 spam profiles. It is observed that the total number of characters seems to be proportional with the number of spam profiles created in a month, where a spam profile is registered with around 100–200 characters. However, in November 2010, there was a sudden drop in the number of spam profile registrations. There is an average of only six new registrations per day and 7 days without any new registrations.

In the second phase, the forum received the highest registration of spam profiles in March 2011 with 32,040 profiles and 4,799,605 characters. There was also a rapid decline in spammers' registration in April and May 2011 as CAPTCHA was implemented for the forum. Nonetheless, the number of profiles registered increased again in June 2011. It is assumed that spammers are using smarter bots to break the CAPTCHA and successfully register anew as there is no change in the forum setting. In the first phase, the storage used to retain all these profiles was 2.6896MB. By end of the second phase, this value amounted to 8.7401MB in June 2011.

5.2.1.2 Spam personal message

Table 5.2 presents the details on spam personal messages created in the forum including the number of personal messages, total number of characters for personal messages, size for each month's spam personal messages and accumulated size for all personal messages.

Table 5.2: Personal message(pm) spam

Month-Year	No. of pm's	No. of characters	Size (MB)	Accumulated size (MB)
Jun-10	0	0	0	0
Jul-10	0	0	0	0
Aug-10	111	14,214	0.0136	0.0136
Sep-10	0	0	0	0.0136
Oct-10	0	0	0	0.0136
Nov-10	0	0	0	0.0136
Feb-11	0	0	0	0.0136
Mar-11	0	0	0	0.0136
Apr-11	30	3,000	0.0028	0.0164
May-11	0	0	0	0.0164
Jun-11	0	0	0	0.0164
Total	141	17,214	0.0164	

5.2.1.3 Spam post

The details on spam posts created in the forum are presented in Table 5.3. This includes the number of spam posts, total number of characters for spam posts, size for each month's spam posts and accumulated size for all spam posts. The total number of characters for spam posts is generated based on the summation of two forum attributes which are post subject and post content.

Table 5.3: Spam post

Month-Year	No. of posts	Size (MB)	Ratio of (post size/no. of posts)	Accumulated size (MB)
Jun-10	39	0.2395	0.0061	0.2395
Jul-10	10,566	89.8141	0.0085	90.0535
Aug-10	25,527	213.6747	0.0084	303.7283
Sep-10	47,297	346.2396	0.0073	649.9679
Oct-10	72,724	638.9952	0.0088	1288.9631
Nov-10	59,268	625.6723	0.0106	1914.6354
Feb-11	21,203	54.1053	0.0026	1968.7407
Mar-11	81,241	253.9601	0.0031	2222.7007
Apr-11	61,901	213.4939	0.0034	2436.1946
May-11	40,650	173.4988	0.0043	2609.6934
Jun-11	30,356	138.5368	0.0046	2748.2302
Total	450,772	2748.2302	0.0061	

In the first phase, the forum received the highest creation of spam posts in October 2010 (72,724) and the lowest in June 2010 (39). Although the highest number of spam posts was received in October 2010, the highest ratio of size of each post was recorded for November 2010. In other words, on average, spammers send larger spam posts that contain longer texts for this particular month. This also indicates that the number of spam posts does not imply the size of the post. Although the forum received a sudden drop of new profile registrations in November 2010, the number of spam posts that was posted in the forum is still quite high. This indicates that the spammers use the same profile created earlier to send spam posts.

In the second phase, the forum received 81,241 spam posts in March 2011, and it was recorded as the highest number of spam posts received in a month. The lowest number of spam posts was recorded for February 2011 with 21,203. The number of spam posts created in the forum declined gradually starting from April 2011. However, compared to the first phase, the number of spam posts received in

the forum declined in the second phase. No amendment or deletion was made to the spam content; thus, the storage consumed to store all spam post messages reached 2,748MB in June 2011.

Table 5.3 also reveals that the average size for spam posts for the whole duration of data collection was recorded at 6B. In the first phase, the lowest average size for spam posts recorded in a month was in June 2010, which is 6B, and the highest was recorded for November 2010, which is 10B. In the second phase, the lowest average size for spam posts recorded in a month was in February 2011, which is near to 3B. The highest average size for spam posts recorded in a month was in June 2011, which is near to 5B. Thus, it is concluded that the average size for a spam post created in this forum ranges from 3B to 10B.

5.2.1.4 Total spam

The details on spam posts created in the forum are presented in Table 5.4. This includes the total number of spam, total number of characters, size for each month's spam and accumulated size for all spam. The total number of characters for spam is generated based on the summation of spam posts, spam post messages and spam profiles.

Table 5.4: Total spam in the discussion forum

Month-Year	Total spam	No. of characters	Size (MB)	Accumulated size (MB)
Jun-10	78	257,429	0.2455	0.2455
Jul-10	12,181	9,444,9617	90.0742	90.3197
Aug-10	31,880	224,943,266	214.5226	304.8423
Sep-10	55,258	364,397,424	347.5165	652.3588
Oct-10	74,350	670,342,690	639.2886	1291.6474
Nov-10	59,458	656,084,770	625.6912	1917.3386
Feb-11	30,113	57,827,566	55.1487	1972.4872
Mar-11	113,281	271,096,021	258.5373	2231.0245
Apr-11	62,318	223,881,765	213.5103	2444.5348
May-11	41,024	181,946,302	173.5175	2618.0523
Jun-11	33,770	145,683,223	138.9343	2756.9867
Total	513,711	2,891,363,782	2756.9867	

It is apparent that the highest number of spam obtained by the forum in the first phase was in October 2010 followed by November 2010. The total number of spam also reflected the storage size consumed in these months. The lowest number of spam obtained by the forum in the first phase was in June 2010 as the forum was just being created and advertised.

For the second phase, it is observed that the forum received the highest number of spam in March 2011 followed by April 2011. Comparing the first two months in the first phase and the second phase, it is apparent that second phase shows a sudden boost in the total number of spam posted in the forum. This may indicate that the spammers are faster in detecting and reaching out to forums where they could send more spam.

5.2.1.5 HoneySpam 2.0 data set characteristics

This subsection investigates the data set characteristics, such as the spam unit preferred to be manipulated by spammers, spam units that consumed more storage and spammers' spamming volume.

Spam unit preferred to be manipulated by spammers.

The total number of spam according to its spam units and their percentage monthly are presented in Figure 5.2 to investigate the spam unit preferred to be manipulated by spammers.

Figure 5.2 reveals that spam posts accounted for 88% of all spam units. Thus, it is apparent that spammers prefer to manipulate spam posts. This is followed by spam profiles that accounted for 12%. While for spam personal messages, when compared to other spam units, it only accumulates to 0% showing that spam personal messages were not preferred to be used by spammers.

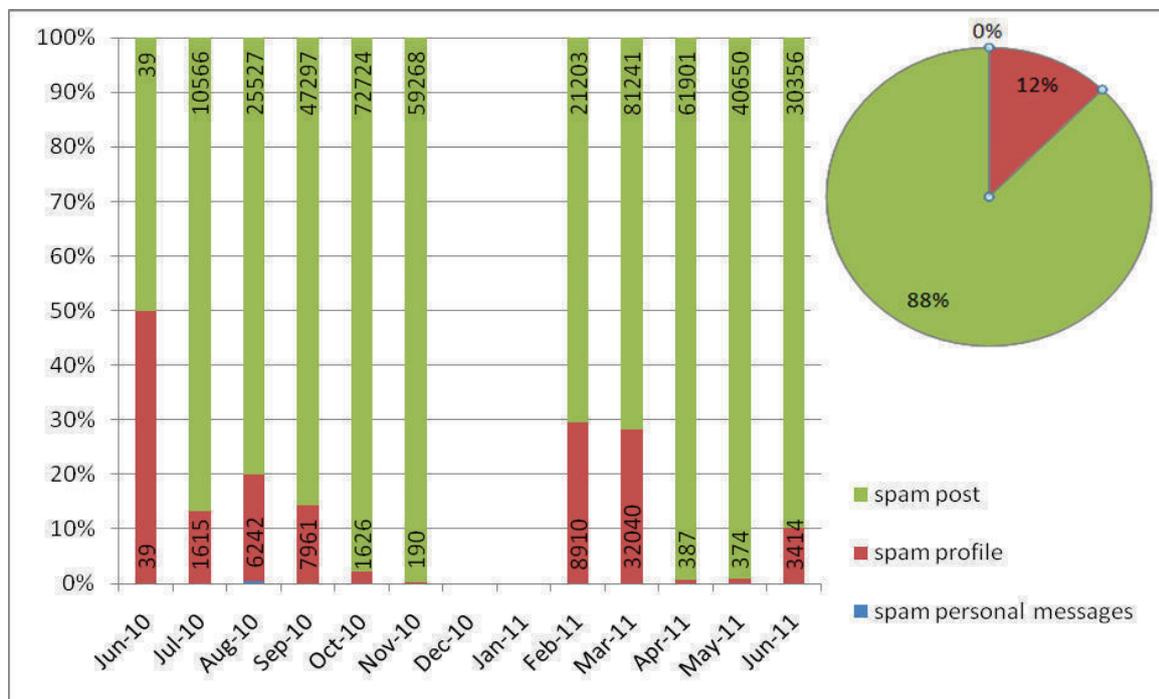


Figure 5.2: Total number of spam units with its percentage according to month

For all months of data collection, it was obvious that more than 70% of the total spam was spammers' spam posts except in June 2010. In this particular month, only 50% of the total spam accounted for

spam posts, while another 50% accounted for spam profiles. This indicates that the first month was specifically focused on spam account registration instead of posting spam posts. However, spam accounts that were registered earlier were used repetitively and enormously to send spam posts.

One of the reasons why spammers prefer to use spam posts is because spam posts have the attribute of high creation flexibility. Thus, spam posts can be created in a massive number using a forum spam automator. Moreover, spam posts will have higher impact view compared to other spam posts. Spam posts possessed both characteristics explained in Section 0. Although the best way to prevent spam posts from being easily manipulated by spammers is to strengthen security before enabling any postings in a forum, this action could also be a hassle for real users.

Spam unit consumed more storage.

Spammers could manipulate all spam units by spamming them to their maximum storage allowed. Even though it was found that all these spam units have the same maximum storage, they could be manipulating spam units at different rates by creating different sizes of spam. Hence, spam units that consumed more storage are presented in Figure 5.3.

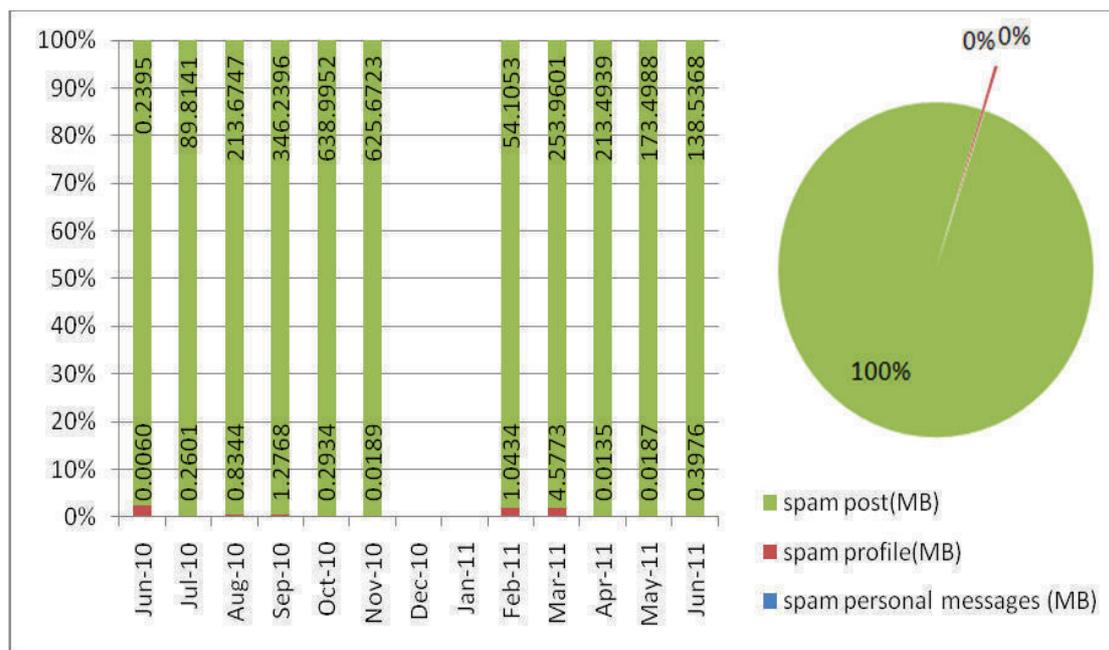


Figure 5.3: Size for each spam units with its percentage according to month

Based on Figure 5.3, it is obvious that 100% of the storage size consumption accounted for spam posts. Although there were a number of spam personal messages and spam profiles, the sum of storage used by both of these spam units was too low compared to spam posts.

Spammers' spamming volume.

From the previous section, it was found that spammers prefer to use spam posts for their spamming activities and spammers send spam of different sizes. As claimed before, spammers registered one

account and used it repetitively to send spam posts. This subsection investigates spammers' spamming volume; thus, the number of spam posts sent by each profile is shown in Figure 5.4.

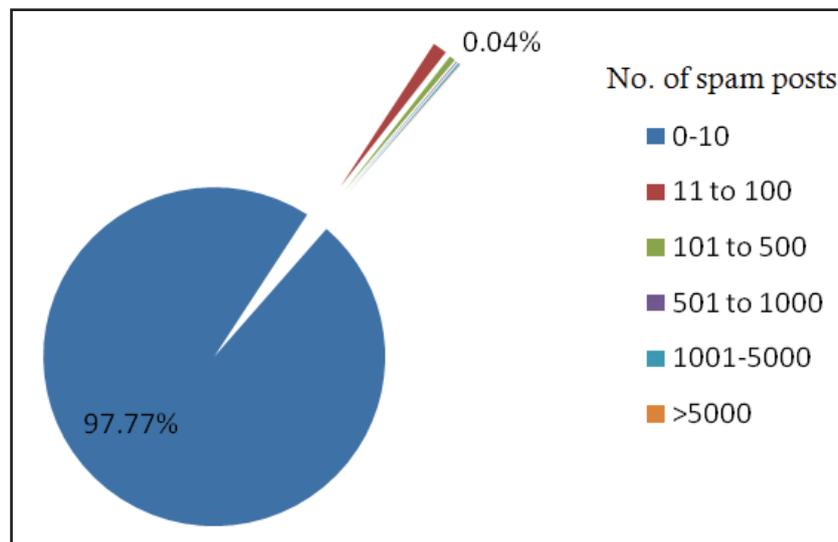


Figure 5.4: Percentage of spam posts sent by spammers divided into six categories

Figure 5.4 reveals that spammers do not spam at a similar volume. Six categories were created based on the number of spam posts to visualize this state, which are 0–10, 11–100, 101–500, 501–1,000, 1,001–5,000 and more than 5,000. The largest percentage accounted for 0–10 (97.77%). Another 0.04% was observed for the more than 5,000 category while other categories accumulated to less than 1%. Additional exploration of the categories with the largest percentage, which are the number of profiles that have created less than 10 posts, is shown in Figure 5.5.

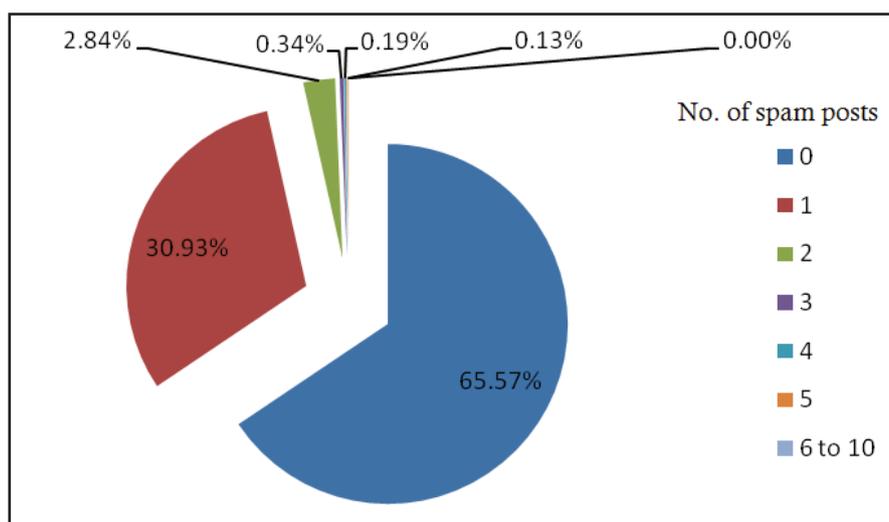


Figure 5.5: Percentage of spam posts sent by spammers for the first category (0–10)

Based on Figure 5.5, it is observed that there are no spammers that have posted 6–10 posts; thus, this category recorded at 0%. The largest percentage was recorded for zero post with 65.57%. Quite a large number of spammers that registered did not create any spam posts. The spammers' motive for

registering without any postings is unclear. However, it is probable that they have a hidden agenda such as to collect the information of other users such as email addresses and use them for other purposes. Nonetheless, they could still advertise and manipulate other attributes contained in their profile.

The second large fraction accumulated to 30.93% was for the category of spammers who posted one post. Nonetheless, each spam post could be spammed at different sizes. Thus, it is possible that even a spammer sends a post, but they manipulate it to the maximum storage with a large number of characters.

This analysis however is done based on the assumption that a spam profile is owned by different individuals. In real life, there is a possibility that a spammer owns more than one profile account. It is insufficient to track IP address geolocations to identify the origins of spamming based on the IP address for each account as the IP address could be rerouted to different sources.

Based on this section, it is concluded that spammers prefer to use spam posts for spamming and thus spam posts use storage the most. Nonetheless, in the next section, the analysis on estimating the cost of storage continues by surveying the storage cost from three different sources.

5.2.2 Storage Cost Survey

In order to generate the value for the storage cost of 1MB storage and connectivity of 1MB bandwidth, three different sources are considered in each subsection as follows:

- Source I : Self-hosted server
- Source II : Commercial web hosting
- Source III : Cloud hosting

5.2.2.1 Self-hosted server

The self-hosted server is considered as our real set-up cost which includes connectivity and storage costs. The server cost is AUD 8,000 for a 146-GB hard disk drive with a lifetime of 5 years. The connectivity cost paid for 200GB on-peak and 200GB off-peak quota for the ADSL connection is AUD 160 per month. Another AUD 50 per year is paid for all services including the domain name. Thus, the related costs are defined as follows:

- *Storage cost perGB per month* $_{source\ 1} = \left(\frac{AUD8000}{146} \right) / 60 = AUD0.91$
- *Connectivity cost perGB for a month* $_{source\ 1} = \frac{AUD160}{400} = AUD0.40$

- *Additional cost*_{source 1} = AUD50

Forum administrators who do not own their own server can always opt for web-hosting packages provided by commercial companies. Related costs used for web-hosting packages are defined in the next subsection.

5.2.2.2 Commercial web hosting

For commercial web hosting, the charges between storage and connectivity are already included in the package. While in terms of charging customers, the provider will charge according to customer usage or the customer will choose a limited package beforehand. If they exceed a certain quota, the customers will be charged accordingly.

Thus, the top five recent commercial companies were examined and 21 basic web hosting packages that they provided were considered. The price per month for all these packages is of a wide range depending on the storage and bandwidth quota given. It is observed that the ratio of storage space to connectivity quota ranges from 1:20 to 1:2 with both the median and average ratio being 1:3.

The costs provided by the commercial companies not only include storage and backup but also contain other costs such as CPU memory, bandwidth and maintenance. Thus, it is assumed that only 25% of the total cost goes towards storage and backup, while the other 75% goes towards other related costs such as bandwidth, server maintenance, human resource cost, dedicated IP addresses, database, email and FTP accounts. The average storage price per GB for these 21 packages is AUD 5.15 per GB per month.

5.2.2.3 Cloud hosting

For cloud hosting packages, customers are allowed to choose their preferred operating system, database, resource management software, web-hosting software and application development servers. Nonetheless, there was not much difference in the prices specified for the three packages surveyed. An average price for three commercial packages (Amazon, Microsoft and Ninefold) is defined as follows:

- *Storage cost perGB per month*_{source 3} = AUD0.13
- *Redundancy storage cost perGB permonth*_{source 3} = AUD 0.118
- *Content delivery per GB per month*_{source 3} = AUD0.195

5.2.3 Estimating the Storage Cost

Based on the formulation provided in Section 5.2, the values need to be obtained and each subsection clearly explains each of the generated values which are as follows:

- Size of storage for all spam units in a month, SS
- Total accumulated cost of storage, TCS
- Current average cost of spam per MB, $ACSM$
- Average total cost of storage for all spam units in a year, $ATCS_{yearly}$
- Estimated total average size of a spam unit, ASC_i

5.2.3.1 Estimating the size of storage for all spam units in a month

Based on Section 0 the value for the size of storage for all spam units in a month, SS , was calculated. These values are similar to those presented in Table 5.4.

5.2.3.2 Calculating the total accumulated cost of storage

The second value, the total accumulated cost of storage, TCS , is calculated based on the first value. The storage cost according to month and the total accumulated cost are calculated based on three sources as explained in Section 0 and the results are presented in Table 5.5.

Table 5.5 shows that the highest total storage cost is calculated for commercial web hosting for the price of AUD 133.90, followed by a self-hosted server, which is AUD 23.66. The lowest total storage cost is calculated for cloud hosting to be AUD 11.53. These costs are very low, but it is apparent that the cost will increase rapidly once the storage used exceeds the standard values provided in the initial package. Hence, it is very important to choose a suitable package in the first place.

5.2.3.3 Estimating current average cost of spam per MB and average total cost of storage for all spams unit per year

The third and fourth equations in the formulation focus on calculating the current average cost of spam per MB, $ACSM$, and the average total cost of storage for all spam units per year, $ATCS$. These values are presented in Table 5.6 and they are calculated based on the storage of spam that existed in the storage by the end of the experiment.

Table 5.5: Storage cost

Month-Year	Accumulated size (MB)	Cost (AUD)		
		Self-hosted server	Commercial web hosting	Cloud hosting
Jun-10	0.2455	0.91	5.15	0.44
Jul-10	90.3197	0.91	5.15	0.44
Aug-10	304.8423	0.91	5.15	0.44
Sep-10	652.3588	0.91	5.15	0.44
Oct-10	1291.6474	1.82	10.30	0.89
Nov-10	1917.3386	1.82	10.30	0.89
Dec-10	1917.3386	1.82	10.30	0.89
Jan-11	1917.3386	1.82	10.30	0.89
Feb-11	1972.4872	1.82	10.30	0.89
Mar-11	2231.0245	2.73	15.45	1.33
Apr-11	2444.5348	2.73	15.45	1.33
May-11	2618.0523	2.73	15.45	1.33
Jun-11	2756.9867	2.73	15.45	1.33
Total		23.66	133.90	11.53

Table 5.6: Storage cost per MB per month

Source	Storage cost	Storage cost per MB	Storage cost per MB per year
Self-hosted server	AUD 23.66	AUD0.0086	AUD0.0079
Commercial web hosting	AUD133.90	AUD0.0486	AUD0.5265
Cloud hosting	AUD 11.53	AUD0.0042	AUD0.0039

5.2.3.4 Estimating the total average size of a spam unit

Spammers create spam of different sizes and send them in different volumes. This value is used as an input in the calculations of the average storage cost in a year for each spam unit, *ASC*. The average size for each spam unit is presented in **Error! Not a valid bookmark self-reference..**

In this case, it is assumed that the estimated number of spam received in a year is 100,000. The average sizes for spam profile, spam personal message and spam post are calculated based on the number of spam units created in the forum and the sum of storage for each spam unit. The lowest average size of a spam unit is calculated for personal messages, which is 11.6MB for 100,000 posts, followed by spam profiles with 13.9MB for 100,000 posts. When compared to other spam units, the

highest average size of a spam unit is calculated for spam posts, with an average of 609.7MB for 100,000 posts.

Table 5.7: Average size of spam units

Spam unit	No. of spam	Storage size (MB)	Average size (MB)	Average size for 100,000 posts (MB)
Profile	62,798	8.7401	0.000139	13.9
Personal message	141	0.0164	0.000116	11.6
Post	450,772	2748.2302	0.006097	609.7

5.2.3.5 Estimating the average cost per year for each spam unit

The analysis on estimating storage cost continues by estimating the average size of each spam unit per year by taking the amount of estimated number of spam received in a year to be 100,000. The storage cost per MB per year is presented in Table 5.6 and the size of 100,000 spam units is obtained from Spammers create spam of different sizes and send them in different volumes. This value is used as an input in the calculations of the average storage cost in a year for each spam unit, *ASC*. The average size for each spam unit is presented in **Error! Not a valid bookmark self-reference..**

In this case, it is assumed that the estimated number of spam received in a year is 100,000. The average sizes for spam profile, spam personal message and spam post are calculated based on the number of spam units created in the forum and the sum of storage for each spam unit. The lowest average size of a spam unit is calculated for personal messages, which is 11.6MB for 100,000 posts, followed by spam profiles with 13.9MB for 100,000 posts. When compared to other spam units, the highest average size of a spam unit is calculated for spam posts, with an average of 609.7MB for 100,000 posts.

Table 5.7. Both inputs are used to calculate the storage cost for 100,000 spam profiles, spam personal messages and spam posts as shown in **Error! Not a valid bookmark self-reference..**

Table 5.8: Average storage cost for 100,000 spam in a year

Spam Unit	Self-hosted server	Commercial web-hosting	Cloud hosting
Profile	AUD0.110	AUD 7.319	AUD0.054
Personal message	AUD0.916	AUD 6.107	AUD0.045
Post	AUD4.817	AUD321.007	AUD2.378

It is observed from the results presented in The analysis on estimating storage cost continues by estimating the average size of each spam unit per year by taking the amount of estimated number of spam received in a year to be 100,000. The storage cost per MB per year is presented in Table 5.6 and the size of 100,000 spam units is obtained from Spammers create spam of different sizes and send them in different volumes. This value is used as an input in the calculations of the average storage cost in a year for each spam unit, *ASC*. The average size for each spam unit is presented in **Error! Not a valid bookmark self-reference..**

In this case, it is assumed that the estimated number of spam received in a year is 100,000. The average sizes for spam profile, spam personal message and spam post are calculated based on the number of spam units created in the forum and the sum of storage for each spam unit. The lowest average size of a spam unit is calculated for personal messages, which is 11.6MB for 100,000 posts, followed by spam profiles with 13.9MB for 100,000 posts. When compared to other spam units, the highest average size of a spam unit is calculated for spam posts, with an average of 609.7MB for 100,000 posts.

Table 5.7. Both inputs are used to calculate the storage cost for 100,000 spam profiles, spam personal messages and spam posts as shown in **Error! Not a valid bookmark self-reference..**

Table 5.8 that the highest cost comes from commercial web-hosting packages for spam posts due to the reason that the basic cost itself is much higher than other sources.

5.2.4 Discussion

The formulation provided in the earlier section involves three spam units which are personal messages, posts and profiles as there are no polls manipulated by spammers in the forum. The formulation for storage cost can also be applied for other web 2.0 applications with the modification in identifying the spam unit related to that particular web 2.0 application. Nonetheless, spammers will manipulate spam units that possess bigger maximum storage with the attribute of high creation flexibility and have a high viewer impact to strengthen their spamming campaign.

As for the storage cost calculated previously, it is clear that the cost will expand continuously unless the forum administrator deletes all these spam contents. Thus, it is concluded that despite any further actions taken, spam definitely wastes network resources. Considering the administrator's workload to detect and read each post and profile created every day, filtering software and CAPTCHA are used to act as the first layer of security. The cost of implementing commercial filtering services for a forum is listed in **Error! Not a valid bookmark self-reference..**

Table 5.9: Commercial filtering services

Filtering services	Prices per month	Notes
Akismet	AUD47	Limited for five sites, unlimited posts per month
Mollom	AUD40	Limited for one site, unlimited spam posts, 1000 legitimate posts per day

Implementing commercial filtering services will definitely increase the total cost calculated earlier. However, considering that these filtering services work effectively, the possibility of the amount of storage exceeding the basic packages is extremely low. Furthermore, if the administrator decides not to implement the filtering services, then they will have to read, detect and eliminate spam manually, which will incur more time, and thus more labour cost is incurred on eliminating spam.

On a side note, the total storage cost was calculated based on three different packages. It is expected that with newer developments on cloud hosting, the cost of storage will become cheaper in the future. Nonetheless, while focusing on storage cost, most of the concerns are focused on the size of a spam unit. However, even if spammers are posting spam posts of smaller size, the danger of spam content embedded is more troublesome as it could probably raise other security issues.

5.3 Loss of Productivity Cost

This section focuses on the experiments done in order to estimate the loss of productivity cost. The thesis first highlights the loss of productivity cost formulation defined in Table 4.1 in Section 4.3.1.2 listed as follows:

- Average time wasted for each spam for an individual in minutes, $ate = \frac{ttw}{m}/x$
- Average spam unit received in a day, $asd = \frac{m}{n*d}$
- Loss of productivity cost for an individual monthly, $LoP_{daily} = \frac{ate*asd}{twh} * ism$
- Loss of productivity cost for an individual annually, $LoP_{annually} = LoP_{daily} * twdy$

where

- Average time wasted for each spam unit for an individual in minutes, ate
- Total time wasted for all spam in minutes, ttw
- Total number of spam, m

- Number of individuals, x
- Number of months, n
- Number of days, d
- Average spam units received in a month, asd
- Total working hours in minutes per month, twh
- Total working days in a year, $twdy$
- Individual salary per month, ism

The data used for this experiment were collected using a survey as explained in Section 0. Thus, Section 5.3.1 first presents the data obtained from the survey used particularly in this cost namely the timing data set. Section 0 provides the calculation involved in producing loss of productivity cost.

5.3.1 Timing Data Set

Timing data sets were collected from a survey on 368 participants using the timing function on 10 Spam 2.0 examples as explained in Section 0. In this section, the study of the cost only focuses on the time taken by participants for spam identification, $t_{identify}$. For clarity, the thesis first defines that time used for spam identification as in the equation:

$$t_{identify} = t_{read} + t_{load} + t_{look} + t_{decide} + t_{click}$$

where

- Time used for spam identification, $t_{identify}$
- Time used to read question, t_{read}
- Time used to load the page, t_{load}
- Time used to look and skim the example, t_{look}
- Time used to identify and decide if its spam or non-spam, t_{decide}
- Time used to click and submit the answer, t_{click}

This experiment was conducted in a controlled setting where an assessment on how long it took for the user to do all these activities was carried out. Based on the design of the web survey, there is a need to minimize the values for t_{load} , t_{look} and t_{click} so that $t_{identify} = t_{read} + t_{decide}$. Hence, the

value obtained from the timing function is the value totally used for spam identification. The time for spam identification can then be categorized into two categories, which are time wasted and cost of misjudgement. The thesis first defines these terms below:

- **Time wasted:** Time wasted in the context of this research is defined as the amount of time taken to handle identified spam which could lead to further actions of managing spam such as reporting, flagging or deleting them.
- **Cost of misjudgement:** Cost of misjudgement in the context of this research is defined as any related costs that serve as the outcomes of incorrectly identified spam as non-spam or non-spam as spam. These include falling and becoming a victim of a spam campaign, facing security attacks, tarnishing a website's reputation or causing distrust, hate and annoyance.

Taken from the user's perspective, time will be considered as a waste only if users think they are dealing with spam. The justification behind this idea is because of further actions that might arise due to incidents of spam such as users possibly having to report or flag or being redirected to another page. The relationship between the researcher's view and respondents' view of spam and non-spam is depicted in Table 5.10. For spam that has been categorized as spam by researchers and seen as spam by respondent as well, it is considered as a waste by the users.

In cases where it is categorized as spam by researchers but identified as non-spam by respondents, the time spent will not be considered as a waste by the users. Nevertheless, the cost of mistakenly identifying spam as non-spam might bring about bad consequences for the users. In this case, these users might fall for the spamming campaign which could further lead to consequences such as being led to face other security attacks such as malware, scams, viruses, etc.

Table 5.10: Relationship between researcher and respondents' view of spam and non-spam and its costs

Researcher's view	Spam		Non-spam	
Respondents' view	Spam	Non-spam	Spam	Non-spam
Related cost	Time wasted	Cost of misjudgement	Time wasted Cost of misjudgement	None

As for entries that were categorized as non-spam but seen as spam by participants, the time spent will be considered as a waste. This is because users might spend more time to report or flag the spam. In addition, the website's reputation or, in our study, the owner's account reputation will be tarnished. Other than that, there were other intangible costs on the line. For example, the messages could create hate and annoyance on the user's side. In the future, a message that comes from a similar owner's account will probably be ignored and deleted, thus impacting future communications. To further

explain the idea presented in Table 5.10, the steps as shown in Figure 5.6 were proposed to determine the cost of identifying Spam 2.0 based on the assessment of respondents' knowledge rather than their own estimation.

Figure 5.6 depicts the steps taken to estimate the cost of identifying Spam 2.0 in terms of time for each participant. Wasted time and total cost of misjudgement will be estimated in these processes. The time spent for each example will be recorded as recorded time, t_r . Initialization of t_r is done for each example. It is important to differentiate between spam and non-spam questions. For the spam examples, if the participants correctly identified spam and answered "yes," then t_r will be considered as wasted time, and the number of responses for time wasted, N_{tw} , will be updated. However, if respondents failed to answer correctly, then t_r will be considered as cost of misjudgement, and the number of responses for cost of misjudgement, N_{com} , will be added by 1. In the case of non-spam examples, if participants mistakenly identified non-spam as spam and answered "no," then t_r will be considered both as time used for spam identification and cost of misjudgement. Thus, both the number of responses for time wasted, N_{tw} , and number of responses for cost of misjudgement, N_{com} , will also be updated. Nonetheless, if the participants correctly identified the example as non-spam, then the t_r value will be discarded. These steps will be repeated for each 10 examples for 368 participants.

The following subsections present the results obtained from these data sets. This includes the basic statistics of time taken for each example, result for time wasted and results for cost of misjudgement.

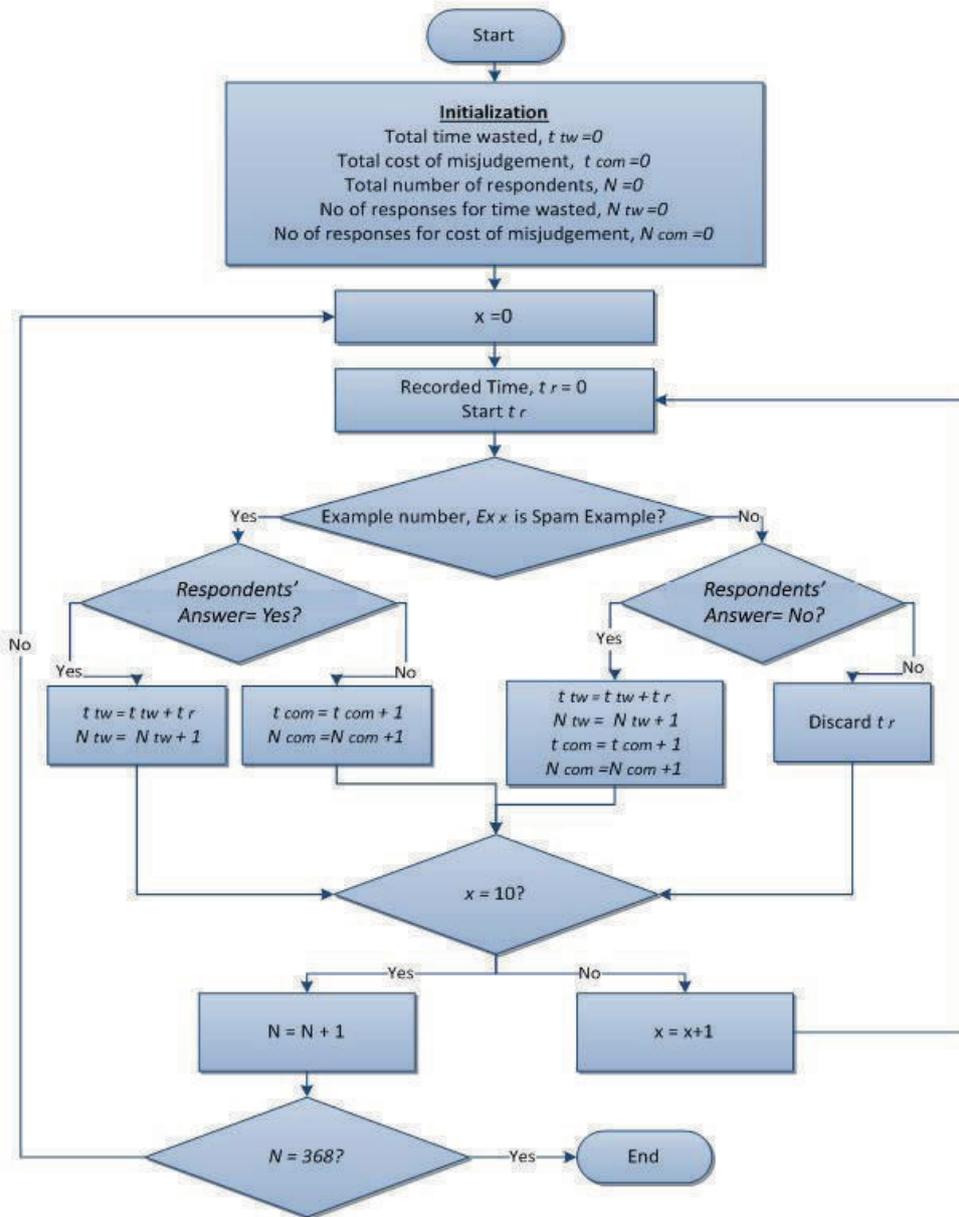


Figure 5.6: Steps to estimate time used for Spam 2.0 identification

5.3.1.1 Basic statistics of timing data set

The basic statistics of the timing data set are presented for each example in **Error! Not a valid bookmark self-reference.**. Outliers were detected in the data set and total time between with and without these outliers were compared. Outliers' data are then decided for the time that exceeds 5 minutes or 300 seconds. Data cleansing is also done for missing values and those with 0 values. There is no definite way of confirming how the times recorded are fully used to answer those questions. However, the research has taken the initiative of informing users to cooperate in giving the best result. On the first page of the survey, it was written that the purpose of the study is to record the time used in identifying Spam 2.0. In this section, the results are reported in seconds as some of the values are very low.

Error! Not a valid bookmark self-reference. shows the basic statistics analysis for time recorded for each response made by the participants. A total time to 62,395.87 seconds was recorded for a total of 3,523 responses. The range of minimum time for answers recorded for all examples is 0.03–1.75 seconds. The lowest minimum recorded time between all 10 examples is 0.03 for Examples 8 and 10 and the highest minimum recorded time is 1.75 for Example 6. On the other hand, the maximum recorded time for all examples for responses excluding the outliers is in the range of 146–294 seconds, which is an equivalent of 2–5 minutes. The highest maximum recorded time between all 10 examples is 294.98 seconds for Example 7 and the lowest maximum recorded time is 146.54 for Example 8. The average time recorded for all examples valued in the range of 10–28 seconds with Example 8 having the lowest average recorded time (10.54 seconds). On the other hand, the highest average recorded time is 27.85 for Example 10.

Table 5.11: Basic statistics of timing data set for each example

Example	No. of responses	Recorded time (seconds)			
		Total	Min	Max	Avg
1	348	9440.34	0.07	263.72	26.74
2	357	4913.24	0.93	157.27	13.65
3	358	7048.67	0.04	271.54	19.63
4	349	6246.72	0.06	276.48	17.75
5	352	5044.73	0.17	260.79	14.21
6	352	6079.64	1.75	250.31	17.08
7	351	4945.23	0.27	194.17	13.81
8	352	3761.32	0.03	146.54	10.54
9	354	5029.88	0.44	294.98	14.01
10	350	9886.11	0.03	193.19	27.85
	3523	62395.87			17.71

Example 8 had the lowest minimum, maximum and average time compared to other examples. This indicates that participants have the tendency to identify correctly in a short time as Example 8 contains a simple spam question with the clearest indications, i.e. the link attached to this message seems dubious, the message content was sent in a colourful text to attract readers and the sender's id also contains the usual spam word such as "*porn*."

5.3.1.2 Time wasted

Based on the analysis of each example, the steps presented in Figure 5.6 are followed and the results are presented as shown in **Error! Not a valid bookmark self-reference.**. Thus, it is estimated that 40,142.41 seconds or 669 minutes, which is equal to 11 hours and 9 minutes, were wasted for spam identification for 2,345 responses. On average, 17 seconds were used by each participant to identify each example. This number might look small; however, there are thousands of spam available on the Internet these days. This cost of time does not even include the time to remove or report those spam entries yet.

Table 5.12: Total time for related costs for spam identification

Researcher's view	Spam		Non-spam	
Respondents' view	Spam	Non-spam	Spam	Non-spam
Total time (seconds)	17064.93	11025.08	23077.48	11228.39
Total time (mins)	~284	~184	~385	~187

5.3.1.3 Cost of misjudgement

Based on Based on the analysis of each example, the steps presented in Figure 5.6 are followed and the results are presented as shown in **Error! Not a valid bookmark self-reference.**. Thus, it is estimated that 40,142.41 seconds or 669 minutes, which is equal to 11 hours and 9 minutes, were wasted for spam identification for 2,345 responses. On average, 17 seconds were used by each participant to identify each example. This number might look small; however, there are thousands of spam available on the Internet these days. This cost of time does not even include the time to remove or report those spam entries yet.

Table 5.12 and the explanation provided earlier in Table 5.10, it is estimated that the cost of misjudgement is 34,102.56 seconds or 568 minutes which is equal to 9 hours and 28 minutes calculated for 1,759 responses. On average, 19.39 seconds were calculated as the cost of misjudgement for each example per respondent. In other words, in nearly 20 seconds, respondents

could have incorrectly made a wrong decision in identifying spam as non-spam or mistakenly identifying non-spam as spam, thus facing other consequences.

5.3.2 Estimating Loss of Productivity Cost

Based on the formulation provided in Section 5.3, the values that need to be obtained are as follows:

- Average time wasted for each spam for an individual (in minutes), ate
- Average spam unit received in a day, asd
- Loss of productivity cost for an individual daily, LOP_{daily}
- Loss of productivity cost for an individual annually, $LOP_{annually}$

5.3.2.1 Estimating average time wasted for each spam for an individual

The average time wasted for each spam for an individual are as calculated in Section 0The average time wasted to identify a spam entry is 17 seconds. An equivalence of seven spam entries will cost a single person 2 minutes for spam identification. Thus,

$$ate = 17 \text{ seconds} = \frac{2}{7} (\text{minutes})$$

5.3.2.2 Estimating average spam units received in a day

It is not suitable to take the value of average spam units received in a day from this data set, as the number of spam provided in this data set was predetermined. Thus, in order to obtain a value from real observation, the values were taken from HoneySpam 2.0 datasets. The average spam units received in a day was taken as one-third proportion as the data were collected 24 hours non-stop. Using data from Table 5.4 in Section 0thus

$$asd = \frac{1027422}{13 * 3 * 30} = 878.14 \sim 878 \text{ spam units}$$

5.3.2.3 Estimating loss of productivity cost per individual per day

In order to calculate the loss of productivity cost per individual per day, LOP_{daily} , there is a need to calculate the total working hours in a month (in minutes), twh . Using the standard full-time working hours in a month (minutes), which is 8 hours,

$$twh = 8 * 60 * 30 = 14,400$$

The thesis obtained the value of individual salary per month, ism , based on the full-time adult ordinary time earnings provided by the Australian Bureau of Statistics (Australian Bureau of Statistics 2013). Using the pre-calculated values of the average time wasted to identify a spam entry, ate , average spam units received in a month, asd , and total working hours in minutes per month, twh , the loss of productivity cost per individual per month is calculated as:

$$LoP_{daily} = ate * asd = \frac{2}{7} * 878 * \frac{AUD1526.80}{14400} = AUD26.60$$

5.3.2.4 Estimating loss of productivity cost per individual per year

In order to estimate the monetary cost of loss of productivity per individual annually, a predefined values of total working days in a year, $twdy = 250$, is used. Hence,

$$LoP_{annually} = 26.60 * 250 = AUD6650$$

5.3.3 Discussion

The formulations introduced in the earlier section require average values such as average time wasted for an individual and average spam units received in a day. Thus, the formulation for loss of productivity can also be applied for other web 2.0 applications with the modification in average values using related spam units. The work and results presented here are one of the earliest works in identifying these values. To obtain such values, the thesis proposed to use the timing function and further eliminate the estimated values given directly by the survey respondents. Still, zero and missing values were obtained for certain respondents. However, the total values are still nearly 95% out of the total respondents. Thus, there should not be much difference in the calculated results.

The results produced for loss of productivity cost should be able to portray how much Spam 2.0 can cost. This will also help to increase awareness among public users on the cost of Spam 2.0 as it not only incurs monetary cost but also causes a waste of time.

5.4 Conclusion

In this chapter, the storage cost and loss of productivity cost formulation were proposed to identify the cost of Spam 2.0. The storage cost was calculated based on HoneySpam 2.0 data sets and storage cost survey. Three spam units were considered in the calculations based on the price costs of self-hosted servers, commercial web hosting and cloud hosting packages. As a conclusion, the range of storage cost of 100,000 spam units in a year is between AUD 0.110 and AUD 321.007 depending on the related costs and size of spam units. Storage cost was explained in Section 5.2.

By contrast, the loss of productivity cost was mostly generated based on timing data sets obtained from the public awareness, knowledge and perception survey. The loss of productivity cost was explained in Section 5.3. Two cost categories were identified from the data set, namely the time wasted and cost of misjudgement. The time wasted was then used in the calculations of the loss of productivity cost to identify the average time used for Spam 2.0 identification. It is approximated that the loss of productivity cost for an individual with an average salary of AUD 1,526.80 in a year is AUD 6,650.

The following chapter provides an in-depth analysis of the survey on the public awareness, knowledge and perception of Spam 2.0.

Chapter 6

Public Awareness, Knowledge and Perception of Spam 2.0

This chapter covers:

- ▶ The respondent demographics of the Spam 2.0 survey;
- ▶ The descriptive analysis of the public awareness of Spam 2.0;
- ▶ The descriptive analysis of the public knowledge of Spam 2.0;
- ▶ The descriptive analysis of the public perception of Spam 2.0; and
- ▶ The analysis of respondents' justification comments.

6.1 Introduction

As mentioned earlier in this thesis, this research focus is twofold, which are the Spam 2.0 cost model and public awareness, knowledge and perception of Spam 2.0. The current literature lacks a comprehensive understanding of public awareness, knowledge and their views on computer security issues specifically Spam 2.0. Hence, it is concluded that the best way to assess the public awareness, knowledge and perception of Spam 2.0 is through an online survey. This chapter provides the statistical analysis of the data collected using the web-based survey as described in Section 4.3.2

This chapter is organized as follows: Section 6.2 starts with an overview of the participating respondents' demographic profiles. Section 6.3 provides a descriptive analysis of the public awareness, knowledge and perception of Spam 2.0 and comments analysis. Discussion is provided in Section 6.4 and conclusions from the survey are discussed in Section 6.5.

6.2 Respondent Demographics

The demographics of the respondents were profiled to gain a clear picture of the sample. This section provides the results gained from Part A in the survey. A demographic profile is summarized in **Error! Not a valid bookmark self-reference.** according to gender, age group, education level, average hours

spent using Internet per day, work/study related to technology field and activities engaged in when using the Internet. Gender, age group and education level are considered as sensitive questions; thus, respondents are allowed to refuse exposing these details. A total of 368 respondents completed the survey.

Table 6.1: Summary of respondent demographics

Items	Categories	Percentage	Frequency
Gender	Male	35.3	130
	Female	61.4	226
	Refused to specify	3.3	12
Age group	18–24	41.3	152
	25–34	35.3	130
	35–49	20.4	75
	50–64	1.4	5
	Refused to specify	1.6	6
Education level	Primary	0.3	1
	Secondary	0.8	3
	Certification	1.6	6
	Diploma/advanced diploma	4.9	18
	Undergraduate	43.2	159
	Postgraduate	45.4	167
	Refused to specify	3.8	14
Average hours spent using Internet per day	Less than 1 hour	3.3	12
	1–5 hours	51.6	190
	6–9 hours	29.9	110
	More than 9 hours	15.2	56
Work/study-related to a technology field	Yes	74.2	273
	No	25.8	95
Activities engaged in when using the Internet	Searching for information	88.3	325
	Gaming	8.4	31
	Chatting and social networking	73.1	269
	Email	73.9	272
	Other	10.6	39

A female majority of respondents answered the survey. The profile data included 35.3% male ($n=130$) and 61.4% ($n=226$). Another 3.3% ($n=12$) refused to specify their gender.

The respondents were asked to indicate their age. Age was divided into four different groups, which are 18–24, 25–34, 35–49 and finally from 50 to 64. Of the respondents, 41.3% are aged from 18 to 24, whereas 35.3% are aged from 25 to 34. Only 1.4% of the respondents are in the age group of 50–64. More than 70% of the respondents are aged below 35. A younger age group majority could be expected due to the fact that this survey was carried out online.

Based on The demographics of the respondents were profiled to gain a clear picture of the sample. This section provides the results gained from Part A in the survey. A demographic profile is summarized in **Error! Not a valid bookmark self-reference.** according to gender, age group, education level, average hours spent using Internet per day, work/study related to technology field and activities engaged in when using the Internet. Gender, age group and education level are considered as sensitive questions; thus, respondents are allowed to refuse exposing these details. A total of 368 respondents completed the survey.

Table 6.1, only 0.3% of 360 respondents has primary education, 0.8% have secondary, 1.6% have certification and another 4.9% have a diploma or advanced diploma. The level of education under the undergraduate and postgraduate categories accounted for 43.2% and 45.4% of the participants, respectively. Of the respondents, 3.8% refused to specify their level of education. Out of 368 respondents, more than 80% completed at least undergraduate studies. This could be expected as the link to the web survey was promoted in educational groups.

Only 3.3% of the respondents spent less than 1 hour on average per day using the Internet. More than 50% of them normally spent 1–5 hours per day. Another 29.9% of the respondents use Internet for an average of 6–9 hours per day.

A majority of the respondents (74.2%) work or study in areas related to a technology field, while the rest (25.8%) work or study in areas not related to a technology field.

The respondents were also asked to specify the activities they engaged in when using the Internet. Of the respondents, 88.3% use the Internet to search for information, whereas 73.9% and 73.1% use the Internet for email and both chatting and social networking, respectively. Another 8.4% use the Internet for gaming, while 10.6% of the respondents specify others including for online banking and entertainment that includes downloading and streaming movies and songs.

It can be observed from the explanation provided in the demographic profile that:

- Approximately one-third of the respondents (61.4%) were female.
- More than three-quarters of the respondents (76.6%) were aged less than 34 years.
- A majority of the respondents had at least undergraduate level of education (88.6%).

- More than two-fifths of the respondents (44.5%) spent at least 6 hours on average per day using the Internet.
- Approximately three-quarters of the respondents' work or study (74.2%) is related to a technology field.
- Most popular activities engaged in when using the Internet specified by the respondents were searching for information (88.3%), followed by email (73.9%) and chatting and social networking (73.1%).

6.3 Descriptive Analysis

This section presents a descriptive analysis of users' awareness, knowledge and perception of Spam 2.0 followed by a summary of observations written for each issue. A small subsection on comments analysis from Part C in the survey is included in this section. This section presents the result obtained in both Part B and Part C of the survey. In order to provide a better understanding for the respondents to be able to answer the survey's questions, the term online spam was used instead of Spam 2.0 as the term "Spam 2.0" sounds too technical for public users.

6.3.1 Awareness of Spam 2.0

This section reports data on awareness-related questions obtained in Part B of the survey as mentioned in Section 0The public awareness of Spam 2.0 includes perceived awareness and actual awareness. Descriptive statistics on perceived awareness are explained in Section 0 and descriptive statistics on actual awareness are explained in Section 0Section 6.3.2.3 focuses on the analysis of the level of awareness.

6.3.1.1 Perceived awareness

To assess the perceived awareness of online spam, respondents were asked the question "*Have you ever heard of online spam?*" Respondents were given the choice to answer either "*yes*" or "*no*." **Error! Not a valid bookmark self-reference.** shows that 91.6% of the 368 respondents indicated that they have heard of online spam, while only 8.4% of the respondents stated that they have never heard of online spam.

Table 6.2: Perceived awareness of online spam

Items	Yes	No
-------	-----	----

	%	Frequency	%	Frequency
Have you ever heard of online spam?	91.6	337	8.4	31

It can be observed from To assess the perceived awareness of online spam, respondents were asked the question “*Have you ever heard of online spam?*” Respondents were given the choice to answer either “*yes*” or “*no*.” **Error! Not a valid bookmark self-reference.** shows that 91.6% of the 368 respondents indicated that they have heard of online spam, while only 8.4% of the respondents stated that they have never heard of online spam.

Table 6.2 which covers the actual awareness of online spam that:

- The total respondents’ perceived awareness is very high (91.6%).

6.3.1.2 Actual awareness

To evaluate the actual awareness of online spam, respondents were asked, “*Have you had any of these experiences while browsing the Internet?*” “YES,” “NO” and “DON’T KNOW” answers were offered as a choice of answer. Questions presented in this subsection are consistent with subsequent questions in the ‘knowledge section’; thus, there were 13 statements of suspicious activities. Out of this 13, only 10 of them are related to online spam. If users have encountered any spam-related activities identified in any of these 10 statements, then they are translated to having actual awareness. The online spam-related activities are as follows:

- Found pages that are only full of repeated keywords.
- Being redirected to an unrelated page from what was expected.
- Found pages with repetitive links.
- Received message considered as unwanted/suspicious/annoying on a web 2.0 application.
- Found pages with unrelated links.
- Received/seen suspicious link on a web 2.0 application.
- Found pages that are only advertising with very little content.
- Received unwanted postings on their social network account.
- Received unwanted friend requests on their social network account.
- Being tagged by unwanted parties on a Web 2.0 application.

Error! Not a valid bookmark self-reference. recapitulates data on the respondents' actual awareness of online spam. Of the respondents, 47% have found pages that are only full of repeated keywords; 39.7% have never experienced this; and 13.3% the respondents chose "DON'T KNOW" as their answer.

Of the respondents, 87% have experienced "*Being redirected to an unrelated page from what was expected.*" Only 9.2% stated that they never had any experience of being redirected to an unrelated page from what was expected and 3.8% chose to answer "DON'T KNOW."

Of the 368 respondents, 64.1% have found pages with repetitive links, 25% have never had any experience in finding pages with repetitive links and 10.9% stated that they do not know.

Of the participating respondents, 78% stated that they have received messages considered as unwanted/suspicious/annoying on a web 2.0 application. Only 10.9% have never received messages considered as unwanted/suspicious/annoying on a web 2.0 application and 11.1% have chosen "DON'T KNOW" as their answer.

Table 6.3: Actual awareness of online spam

Items	Yes		No		Don't know	
	%	Freq.	%	Freq.	%	Freq.
Found pages that are only full of repeated keywords	47	173	39.7	146	13.3	49
Being redirected to an unrelated page from what was expected	87	320	9.2	34	3.8	14
Found pages with repetitive links	64.1	236	25	92	10.9	40
Received message considered as unwanted/suspicious/annoying on a web 2.0 application	78	287	10.9	40	11.1	41
Found pages with unrelated links	82.3	303	13.3	49	4.3	16
Received/seen suspicious link on a web 2.0 application	63	232	17.7	65	19.3	71
Found pages that are only advertising with very little content	77.4	285	15.8	58	6.8	25
Received unwanted postings on their social network account	67.4	248	27.7	102	4.9	18
Received unwanted friend requests on their social network account	77.4	285	18.8	69	3.8	14
Being tagged by unwanted parties on a Web 2.0 application	49.5	182	36.1	133	14.4	53

Of the respondents, 82.3% have found pages with unrelated links, whereas 13.3% have never experienced this. Only 4.3% of the respondents chose "DON'T KNOW" as their answer.

Of the total participants of the survey, 63% have at least once received/seen suspicious links on a web 2.0 application; 17.7% have never received/seen suspicious links on a web 2.0 application. Only 19.3% have chosen “DON’T KNOW” as their answer.

Of the respondents, 77.4% have ever found pages that are only advertising with very little content; 15.8% stated that they have never had any experience of being redirected to an unrelated page from what was expected; and 6.8% chose to answer “DON’T KNOW.”

Of the participating respondents, 67.4% stated that they have received unwanted postings on their social network account. Only 27.7% have never received unwanted postings on their social network account and 4.9% have chosen “DON’T KNOW” as their answer.

Of the respondents, 77.4% have received unwanted friend requests on their social network account, whereas 18.8% have never experienced this and 3.8% of the respondents chose “DON’T KNOW” as their answer.

Only 49.5% of the respondents have the experience of being tagged by unwanted parties on a Web 2.0 application; 36.1% stated that they have never had any experience of being redirected to an unrelated page from what was expected; and 14.4% chose to answer “DON’T KNOW.”

It can be observed from To evaluate the actual awareness of online spam, respondents were asked, “Have you had any of these experiences while browsing the Internet?” “YES,” “NO” and “DON’T KNOW” answers were offered as a choice of answer. Questions presented in this subsection are consistent with subsequent questions in the ‘knowledge section’; thus, there were 13 statements of suspicious activities. Out of this 13, only 10 of them are related to online spam. If users have encountered any spam-related activities identified in any of these 10 statements, then they are translated to having actual awareness. The online spam-related activities are as follows:

- Found pages that are only full of repeated keywords.
- Being redirected to an unrelated page from what was expected.
- Found pages with repetitive links.
- Received message considered as unwanted/suspicious/annoying on a web 2.0 application.
- Found pages with unrelated links.
- Received/seen suspicious link on a web 2.0 application.
- Found pages that are only advertising with very little content.

- Received unwanted postings on their social network account.
- Received unwanted friend requests on their social network account.
- Being tagged by unwanted parties on a Web 2.0 application.

Error! Not a valid bookmark self-reference. recapitulates data on the respondents' actual awareness of online spam. Of the respondents, 47% have found pages that are only full of repeated keywords; 39.7% have never experienced this; and 13.3% the respondents chose "DON'T KNOW" as their answer.

Of the respondents, 87% have experienced "*Being redirected to an unrelated page from what was expected.*" Only 9.2% stated that they never had any experience of being redirected to an unrelated page from what was expected and 3.8% chose to answer "DON'T KNOW."

Of the 368 respondents, 64.1% have found pages with repetitive links, 25% have never had any experience in finding pages with repetitive links and 10.9% stated that they do not know.

Of the participating respondents, 78% stated that they have received messages considered as unwanted/suspicious/annoying on a web 2.0 application. Only 10.9% have never received messages considered as unwanted/suspicious/annoying on a web 2.0 application and 11.1% have chosen "DON'T KNOW" as their answer.

Table 6.3 which covers the actual awareness of online spam that:

- Two of the highest suspicious spam-related activities that the respondents have ever experienced were recorded for the statement "*Being redirected to an unrelated page from what was expected*" (87%), followed by the statement "*Found pages with unrelated links*" (82.3%). The percentages of respondents choosing "DON'T KNOW" were low for both of these questions (3.8% and 4.3%, respectively).
- The highest suspicious spam-related activities that the respondents have never experienced were recorded for the statement "*Found pages that are only full of repeated keywords*" (39.7%) and "*Being tagged by unwanted parties on a Web 2.0 application*" (36.1%).
- The percentages of respondents who chose to answer "DON'T KNOW" to statements such as "*Found pages that are only full of repeated keywords*" (13.3%), "*Being tagged by unwanted parties on a Web 2.0 application*" (14.4%) and "*Received/seen suspicious link on a web 2.0 application*" (19.3%) were relatively high compared to other questions.

6.3.1.3 Level of awareness

In order to determine the level of awareness of Spam 2.0, a score of 1 was given for each right answer for actual awareness (refer to To evaluate the actual awareness of online spam, respondents were asked, “*Have you had any of these experiences while browsing the Internet?*” “YES,” “NO” and “DON’T KNOW” answers were offered as a choice of answer. Questions presented in this subsection are consistent with subsequent questions in the ‘knowledge section’; thus, there were 13 statements of suspicious activities. Out of this 13, only 10 of them are related to online spam. If users have encountered any spam-related activities identified in any of these 10 statements, then they are translated to having actual awareness. The online spam-related activities are as follows:

- Found pages that are only full of repeated keywords.
- Being redirected to an unrelated page from what was expected.
- Found pages with repetitive links.
- Received message considered as unwanted/suspicious/annoying on a web 2.0 application.
- Found pages with unrelated links.
- Received/seen suspicious link on a web 2.0 application.
- Found pages that are only advertising with very little content.
- Received unwanted postings on their social network account.
- Received unwanted friend requests on their social network account.
- Being tagged by unwanted parties on a Web 2.0 application.

Error! Not a valid bookmark self-reference. recapitulates data on the respondents’ actual awareness of online spam. Of the respondents, 47% have found pages that are only full of repeated keywords; 39.7% have never experienced this; and 13.3% the respondents chose “DON’T KNOW” as their answer.

Of the respondents, 87% have experienced “*Being redirected to an unrelated page from what was expected.*” Only 9.2% stated that they never had any experience of being redirected to an unrelated page from what was expected and 3.8% chose to answer “DON’T KNOW.”

Of the 368 respondents, 64.1% have found pages with repetitive links, 25% have never had any experience in finding pages with repetitive links and 10.9% stated that they do not know.

Of the participating respondents, 78% stated that they have received messages considered as unwanted/suspicious/annoying on a web 2.0 application. Only 10.9% have never received messages considered as unwanted/suspicious/annoying on a web 2.0 application and 11.1% have chosen “DON’T KNOW” as their answer.

Table 6.3). The minimum score that one can obtain is 0 and the maximum score is 10. The respondents’ scores are shown in **Error! Not a valid bookmark self-reference..** Actual awareness is the actual awareness level obtained by a respondent, while perceived awareness is the respondents’ self-rated level of awareness.

Table 6.4: Respondents’ score for actual awareness

Score	%	Frequency
0	1.09	4
1	1.90	7
2	1.90	7
3	5.71	21
4	4.62	17
5	9.51	35
6	14.67	54
7	13.86	51
8	16.03	59
9	16.03	59
10	14.67	54

Perceived awareness as shown in To assess the perceived awareness of online spam, respondents were asked the question “*Have you ever heard of online spam?*” Respondents were given the choice to answer either “*yes*” or “*no*.” **Error! Not a valid bookmark self-reference.** shows that 91.6% of the 368 respondents indicated that they have heard of online spam, while only 8.4% of the respondents stated that they have never heard of online spam.

Table 6.2 can be categorized into two, those who have heard of online spam and those who have never heard of online spam. In order to compare between perceived awareness and actual awareness, actual awareness is divided into two categories. Those who score 0 for actual awareness will be compared to those who have never heard of online spam for perceived awareness. Other scores in actual awareness are comparable to those who have heard of online spam. A comparison between perceived awareness

and actual awareness is shown in

Figure 6.1.

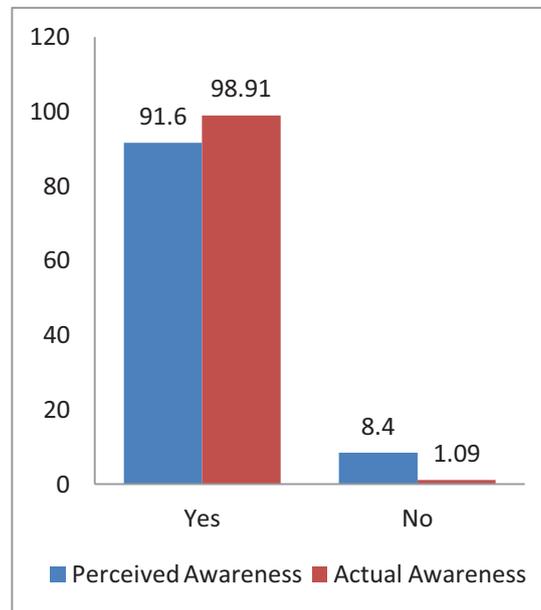


Figure 6.1: Comparison between perceived awareness and actual awareness.

Of the respondents, 8.4% thought that they are not aware of Spam 2.0; however, only 1.09% truly are not aware of Spam 2.0. On the other hand, 91.6% thought they are aware of Spam 2.0; however, a higher percentage of respondents (98.91%) are actually aware of the Spam 2.0 problem.

From the respondents' score tabularized in In order to determine the level of awareness of Spam 2.0, a score of 1 was given for each right answer for actual awareness (refer to To evaluate the actual awareness of online spam, respondents were asked, "Have you had any of these experiences while browsing the Internet?" "YES," "NO" and "DON'T KNOW" answers were offered as a choice of answer. Questions presented in this subsection are consistent with subsequent questions in the 'knowledge section'; thus, there were 13 statements of suspicious activities. Out of this 13, only 10 of them are related to online spam. If users have encountered any spam-related activities identified in any of these 10 statements, then they are translated to having actual awareness. The online spam-related activities are as follows:

- Found pages that are only full of repeated keywords.
- Being redirected to an unrelated page from what was expected.
- Found pages with repetitive links.

- Received message considered as unwanted/suspicious/annoying on a web 2.0 application.
- Found pages with unrelated links.
- Received/seen suspicious link on a web 2.0 application.
- Found pages that are only advertising with very little content.
- Received unwanted postings on their social network account.
- Received unwanted friend requests on their social network account.
- Being tagged by unwanted parties on a Web 2.0 application.

Error! Not a valid bookmark self-reference. recapitulates data on the respondents' actual awareness of online spam. Of the respondents, 47% have found pages that are only full of repeated keywords; 39.7% have never experienced this; and 13.3% the respondents chose "DON'T KNOW" as their answer.

Of the respondents, 87% have experienced "*Being redirected to an unrelated page from what was expected.*" Only 9.2% stated that they never had any experience of being redirected to an unrelated page from what was expected and 3.8% chose to answer "DON'T KNOW."

Of the 368 respondents, 64.1% have found pages with repetitive links, 25% have never had any experience in finding pages with repetitive links and 10.9% stated that they do not know.

Of the participating respondents, 78% stated that they have received messages considered as unwanted/suspicious/annoying on a web 2.0 application. Only 10.9% have never received messages considered as unwanted/suspicious/annoying on a web 2.0 application and 11.1% have chosen "DON'T KNOW" as their answer.

Table 6.3). The minimum score that one can obtain is 0 and the maximum score is 10. The respondents' scores are shown in **Error! Not a valid bookmark self-reference..** Actual awareness is the actual awareness level obtained by a respondent, while perceived awareness is the respondents' self-rated level of awareness.

Table 6.4, the level of awareness was categorized into five categories: very low, low, intermediate, high and very high. Scores between 0 and 2 will fall into very low, 3 and 4 into low, 5 and 6 into intermediate, 7 and 8 into high and between 9 and 10 into very high. The levels of awareness for all respondents are recapitulated in **Error! Not a valid bookmark self-reference..**

Based on **Error! Not a valid bookmark self-reference.**, only 4.9% of the total respondents are categorized as having a very low level of awareness. The low categories include 10.3% of the respondents. Of the respondents, 24.2% are categorized under the intermediate level of awareness, 29.9% are categorized as having a high level of awareness and 30.7% are considered to have a very high level of awareness.

Table 6.5: Respondents level of awareness

Level of awareness	%	Frequency
Very low	4.9	18
Low	10.3	38
Intermediate	24.2	89
High	29.9	110
Very high	30.7	113

From the tables and figure presented in this subsection, it can be observed that:

- There is a small percentage difference between perceived awareness and actual awareness. A smaller percentage of respondents thought that they are not aware of Spam 2.0.
- The highest percentage for the respondents' level of awareness is recorded for the very high category with 30.7%.
- The lowest percentage for the respondents' level of awareness is recorded for the very low category with 4.9%.

6.3.2 Knowledge of Spam 2.0

The public knowledge of Spam 2.0 includes two categories of knowledge, perceived knowledge and actual knowledge. Descriptive statistics on perceived knowledge and actual knowledge are explained separately in Section 6.3.2.1 and Section 6.3.2.2. This section reports data on knowledge-related questions obtained in Part B of the survey as mentioned in Section 6.3.2.1. In addition, survey data from Part C is also included in the Actual Knowledge III assessment. Section 6.3.2.3 focuses on the analysis of the level of knowledge.

6.3.2.1 Perceived knowledge

Perceived knowledge is defined as the respondent's self-rated level of knowledge. The participants' perceived knowledge was measured based on their rating of their own overall knowledge of online spam: none, poor, fair, good or expert. The results for perceived knowledge of online spam tabulated

in **Error! Not a valid bookmark self-reference.** show that only 1.6% of respondents stated that they have no knowledge at all about online spam. Of the respondents, 26.4% claimed that they have poor knowledge on online spam. More than half of the respondents (53%) considered themselves as having fair knowledge, while the respondents who considered themselves as having a good knowledge of online spam were reported to be 15.8%. Only 12 respondents equalling to 3.3% rated their knowledge of online spam as expert.

Table 6.6: Perceived knowledge of online spam

Items	My overall knowledge of online spam can best be described as	
	%	Frequency
None	1.6	6
Poor	26.4	97
Fair	53.0	195
Good	15.8	58
Expert	3.3	12

- Through the results presented in Perceived knowledge is defined as the respondent's self-rated level of knowledge. The participants' perceived knowledge was measured based on their rating of their own overall knowledge of online spam: none, poor, fair, good or expert. The results for perceived knowledge of online spam tabulated in **Error! Not a valid bookmark self-reference.** show that only 1.6% of respondents stated that they have no knowledge at all about online spam. Of the respondents, 26.4% claimed that they have poor knowledge on online spam. More than half of the respondents (53%) considered themselves as having fair knowledge, while the respondents who considered themselves as having a good knowledge of online spam were reported to be 15.8%. Only 12 respondents equalling to 3.3% rated their knowledge of online spam as expert.

Table 6.6, it is observed that:

- The perceived knowledge for more than half of the total respondents is fair (53%).
- Only a small percentage of respondents rated themselves as having no knowledge at all (1.6%) and being expert (3.3%).

6.3.2.2 Actual knowledge

Actual knowledge is the actual knowledge level obtained by a respondent. Actual knowledge is assessed using three sets of questions, asked to respondents:

- Actual Knowledge I: Which of these actions do you think is considered as online spam?
- Actual Knowledge II: Assess the statement below and please choose the appropriate response for each item.
- Actual Knowledge III: Do you think the screen below contains spam?

Actual Knowledge I

Actual knowledge in this question is evaluated through statements presented in this question. These 13 statements were a list of suspicious activities that could happen to respondents when using the Internet. Respondents then need to decide if they consider these activities as online spam or not. Respondents are also allowed to opt for “DON’T KNOW” as their answer. These statements are as follows:

- Found pages that are only full of repeated keywords.
- Being redirected to an unrelated page from what was expected.
- Found pages with repetitive links.
- Received message considered as unwanted/suspicious/annoying on a web 2.0 application.
- Being hacked/hijacked is an action of online spam.
- Found pages with unrelated links.
- Attacked by virus.
- Received/seen suspicious link on a web 2.0 application.
- Found pages that are only advertising with very little content.
- Being asked passwords and/or credit card details by a trustworthy-looking entity in the Internet communication.
- Received unwanted postings on their social network account.
- Received unwanted friend requests on their social network account.
- Being tagged by unwanted parties on a Web 2.0 application.

Error! Not a valid bookmark self-reference. sums up data on respondents’ replies in terms of correctness for 13 statements on online spam-related activities.

For the statement “*Found pages that are only full of repeated keywords,*” 55.2% of the respondents managed to answer the question correctly, whereas 24.7% answered this incorrectly. Another 20.1% of the respondents chose “DON’T KNOW” as their answer.

Table 6.7: Respondents' data to assess knowledge on online spam-related activities

Items	Correct		Incorrect		Don't know	
	%	Freq.	%	Freq.	%	Freq.
Found pages that are only full of repeated keywords	55.2	203	24.7	91	20.1	74
Being redirected to an unrelated page from what was expected	79.6	293	13.9	51	6.5	24
Found pages with repetitive links	65.8	242	19.3	71	14.9	55
Received message considered as unwanted/suspicious/annoying on a web 2.0 application	81.8	301	9.2	34	9.0	33
Being hacked/hijacked is an action of online spam	36.1	133	57.3	211	6.5	24
Found pages with unrelated links	76.6	282	17.4	64	6	22
Attacked by virus	26.9	99	67.7	249	5.4	20
Received/seen suspicious link on a web 2.0 application	80.2	295	7.9	29	12	44
Found pages that are only advertising with very little content	60.1	221	26.6	98	13.3	49
Being asked passwords and/or credit card details by a trustworthy-looking entity in the Internet communication	23.4	86	68.8	253	7.9	29
Received unwanted postings on their social network account	84.8	312	10.9	40	4.3	16
Received unwanted friend requests on their social network account	72.8	268	20.4	20.4	6.8	25
Being tagged by unwanted parties on a Web 2.0 application	68.5	252	17.4	64	14.1	52

Of the respondents, 79.6% managed to assess the statement “*Being redirected to an unrelated page from what was expected*” correctly. Only 13.9% assessed this statement incorrectly and 6.5% stated that they do not know the answer to this statement.

Of the total respondents, 65.8% assessed the statement “*Found pages with repetitive links*” and managed to answer them correctly, while 19.3% got it wrong and 14.9% chose to answer “DON’T KNOW.”

Of the participating respondents, 81.8% managed to answer correctly for the statement “*Received message considered as unwanted/suspicious/annoying on a web 2.0 application,*” whereas 9.2% answered this statement incorrectly. Only 9.0% of the respondents stated that they do not know the answer to this statement.

Only 36.2% of total respondents managed to answer correctly for the statement “*Being hacked/hijacked is an action of online spam.*” More than half of the respondents (57.3%) incorrectly assessed this statement, whereas 6.5% chose “DON’T KNOW” as their answer.

For the statement “*Found pages with unrelated links,*” 76.6% of the respondents managed to answer the question correctly, whereas 17.4% answered this incorrectly. Only 6% of the respondents chose “DON’T KNOW” as their answer.

Only 26.9% of the total respondents managed to answer correctly for the statement “*Attacked by virus,*” whereas 67.7% of the respondents incorrectly assessed this statement and 5.4% chose “DON’T KNOW” as their answer.

Of the respondents, 80.2% managed to recognize “*Received/seen suspicious link on a web 2.0 application*” as one of the actions of online spam. Only 7.9% incorrectly answered this statement and 12% stated that they do not know the answer.

Of the respondents, 60.1% correctly chose the right answer in identifying “*Found pages that are only advertising with very little content*” as the online spam actions. Only 26.6% of respondents made the wrong assessment and 13.3% chose to answer “DON’T KNOW.”

Only 23.4% of the respondents managed to assess the statement “*Being asked passwords and/or credit card details by a trustworthy-looking entity in the Internet communication*” correctly. Of the respondents, 68.8% assessed this statement incorrectly and 7.9% stated that they do not know the answer to this statement.

Of the total respondents, 84.8% assessed the statement and correctly considered “*Received unwanted postings on their social network account*” as one of the online spam actions, while 10.9% got it wrong and 4.3% chose to answer “DON’T KNOW.”

Of the participating respondents, 72.8% managed to correctly recognize the statement “*Received unwanted friend requests on their social network account*” as one of the online spam actions, whereas

20.4% incorrectly answered this statement. Another 6.8% of the respondents opted to answer “DON’T KNOW.”

For the last statement in this question, respondents were asked to assess whether “*Being tagged by unwanted parties on a Web 2.0 application*” is considered as online spamming or not. Of the respondents, 68.5% correctly identified this and 17.4% got it wrong, while 14.1% stated that they do not know the answer.

- Through the results presented in Error! **Not a valid bookmark self-reference.** sums up data on respondents’ replies in terms of correctness for 13 statements on online spam-related activities.

For the statement “*Found pages that are only full of repeated keywords,*” 55.2% of the respondents managed to answer the question correctly, whereas 24.7% answered this incorrectly. Another 20.1% of the respondents chose “DON’T KNOW” as their answer.

Table 6.7, it is observed that:

- Three of the actions that were identified correctly with the highest percentage (more than 80%) were “*Received unwanted postings on their social network account,*” “*Received message considered as unwanted/suspicious/annoying on a web 2.0 application*” and “*Received/seen suspicious link on a web 2.0 application.*” The percentages for respondents who answered this correctly are 84.8, 81.8 and 80.2, respectively.
- The actions that were identified correctly with the lowest percentage (less than 30%) were “*Attacked by virus*” and “*Being asked passwords and/or credit card details by a trustworthy-looking entity in the Internet communication.*” The percentages for respondents who answered this correctly are 26.9 and 23.4, respectively. These statements were also recorded for the actions that were identified wrongly with the highest percentage (more than 60%).
- The action with the highest percentage that the respondents chose to answer “DON’T KNOW” is “*Found pages that are only full of repeated keywords*” reported at 20.1%.
- The actions that the respondents had the most misconceptions or have no knowledge about are “*Attacked by virus*” and “*Being asked passwords and/or credit card details by a trustworthy-looking entity in the Internet communication*” with 73.1% and 76.7%, respectively.

Actual Knowledge II

Respondents were asked to assess 10 statements on online spam facts and choose the appropriate response for each item. For this question, respondents were given the option to choose “TRUE,” “NOT TRUE” and “DON’T KNOW” as their answer. Most of the questions are basic knowledge and categorized as easy; however, there are two technical questions which can be categorized as tough

questions. The statements cover the causes, impacts, suitable solutions or related information about online spam. Ten statements on online spam used in this question are as follows:

- Online spam can be used as a part of phishing attack.
- Online spam can be used to disseminate malware.
- Online spam can be used to promote affiliate websites.
- There is no difference between online spam and email spam.
- All online spam can be detected and treated using existing anti-spam techniques that are designed for email spam.
- Online spam has lower viewer impact than email spam.
- Auto registration software can be used to register spam accounts.
- Online spam can be found on legitimate websites.
- Online spam can be used to provide false information to users.
- Online spam can lead to other crimes such as fraud.

The respondents' correctness in answering these statements is then summarized in Table 6.8.

Table 6.8: Respondents' data to assess knowledge on online spam.

Items	Correct		Incorrect		Don't know	
	%	Freq.	%	Freq.	%	Freq.
Online spam can be used as a part of phishing attack	77.2	284	2.7	10	20.1	74
Online spam can be used to disseminate malware	72	265	7.2	26	20.9	77
Online spam can be used to promote affiliate websites	67.9	250	13	48	19	70
There is no difference between online spam and email spam	36.4	134	28.8	106	34.8	128
All online spam can be detected and treated using existing anti-spam techniques that are designed for email spam	25	92	35.6	131	39.4	145

Online spam has lower viewer impact than email spam	29.9	110	25.5	94	44.6	164
Auto registration software can be used to register spam accounts	45.4	167	5.2	19	49.5	182
Online spam can be found on legitimate websites	56.5	208	13.3	49	30.2	111
Online spam can be used to provide false information to users	77.7	286	5.2	19	17.1	63
Online spam can lead to other crimes such as fraud	81.5	300	3	11	15.5	57

Of the respondents, 77.2% have managed to assess the statement “*Online spam can be used as a part of phishing attack*” correctly. Only 2.7% of the respondents have incorrectly answered this statement and 20.1% stated that they do not know the answer.

Of the respondents, 72% have correctly assessed the statement “*Online spam can be used to disseminate malware,*” while 7.2% have made the wrong assessment and another 20.9% chose to answer “DON’T KNOW.”

Of the respondents, 67.9% managed to assess the statement of “*Online spam can be used to promote affiliate websites*” correctly; 13% assessed this statement incorrectly; and 19% stated that they do not know the answer to this statement.

Of the total respondents, 36.4% decided that the statement of “*There is no difference between online spam and email spam*” is wrong and correctly assessed this statement, whereas 28.8% incorrectly assessed this statement and 34.8% chose to answer “DON’T KNOW.”

For the technical statement “*All online spam can be detected and treated using existing anti-spam techniques that are designed for email spam,*” only 25% of the respondents managed to assess correctly, whereas 35.6% incorrectly answered this statement. Another 39.4% of the respondents have opted to answer “DON’T KNOW.”

Of the total respondents, 29.9% have correctly assessed the technical statement of “*Online spam has lower viewer impact than email spam,*” while 25.5% have made the wrong assessment and another 44.6% chose to answer “DON’T KNOW.”

Through the results presented in Table 6.8, it is apparent that:

- The statement that was assessed correctly with the highest percentage (more than 80%) was “*Online spam can lead to other crimes such as fraud.*”

- The statements that were assessed correctly with the lowest percentage (less than 30%) were “*All online spam can be detected and treated using existing anti-spam techniques that are designed for email spam*” and “*Online spam has lower viewer impact than email spam.*” The percentages for respondents who have answered this correctly are only 25% and 29.9%, respectively.
- The statements that the respondents chose to answer “DON’T KNOW” with percentages more than 30% are “*Auto registration software can be used to register spam accounts,*” “*Online spam has lower viewer impact than email spam,*” “*All online spam can be detected and treated using existing anti-spam techniques that are designed for email spam*” and “*There is no difference between online spam and email spam.*” The percentages for respondents who have chosen to answer “DON’T KNOW” for these statements are 49.5%, 44.6%, 39.4% and 34.8%, respectively.
- More than half of the respondents have misconceptions and do not have knowledge on the statements “*Auto registration software can be used to register spam accounts,*” “*There is no difference between online spam and email spam,*” “*Online spam has lower viewer impact than email spam*” and “*All online spam can be detected and treated using existing anti-spam techniques that are designed for email spam.*” The corresponding percentages for all statements were 54.7%, 63.6%, 70.1% and 75%, accordingly.

Actual Knowledge III

This section specifically reports data on spam identification questions obtained in Part C of the survey as mentioned in Section 0. Respondents were asked to assess 10 examples of spam and identify if it is spam or not. Thus, in these questions, the respondents’ actual knowledge is assessed indirectly. The respondents are allowed to choose “YES,” “NO” or “DON’T KNOW” as their answer. The respondents’ correctness in identifying spam is then summarized in Table 6.9.

Of the respondents, 61.7% have managed to answer Example 1 correctly by identifying it as spam, whereas 11.1% have incorrectly answered this question and another 27.2% stated that they do not know the answer.

For Example 2, 70.9% have correctly identified it as spam, while 8.7% of respondents have made the wrong assessment and another 20.4% chose to answer “DON’T KNOW.”

Only 23.6% of the respondents managed to correctly identify Example 3 as non-spam, whereas 42.1% made the wrong decision in identifying the example as spam and another 34.2% stated that they do not know the answer to this question.

Of the total respondents, 38% correctly identified Example 4 as non-spam; 32.1% incorrectly identified the example as spam; and 29.9% chose to answer “DON’T KNOW.”

For Example 5, 27.7% of the respondents managed to correctly assess the spam identification, whereas 52.4% failed to answer this question correctly. Another 19.8% of the respondents have opted to answer “DON’T KNOW.”

Table 6.9: Respondents’ data to assess knowledge on spam identification

Items	Correct		Incorrect		Don’t Know	
	%	Freq.	%	Freq.	%	Freq.
Example 1	61.7	227	11.1	41	27.2	100
Example 2	70.9	261	8.7	32	20.4	75
Example 3	23.6	87	42.1	155	34.2	126
Example 4	38.0	140	32.1	140	29.9	110
Example 5	27.7	102	52.4	193	19.8	73
Example 6	27.7	102	54.9	202	17.4	64
Example 7	81.3	299	6.3	23	12.5	46
Example 8	91.0	335	1.9	7	7.1	26
Example 9	26.4	97	41.3	152	32.3	119
Example 10	47.3	174	31.3	115	21.5	79

Example 6 presents a screen capture of non-spam. Only 27.7% of the respondents have correctly identified it as non-spam, while 54.9% have made the wrong assessment and 17.4% chose to answer “DON’T KNOW.”

Of the respondents, 81.3% managed to identify Example 7 as spam correctly. Only 6.3% answered this question incorrectly and 12.5% stated that they do not know the answer.

For Example 8, 91% of the respondents managed to correctly identify it as spam. Only 1.9% of them have mistakenly answered this question and 7.1% chose to answer “DON’T KNOW.”

Of the total respondents, 26.4% correctly identified Example 9 as spam; 41.3% incorrectly identified the example as non-spam; and 32.3% stated that they do not know the answer to this question.

For the final question in this subsection, Example 10 is a non-spam. Of the respondents, 47.3% correctly made the right identification while 31.3% incorrectly identified it as spam. Another 21.5% of the respondents have opted to answer “DON’T KNOW.”

Through the results presented in Table 6.9, it is observed that:

- The examples that were answered correctly with the highest percentage (more than 80%) were Examples 8 and 7 with 91% and 81.3%, respectively.
- The examples that were assessed correctly with the lowest percentage (less than 30%) were Examples 3, 5, 6 and 9. The percentages for respondents who have answered this correctly are only 23.6%, 27.7%, 27.7% and 26.4%, respectively.
- More than half of the respondents have answered Examples 5 and 6 incorrectly with the percentages recorded at 52.4% and 54.9%, respectively.
- The highest percentages for respondents who have chosen to answer “DON’T KNOW” were recorded for Example 3 (34.2%) and Example 9 (32.3%).

6.3.2.3 Level of Knowledge

In order to determine the level of knowledge of Spam 2.0, a score of 1 was given for each right answer given for Actual Knowledge I, II and III (refer to Error! **Not a valid bookmark self-reference.** sums up data on respondents’ replies in terms of correctness for 13 statements on online spam-related activities.

For the statement “*Found pages that are only full of repeated keywords,*” 55.2% of the respondents managed to answer the question correctly, whereas 24.7% answered this incorrectly. Another 20.1% of the respondents chose “DON’T KNOW” as their answer.

Table 6.7, Table 6.8 and Table 6.9). The minimum score that one can obtain is 0 and the maximum score is 33. From these respondents’ score, the level of awareness was categorized into five categories: none, poor, fair, good and expert so that it is comparable to the five levels of perceived knowledge. Scores between 0 and 6 will fall into none, 7–13 into poor, 14–20 into fair, 21–27 into good and 28–33 into expert. The levels of awareness for all respondents are recapitulated in **Error! Not a valid bookmark self-reference.**

Table 6.10: Respondents’ level of knowledge

Actual Knowledge	%	Frequency
None	0.82	3
Poor	12.23	45
Fair	48.10	177
Good	37.50	138
Expert	1.36	5

Based on In order to determine the level of knowledge of Spam 2.0, a score of 1 was given for each right answer given for Actual Knowledge I, II and III (refer to Error! **Not a valid bookmark self-reference.** sums up data on respondents' replies in terms of correctness for 13 statements on online spam-related activities.

For the statement "*Found pages that are only full of repeated keywords,*" 55.2% of the respondents managed to answer the question correctly, whereas 24.7% answered this incorrectly. Another 20.1% of the respondents chose "DON'T KNOW" as their answer.

Table 6.7, Table 6.8 and Table 6.9). The minimum score that one can obtain is 0 and the maximum score is 33. From these respondents' score, the level of awareness was categorized into five categories: none, poor, fair, good and expert so that it is comparable to the five levels of perceived knowledge. Scores between 0 and 6 will fall into none, 7–13 into poor, 14–20 into fair, 21–27 into good and 28–33 into expert. The levels of awareness for all respondents are recapitulated in Error! **Not a valid bookmark self-reference.**

Table 6.10, only 0.82% of the total respondents are categorized as "none." The "poor" category includes 12.23% of the respondents. Of the respondents, 48.10% are categorized under a "fair" level of knowledge, whereas 37.5% are categorized as having a "good" level of knowledge. Only 1.36% of the respondents are considered "expert."

The thesis now compares perceived knowledge with level of knowledge. Perceived knowledge as shown in Perceived knowledge is defined as the respondent's self-rated level of knowledge. The participants' perceived knowledge was measured based on their rating of their own overall knowledge of online spam: none, poor, fair, good or expert. The results for perceived knowledge of online spam tabulated in Error! **Not a valid bookmark self-reference.** show that only 1.6% of respondents stated that they have no knowledge at all about online spam. Of the respondents, 26.4% claimed that they have poor knowledge on online spam. More than half of the respondents (53%) considered themselves as having fair knowledge, while the respondents who considered themselves as having a good knowledge of online spam were reported to be 15.8%. Only 12 respondents equalling to 3.3% rated their knowledge of online spam as expert.

Table 6.6 was categorized into five categories; thus, both of them are comparable. The comparison between perceived knowledge and level of knowledge is shown in Figure 6.2.

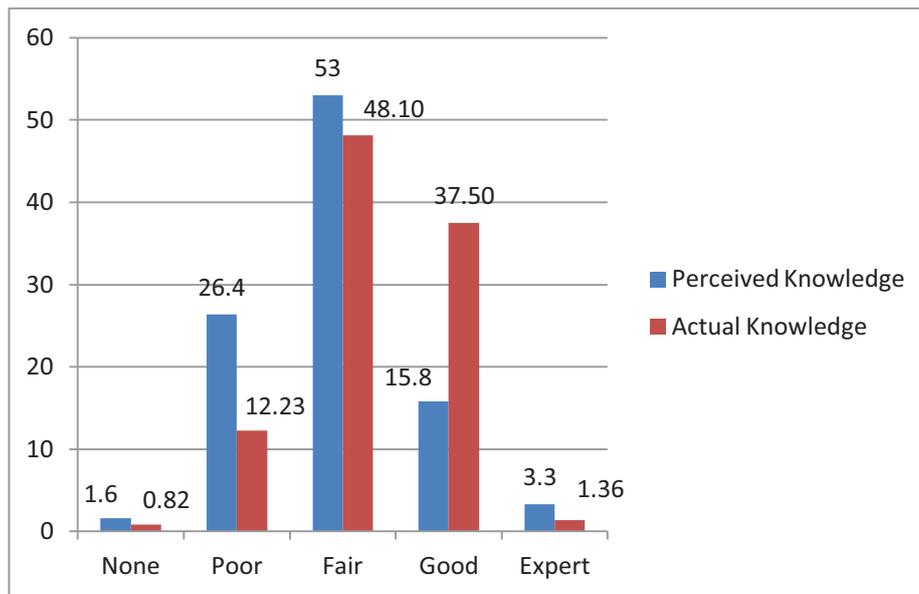


Figure 6.2: Comparison between perceived knowledge and level of knowledge

The perceived knowledge of 1.6% of the total respondents was categorized under “none”; however, a lower percentage (0.82%) was categorized under the same category for actual knowledge. For the “poor” category, 26.4% perceived knowledge was recorded for the total respondents, while for actual knowledge, it includes only 12.23%. Of the total respondents, 53% were categorized as having a “fair” level of perceived knowledge. However, for actual knowledge, there is a difference of approximately 5% totalling to 48.1% of respondents categorized under the “fair” category.

There is a big difference (in percentage) between perceived knowledge and actual knowledge categorized under the “good” category with 15.8% for perceived knowledge and 37.5% for actual knowledge. For the “expert” group, 3.3% of the respondents thought they are “expert.” However, only 1.36% is categorized for having an “expert” level of knowledge.

From Table 6.13 and Figure 6.2 presented in this subsection, it can be observed that:

- The highest percentage for respondents’ level of knowledge is recorded for the “fair” category with 40.1%.
- Both the “none” and “expert” category levels include the lowest percentages of respondents, which are 0.82% and 1.36%, respectively.
- All categories except “good” show a lower percentage for perceived knowledge compared to actual level of knowledge.

6.3.3 Perception of Spam 2.0

This section reports data on perception-related questions obtained in Part B of the survey as mentioned in Section 0. In this section, the thesis provides descriptive results for questions related to perception discretely. This includes perception towards crime, perception towards crime's punishment, fear of crime, perception towards seriousness of crime and perception towards crime's motivation.

6.3.3.1 Perception towards crime

Perception towards crime is assessed through two questions. For both of these questions, respondents are allowed to opt for "YES," "NO" and "MAYBE" answer. The questions asked to respondents were as follows:

- Do you think that online spam is a problem?
- Do you think that spamming is acceptable?

Error! Not a valid bookmark self-reference. recapitulates data on the respondents' perception towards the crime of online spam. The majority of respondents think that online spam is a problem (77.2%). Only 2.4% think that online spam is not a problem. Another 20.4% have chosen the answer "MAYBE."

Table 6.11: Percentage of respondents' perception towards crime

Items	Yes		No		Maybe	
	%	Freq.	%	Freq.	%	Freq.
Do you think that online spam is a problem?	77.2	284	2.4	9	20.4	75
Do you think that spamming is acceptable?	6.3	23	72.8	268	20.9	77

It is shown in Perception towards crime is assessed through two questions. For both of these questions, respondents are allowed to opt for "YES," "NO" and "MAYBE" answer. The questions asked to respondents were as follows:

- Do you think that online spam is a problem?

- Do you think that spamming is acceptable?

Error! Not a valid bookmark self-reference. recapitulates data on the respondents' perception towards the crime of online spam. The majority of respondents think that online spam is a problem (77.2%). Only 2.4% think that online spam is not a problem. Another 20.4% have chosen the answer "MAYBE."

Table 6.11 that 72.8% of the respondents think that spamming is not acceptable. Another 6.3% think that spamming is acceptable, whereas 20.9% have answered "MAYBE."

Through the results presented in Perception towards crime is assessed through two questions. For both of these questions, respondents are allowed to opt for "YES," "NO" and "MAYBE" answer. The questions asked to respondents were as follows:

- Do you think that online spam is a problem?
- Do you think that spamming is acceptable?

Error! Not a valid bookmark self-reference. recapitulates data on the respondents' perception towards the crime of online spam. The majority of respondents think that online spam is a problem (77.2%). Only 2.4% think that online spam is not a problem. Another 20.4% have chosen the answer "MAYBE."

Table 6.11, it is observed that most respondents think that online spam is a problem and not acceptable.

6.3.3.2 Perception towards crime's punishment

The perception towards crime's punishment is evaluated through two questions as follows:

- Do you think that confessed spammers should be punished?
- Do you think that convicted spammers should be allowed to work in the computing field?

For both of these questions, the respondents are allowed to opt for "YES," "NO" or "MAYBE" answer. The results for these questions are revealed in **Error! Not a valid bookmark self-reference..**

Table 6.12: Percentage of respondents' perception on crime's punishment

Items	Yes		No		Maybe	
	%	Freq.	%	Freq.	%	Freq.
Do you think that confessed spammers should be punished?	57.1	210	4.9	18	37.8	139
Do you think that convicted spammers should be allowed to work in computing field?	24.5	90	35.5	130	39.9	147

Of the total respondents, 57.1% think that confessed spammers should be punished; 4.9% think that they should not be punished; and 37.8% have chosen "MAYBE" as their answer.

Only 24.5% of the respondents think convicted spammers should be allowed to work in the computing field; 35.5% stated that they think the convicted spammers should not be allowed to work in the computing field; and 39.9% have chosen the answer "MAYBE."

Through the results presented in The perception towards crime's punishment is evaluated through two questions as follows:

- Do you think that confessed spammers should be punished?
- Do you think that convicted spammers should be allowed to work in the computing field?

For both of these questions, the respondents are allowed to opt for "YES," "NO" or "MAYBE" answer. The results for these questions are revealed in **Error! Not a valid bookmark self-reference..**

Table 6.12, it is observed that:

- Half of the respondents agreed that confessed spammers should be punished. Only a smaller percentage of respondents agreed that convicted spammers should be allowed to work in the computing field.
- Nearly a quarter of the total respondents think that convicted spammers should be allowed to work in the computing field.
- Nearly two-fifths of the respondents could not give a definite answer on punishing spammers.

6.3.3.3 Fear of crime

The fear of crime is evaluated through two questions as follows:

- How vulnerable are you to spam?
- What is the likelihood that you will be spammed?

For the question “*How vulnerable are you to spam?*”, the respondents were asked to choose their answer based on a 5-point Likert scale of how vulnerable they think they are. The responses are presented in Table 6.13. Of the total participants of the survey, 12.2% have chosen very vulnerable; 41.3% felt somewhat vulnerable; 21.2% felt indifferent; and 22.6% have chosen not very vulnerable as their answer. Only 2.7% have chosen not vulnerable at all.

Table 6.13: Respondents’ perceived vulnerability to spam

Items	How vulnerable are you to spam?	
	%	Frequency
Very vulnerable	12.2	45
Somewhat vulnerable	41.3	152
Indifferent.	21.2	78
Not very vulnerable	22.6	83
Not vulnerable at all	2.7	10

Table 6.14 presents the data for another question to assess the respondents’ fear of crime. The question asked to the respondents was “*What is the likelihood that you will be spammed?*” Of 368 respondents, 11.4% stated very unlikely; 21.7% stated that it is unlikely for them to be spammed; 32.9% stated that it is undecided; and 27.4% felt that it is likely for them to be spammed in the future. Only 6.5% of the respondents think that it is very likely for them to be spammed.

Table 6.14: Respondents’ perceived likelihood of being spammed

Items	What is the likelihood that you will be spammed?	
	%	Frequency
Very unlikely	11.4	42
Unlikely	21.7	80
Undecided	32.9	121
Likely	27.4	101
Very likely	6.5	24

From Table 6.13 and Table 6.14 on reported data from questions related to the fear of crime, it can be observed that:

- More than half of the respondents (51.3%) think that they are vulnerable to spam.
- Approximately a quarter of the respondents think that they are not vulnerable to spam.
- Of the respondents, 33.9% think that there is a high likelihood of being spammed. Nearly the same percentage of respondents (33.1%) thinks that there is a low likelihood of being spammed.

6.3.3.4 Perception towards seriousness of crime

The perception towards the seriousness of crime is evaluated based on the question where the respondents are asked to indicate how strongly they disagree or agree with eight statements. These statements are listed as follows:

- Virus is a serious problem
- Hacking is a serious problem.
- Phishing is a serious problem.
- Computer fraud is a serious problem.
- Online spam is a serious problem.
- Email spam is a serious problem.
- Spyware is a serious problem.
- Worm is a serious problem.

The purpose of this question is to evaluate how serious the problem of online spam is seen by Internet users compared to other computer crimes. For each statement, respondents need to give a value on a scale of 1–7 where 1 means strongly disagree and 7 means strongly agree. The results are presented in Table 6.15.

A higher average value means more people rated the statement at a higher number. The higher number (closer to 7 as the maximum) means the stronger respondents agree that it is a serious problem. Thus, our focus is only on the average values. There was also not much difference in the values of standard deviation.

Table 6.15: Respondents' perception on seriousness of computer security problem

Statement	Average value	Standard deviation
Virus is a serious problem.	6.5353	.98149
Hacking is a serious problem.	6.5272	1.04103
Phishing is a serious problem.	6.1495	1.16813
Computer fraud is a serious problem.	6.3723	1.09991
Online spam is a serious problem.	5.9402	1.25353
Email spam is a serious problem.	5.75	1.40745
Spyware is a serious problem.	6.0462	1.17677
Worm is a serious problem.	6.0598	1.26651

For the statement “*Virus is a serious problem,*” the average value is calculated as 6.5353. Average values of 6.5272, 6.1495 and 6.3723 are recorded for the seriousness of hacking, phishing and computer fraud, accordingly. An average value of 5.9402 is obtained for the statement “*Online spam is a problem.*” For email spam, the average value rated by the respondents was 5.75. In rating the statement “*Spyware is a serious problem,*” the average value is calculated as 6.0462. The average value of rating from respondents for the worm seriousness problem is 6.0598.

From Table 6.15, it is observed that:

- The rank for seriousness of a computer crime problem is virus → hacking → computer fraud → phishing → worm → spyware → online spam → email spam.

6.3.3.5 Perception towards crime's motivation

The perception towards crime's motivation is evaluated based on the question where the respondents are asked to indicate how relevant a statement is towards spammers' motives. These statements are listed as follows:

- To make money.
- For malicious reason.
- To obtain a higher rank in search engine.
- To promote their product and services.
- For religious purposes.
- For fun.

The purpose of this question is to evaluate spammers' motivation as thought by the Internet users. For each statement, respondents need to give a value on a scale of 1–6 where 1 means the most relevant answer and 6 means the most irrelevant answer. The results are recapitulated in **Error! Not a valid bookmark self-reference..**

Table 6.16: Respondents' perception on crime's motivation

Statement	Average value	Standard deviation
To make money.	2.3342	1.56415
For malicious reason.	3.1467	1.42771
To obtain a higher rank in search engine.	3.625	1.27462
To promote their product and services.	3.1033	1.3165
For religious purposes.	5.1413	1.499
For fun.	3.6495	1.75299

For this question, a lower average value means more people rated the statement at a lower number. The lower number (1 minimum and 6 maximum) means the stronger respondents agree that a motive is relevant. Thus, our focus is only on the average values. There was also not much difference in the values of standard deviation.

For the statement "*To make money*," the average value is calculated as 2.3342. Average values of 3.1467, 3.625 and 3.1033 are calculated for the statement "*For malicious reason*," "*To obtain a higher rank in search engine*" and "*To promote their product and services*," accordingly. An average value of 5.1413 is obtained for the statement "*For religious purposes*." In rating the statement "*For fun*" as the motives of spamming, the average value is calculated as 3.6495.

Through the results presented in The purpose of this question is to evaluate spammers' motivation as thought by the Internet users. For each statement, respondents need to give a value on a scale of 1–6 where 1 means the most relevant answer and 6 means the most irrelevant answer. The results are recapitulated in **Error! Not a valid bookmark self-reference..**

Table 6.16, it is observed that:

- Based on respondents' perception, the rank for spammers' motivation is as follows:
 - To make money.
 - To promote their product and services.
 - For malicious reason.

- To obtain a higher rank in search engine.
- For fun.
- For religious purposes.

6.3.4 Justification Comments

This section reports data gathered from comments given on the questions in Part C of the survey as mentioned in Section 0. Comments analysis is done based on respondents' justifications on spam examples in Spam 2.0 identification. The thesis first provides basic statistics on the number of comments recorded for each question in Section 0. Based on these comments, the research first identifies and classifies the spam characteristics based on the justifications in Section 0 and non-spam characteristics in Section 6.3.4.2. Section 0 presents comments analysis according to 10 spam examples.

6.3.4.1 Basic statistics on frequency of comments

In total, there were 1,058 comments given by the participants. Figure 6.3 shows the basic statistics for comments that were given by the respondents to justify their answers.

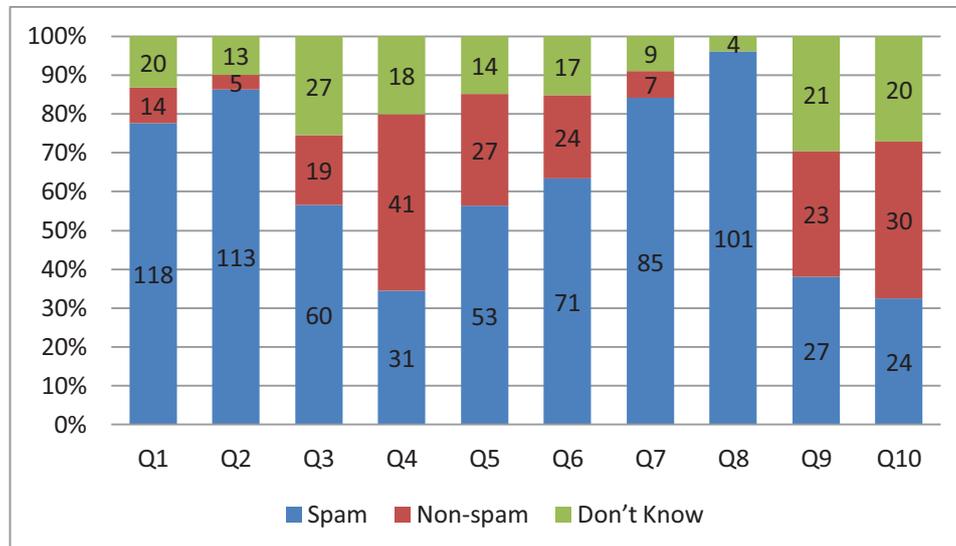


Figure 6.3: Basic statistics for comments

Based on Figure 6.3, there were a total of 152 comments for Example 1. Of these, 118 comments were given to justify why respondents think the answer is spam and 14 comments for non-spam answers' justification. The remaining 20 comments explained why they chose "DON'T KNOW" as their answer.

For Question 2, 131 comments were given by respondents with 113 justifying why they think Example 2 is spam. Five comments were given to justify why it is non-spam and 13 for answering “DON’T KNOW.”

A total of 106 comments were given by respondents for Example 3, of which 60 were to justify the answer why they thought it is spam, 19 for non-spam and 27 comments for “DON’T KNOW.”

For Example 4, there were only 90 comments left by respondents, of which 31 comments were to justify why respondents have answered spam, 41 for the non-spam answer and 18 for answering “DON’T KNOW.”

There were a total of 94 comments for Example 5 with 53 comments given to justify the answer for spam. Another 27 comments were to justify the non-spam answer and another 14 comments for answering “DON’T KNOW.”

For Example 6, there were 71 comments left to justify spam, 24 comments to justify non-spam and 17 comments to justify their “DON’T KNOW” answer.

A total of 101 comments were given by respondents for Example 7, of which 85 were to justify the answer why they thought it as spam, 7 for non-spam and 9 comments for “DON’T KNOW.”

For Example 8, there were a total of 105 comments left by the respondents. However, 101 of this number were left to justify why they thought it is spam and only 4 comments were to justify “DON’T KNOW.”

For Example 9, there were only 71 comments left by respondents making it the least number of comments received from respondents for justification. Of the 71 comments, 27 were to justify why respondents have answered spam, 23 for the non-spam answer and 21 for the “DON’T KNOW” answer.

There were a total of 74 comments for Example 10 with 24 comments given to justify the answer for spam. Another 30 comments was to justify the non-spam answer and 20 comments for answering “DON’T KNOW.”

Based on Figure 6.3, it can be seen that nearly 60% of the comments were to justify answers for spam except for Examples 4, 9 and 10. There were less than 20% of the comments given to justify why respondents have answered “DON’T KNOW” except for Examples 3 and 9. Nonetheless, out of all these 10 examples, Examples 3, 4, 5, 6 and 10 are real non-spam examples. The thesis further analyses these comments and identifies spam characteristics from these comments in the next subsection.

6.3.4.2 Spam characteristics

This subsection presents the characteristics used by respondents to describe what spam is. These comments will not add up to the total of all comments received, as comments that contain several characteristics are treated separately. These characteristics are based on the comments given by respondents. The comments were analysed and divided into nine categories:

- Spam keyword
- Repetitive
- Unsolicited
- Suspicious
- Malicious and threats
- Unrelated/irrelevant
- Mistakes in written language
- Advertisement/promotion
- Access and authorization

Any comment that mentioned spam keywords or spam-like words or symbols is classified into this category. There were a total of 10 comments with 2 that come from Example 1, 2 from Example 8 and 6 from Example 2. Some of the comments included are “*The word OMG shows a typical online spam ;)*” and “*I can’t believe this works... is a typical spam tagline.*”

On the other hand, for any comments that have mentioned “repetitive” or any synonyms, they are classified into this category. Comments adding up to 37, 13, 5 and 1 were found with related words from Examples 1, 3, 5 and 8, respectively, with a total of 55 comments classified under this category. This includes “*Repeated words,*” “*Repetitive sentences*” and “*There are repeating comments on the link.*”

Any comment that contains unwanted, unnecessary or uninvited words is classified into this unsolicited category. A total of 35 comments were classified under this category. This includes “*There is unwanted link for application in this page,*” “*Unwanted response that u will receive when u click the button*” and “*This is a notice I don’t want to get.*”

For the suspicious category, any comments that contain the word “suspicious” or any sentences that will cause suspicion are classified into this category. Words that are related to causing suspicion are

fishy, weird, strange, vague, awkward, inappropriate, unknown and dodgy. Irregular activities could also raise suspicion. Apart from things that can be identified from available situations, users could also be suspicious if they find any unavailable, incomplete or inaccessible items. Hence, in total, there are 304 comments that were classified into this category with 28 comments from Example 1, 48 comments from Example 2, 25 comments from Example 3, 9 comments from Example 4 and 34 comments from Example 5. Fifty-five comments from Example 6 were categorized under this category followed by Example 7 with 41, Example 8 with 30, Example 9 with 21 and 13 comments for Example 10. The thesis now provides some examples of the comments for each of the words listed above:

- *“Contain suspicious link.”*
- *“I don't know Raymond Edwards, why do i need to click on the links and look at his pictures?? Very fishy....”*
- *“Didnt realised a white guy promoting a south east asia country! its weird.”*
- *“Strange (vague) text message and how it direct the reader to do action (to click) without certain description about what will happen later.”*
- *“Awkward Link.”*
- *“There is inappropriate link appear on his profile.”*
- *“Unknown links appears.”*
- *“Dodgy looking like button compared to normal above.”*
- *“It tries to promote a app engine – which is not a regular activity of the user.”*
- *“As the main attachment is not available, it might contain spam.”*

Any comments that contain “malicious” or other threats will be classified into this “Malicious and threats” category. A total of 67 comments contain related words that fall into this category. For instance, *“It's a malicious attempt to access our contact list to send malicious programs,”* *“The provided links might harm the pc/laptop or contains virus link”* and *“Have to click before the next step is revealed. misleading”* are some of the comments classified into this category.

There are 25 comments that were classified under unrelated or irrelevant categories for having these words in their comments. Five comments each were classified from Examples 2 and 3, four comments each were classified from Examples 1 and 8, three comments from Example 6 and one from Examples

4, 5, 7 and 10. Some of the comments included in this category are “*Unrelated link pops up on an unrelated posts,*” “*Totally irrelevant content*” and “*The link is not related to the info.*”

Any comments that were found to be related with grammar, spelling and language mistakes are classified into this category. Only seven comments were found related to this category such as “*Wrong spelling on the name,*” “*Littered with grammatical errors. If it's written by a Caucasian, I'd expect it to be error free :)*” and “*He mixed two languages.*” Four comments were categorized from Example 1, and one each from Examples 3, 4 and 8.

A total of 59 comments from respondents have mentioned advertisement or promotion as a way of identifying spam. If the respondents mentioned about increasing page rank or number of friends, those comments were classified into this category as it is also an act of advertising and promoting their page or profile. Ten comments were identified from Example 1; 14 from Question 2, 15 from Example 3; 7 from each Examples 4 and 8; and another 2 each from Examples 5, 6 and 7. Examples of the comments given are “*It tries to promote a app engine – which is not a regular activity of the user,*” “*Repetition advertisement*” and “*Why would someone want you to see his pictures by following links?? he can just put the pics in the online album, unless he wants some traffic to his site.*”

For the “Access and authorization” category, comments relating to issues of authorization such as autopost and subscription will be identified. Autopost is considered as an unauthorized act while subscription is usually done to get authorization and access to posting activities. Mass distribution is also included under the “Access and authorization” category as it means that many people will be able to have access to read the spam unit. A total of 24 comments were classified under this category, such as “*The link will broadcast to all your contact in FB once it has been clicked,*” “*The auto-posted link from an application and the repetitive comments on a link*” and “*If the user is not subscribing the Ads.*”

From this result, it can be seen that any spam unit that is found to be suspicious contains malicious and other possible threats, with the intention to advertise or promote unsolicited, irrelevant content, impacting access and authorization issues and which is repetitive can be considered as spam by public users. In the next subsection, the thesis analyses the respondent’s justification of deciding the examples to be non-spam.

6.3.4.3 Non-spam characteristics

In this subsection, the thesis presents the characteristics of what non-spam is based on the respondents’ comments. As in the previous subsection, comments that contain several characteristics are treated separately.

From the data collected, it is found that respondents will think that the examples are non-spam if they trust the sender, social networking sites and source of application. Some of them also think the examples are non-spam because they are just ordinary advertisements. Examples are also classified as non-spam depending on the available content, whether it is acceptable and looks genuine. However, there are also respondents who justified their chosen answer as non-spam because they are uncertain and have never had any similar experience.

6.3.4.4 Analysis based on spam examples

For Example 1, most respondents have correctly identified the example as spam. The two most cited reasons of the respondents' justification as to why they classified this example as spam is the repetitiveness of spam units in the examples (37 comments) and that they are suspicious of the available content (28 comments).

Example 2 also showed a high percentage of respondents who managed to correctly identify the example as spam. Based on the comments, the respondents justified this example as spam because they are suspicious (48 comments) and they think that the example contains malicious threats (27 comments).

For Example 3, most of the participants have incorrectly identified this example as spam. There are participants who correctly identified it as non-spam for the reason that it is not spam if the user is not subscribing to the ads. However, links provided on the screen cause respondents to be suspicious (25 comments). Some examples of the comments given by respondents are "*Unnecessary link,*" "*Link provided looks suspicious,*" "*Mail doesn't look like proper promotion email from Airlines. Also, the domain in below link doesn't belong to Firefly Airlines*" and "*The link given seems to be suspicious.*" They also classified this example as spam because it is an advertisement. However, there are several respondents who managed to correctly identify this example as non-spam and their comments are as follows:

- "*I think it just ordinary internet marketing*"
- "*This is merely an advertisement.*"
- "*It is a straight forward and guided information.*"
- "*It was from a secured website*".

For Example 4, most participants incorrectly identified the example as spam. With only 31 of the respondents leaving their comments for justification, 9 of the comments were related to being suspicious, followed by 7 of the comments of the view that it is spam because it is as an advertisement. Some of the comments are "*Dubious information,*" "*Suspicious account holder and*

suspicious event promoted” and *“Grammatical error with poor written English.”* Nonetheless, there are participants who correctly stated the reason that if the invitation comes from a friend, then it is not considered as spam. Some of the justifications given by the respondents who stated that this is non-spam are as follows:

- *“It contains valid information.”*
- *“Appears to be a genuine message...”*
- *“It is a clear invitation.”*
- *“If he is my friend, then no, its not a spam.”*
- *“There's a specific group indicated.”*

For Example 5, most participants who correctly identified it as non-spam have given the reason that if they know the sender, then it will not be considered as spam. However, similar to the previous questions, links embedded in the message cause respondents to feel suspicious. Thirty-five of the respondents’ justification is that the content is suspicious. A few justifications given by the respondents are *“Repeated link,” “Too many links, could have just attached photos here instead of the links,” “Suspicious links,” “Unwanted links lead to unwanted spams”* and *“There is too much link.”* Some of the respondents’ justifications for non-spam are also based on trust in the web application. Two of the comments provided by the respondents show that they are knowledgeable which is about the authorization key contained in the link. Their comments are *“There is an authentication key in order to open the picture”* and *“There is authorization key in each link.”*

For Example 6, most respondents incorrectly identified the example as spam. However, this is because they think that the recipient received the message from an unknown sender. From the comments left, 55 of the respondents think that it is suspicious for the sender to request the email address through YM. Some of the comments left are *“Suspicious message,” “Its kinda weird message”* and *“Asking for email add.”* Some of the comments stated that the email addresses are available for the sender once they add the contacts in the lists. However, in real life, other users can be added to YM lists using their id without knowing the email addresses. Hence, the activity of requesting the email address could be considered as an unsuspecting activity.

Out of all the 10 questions, Example 7 showed the second highest percentage where the respondents correctly identified the example as spam. Forty-two comments stated that it is suspicious. Most of the comments categorized under the suspicious category were about the link. Some of these comments are as follows:

- *“The offered link is suspicious.”*

- *“The link is weird.”*
- *“All the colours seem strange.”*
- *“74uc2u? Suspicious link.”*
- *“Weird request, sounds like free porn.”*

Example 8 was recorded as the highest percentage where the respondents managed to answer correctly. Most of the respondents managed to identify the example as spam. Thirty of the comments were categorized under suspicious, eight comments were categorized under unsolicited, another eight comments were categorized under “Malicious and threats” and seven comments were categorized under advertisement. Some of the comments are listed below:

- *“Suspicious words and link.”*
- *“The link is so suspicious, seems to be from a non-legitimate source.”*
- *“Dodgy link and message.”*
- *“Really spam, no news, no regards, link is unacceptable...”*
- *“Porn ads are potentially malicious.”*
- *“Video of porn are more likely have a hackers activities.”*
- *“Using catchy word such as porn is a kind of attracting somebody to come to their site.”*

Nonetheless, this question provides many indications to be categorized as spam. However, most respondents just mentioned the suspicious link.

Example 9 contains a spam example and most of the respondents incorrectly identified this example as non-spam. Some of the justifications for stating that it was non-spam are as follows:

- *“Blank message.”*
- *“Can’t see any message.”*
- *“This shouldn’t be any spam.. coz there is no link there.. but it still depends.. should be careful to know whether a person adding u is trustworthy friend or unknown.”*

Nonetheless, the justifications provided are unclear. It is obvious that it comes from a user who was not listed in personal contacts and this reason was also mentioned in some of the respondents’

comments. Furthermore, it is suspicious that the sender was able to send a blank message. Such comments are as follows:

- *“No message typed and the user is not your friend, totally spam.”*
- *“As long they're not in your friend's list, this can be consider as a spam.”*
- *“Who is copo_454? unsolicited instant message.”*
- *“U cant send any blank messages.”*

For Example 10, half of the respondents (52.8%) incorrectly identified the example as spam. Based on the comments left by the respondents, they thought that it contains a suspicious link and the message comes from an unknown sender. On the other hand, most of the respondents who stated that it was non-spam stated that the link is not suspicious. Some of the comments are as follows:

- *“Typical IM messages with seems to be valid link.”*
- *“Link to specific event.”*
- *“Ok, the link provided with clear purpose of what it is.”*
- *“Check the link and working.”*

6.4 Discussion

Generally, there was not much difference between perceived awareness and actual awareness. Nonetheless, perceived awareness and actual awareness were only compared between those who think that they have zero knowledge about Spam 2.0 and those who think that they have knowledge of Spam 2.0. Through observation of the level of awareness that was calculated from actual awareness, most of the respondents are categorized as having a high level of awareness (more than half of the respondents were categorized into high and very high levels of awareness) and it is evident that most of the respondents are aware, can identify and have experiences with basic spamming techniques used by spammers. Hence, it is concluded that the public awareness level of Spam 2.0 is quite high.

Nonetheless, there are several spamming activities that seemed to be unfamiliar to the respondents such as *“Found ages that are only full of repeated keywords,”* *“Being tagged by unwanted parties”* and *“Received/seen suspicious link on a Web 2.0 application.”* Undeniably, the respondents could have never encountered them but it is also possible that they could not have recognized these actions which explains the reason why the percentages of respondents who have answered “DON'T KNOW” for these questions are relatively high (13.3%, 14.4% and 19.3%, respectively), compared to the other activities.

On the topic of knowledge of Spam 2.0, there were only a small percentage of respondents who stated they have no knowledge at all or they are experts on Spam 2.0. Most of the respondents think that they have fair knowledge. Similarly, through a detailed observation of the level of knowledge which was calculated from actual knowledge, the respondents' level of knowledge on Spam 2.0 is quite high with more than 80% of the respondents having at least fair knowledge.

When compared to the level of knowledge that was calculated from actual knowledge, more respondents thought they have less knowledge. It is evident from looking at the percentage of perceived knowledge, which is smaller than the percentage of the level of knowledge for the categories of none, poor and fair. Still, there was also a difference of 1.94% of respondents who thought they were expert, but when considering their actual knowledge, they were not categorized under expert.

The levels of knowledge were calculated based on three different questions which are Actual Knowledge I, Actual Knowledge II and Actual Knowledge III. Questions on Actual Knowledge I revealed some of the non-spam activities such as "*Account being hacked/hijacked,*" "*Attacked by virus*" and "*Being asked username, passwords and/or credit card details by a trustworthy-looking entity in the Internet communication*" are falsely viewed as online spam and do not have the knowledge on these actions. This is evident from the percentages of the respondents who have both correctly and incorrectly identified these statements. From the percentages of the respondents who have chosen "DON'T KNOW", it is evident that the respondents do not have the knowledge on Spam 2.0 actions such as "*Found pages that are only full of repeated keywords.*"

Aside from the percentages of wrong answers, the percentages of respondents choosing "DON'T KNOW" as the answer are also important. As expected, the questions about the technical side in Actual Knowledge II and the ones considered difficult have the highest percentages of respondents choosing "DON'T KNOW." These questions are "*Online spam has lower viewer impact than email spam*" and "*Auto registration software can be used to register spam accounts*" (44.6% and 49.5%, respectively).

Considering those who answered the questions incorrectly and those who chose "DON'T KNOW" as not knowledgeable, it can be seen that there are some facts which are not well understood by the general public. Of the respondents, 75% had a misconception that all Spam 2.0 can be detected and treated using the existing anti-spam techniques that are designed for email spam; 63.6% of the respondents do not know that there is a difference between Spam 2.0 and email spam. Other facts about Spam 2.0 were known by at least more than 50% of respondents.

Looking at Actual Knowledge III, most respondents managed to answer Questions 7 and 8 correctly as they were probably the easiest questions compared to other questions as the screens fulfil a lot of

spam indicators. For other questions, approximately three-quarters of the respondents have a misconception and do not have the knowledge on identifying spam presented in Examples 3, 5, 6 and 9 with the percentages 76.3%, 72.2%, 72.3% and 73.6% recorded respectively. The participants tend to decide that those examples are spam even though it is not. Spammers will try to pose as real users and the messages they post would seem genuine, but it is also interesting to see how genuine messages from friends have been categorized as spam.

Comparing the level of awareness with level of knowledge, it is obvious that more than half of the respondents were categorized as having high and very high levels of awareness, but only 19.1% of the respondents have good and expert knowledge. It is possible that even for those who have high awareness; they obtained only basic knowledge about Spam 2.0. Thus, when considering difficult questions or higher level complicated questions, even most of the high-categorized-awareness respondents could not answer these questions correctly.

The correlation between awareness and knowledge was also tested to see if there is any association between them. The correlation is done after the two types of data are copied into SPSS. Bivariate correlation using the non-parametric Spearman correlation was chosen and they were correlated with $r = 0.452$. The linear association between this awareness and knowledge is considered as medium correlation. Hence, it provides an indication that there could be a positive relationship between those who are aware and those who are knowledgeable about Spam 2.0.

Most respondents agreed that Spam 2.0 is a problem and spamming is unacceptable. Nonetheless, one-fifth of the respondents for both questions chose to answer "DON'T KNOW" indicating that there are a small group of respondents who do not think that spam is a problem. Through the questions to determine the punishment for spammers, more than half of the respondents think that confessed spammers should be punished. With 24.5% of the respondents thinking that convicted spammers should be allowed to work in the computing field and with nearly two-fifths of the respondents choosing "MAYBE" for each of the crime's punishment showed that respondents' attitudes towards punishing spammers severely disagree with each other. Furthermore, this is supported by our result for perception towards the seriousness of crime. This question provides an indication that both email and online spam were not seen to be as serious as other computer crimes where both of these crimes fall in the last place compared to other computer crimes.

More than half of the respondents (53.5%) think that they are vulnerable to spam. However, there are 25.3% who think that they are not vulnerable to spam. The next question shows a similar pattern; with 33.1% of the respondents who think that there is a low likelihood of them being spammed and 33.9% who think that there is a high likelihood of them being spammed. The fear of crime could be low if the respondents think that the applications provide sufficient protection against spam or they might also think they have enough skill and knowledge to handle spam. While the respondents believed that

the motivation behind spamming is to make money, it is also possible that they could not see themselves as victims.

Based on the comments analysis on justifications provided by the respondents, some of the respondents will consider a message as spam if the sender is unknown. However, the comments provided by the respondents relatively correspond to links; thus, they disregard other potential sources to identify if it as spam or not, such as the account owner and username. A few respondents are highly analytical and provide good justifications showing that they have good knowledge on this topic. In addition, it is also identified that the respondents have a tendency to classify the examples into spam and give their justification for spam answers more than non-spam answers.

From the users' perspective, the most common reason why the participants will categorize an example as spam is if they see it to be suspicious. Hence, the lesson learnt from this survey is in order to create a better advertisement or to avoid messages being seen as spam, the sender has to avoid creating suspicious content.

6.5 Conclusion

This chapter presents a detailed explanation of the public awareness, knowledge and perception of Spam 2.0 survey results. In conclusion, even though most respondents are aware of online spam, it is also apparent that respondents have inadequate knowledge about online spam. It is also apparent that the public have basic knowledge on Spam 2.0; however, most of them failed at recognizing harder Spam 2.0. Thus, this study shows that there are bigger opportunities for spammers to trick users into becoming their victims if spammers use complicated and smart tricks. Furthermore, it was also evident that Spam 2.0 is not seen to be as serious as other computer crimes.

The following chapter provides the thesis conclusion and future work that could be done to further improve the research.

Chapter 7

Conclusion and Future Work

This chapter covers:

- ▶ The current issues and problems with the cost model for Spam 2.0;
- ▶ The current issues and problems with the public awareness, knowledge and perception of Spam 2.0;
- ▶ Solutions proposed by this dissertation to address the Spam 2.0 cost model;
- ▶ Solutions proposed by this dissertation to address the public awareness, knowledge and perception of Spam 2.0; and
- ▶ Conclusion and future works.

7.1 Introduction

The advancement in using the Internet as the medium of communication has allowed Internet users to basically add value to the web 2.0 applications liberally. The opportunities that facilitate and promote information sharing are based on the read/write concept where users can maximize the ability to interact, collaborate and contribute content on the World Wide Web. This user-centred design notion in the web 2.0 platform that allows users to generate and consume information has provided an open door for spammers to carry out their spamming activities by posting inappropriate, unsolicited and irrelevant content. The contents that are embedded in web 2.0 applications created in order to trick users into clicking other pages and promoting fake products are called *spam 2.0*. Spam 2.0 not only degrades the quality of information and user's trust in particular web 2.0 applications but, worse than that, could also bring about severe consequences relating to computer security.

As explained in Chapters 1 and 2, this new type of spam definitely has its cost. Nonetheless, most of the existing research focuses only on developing better spam filtering techniques. While in the email spam domain, there exists a lot of spam calculators or work relating to identifying the cost of email spam, to the best of the author's knowledge, there was none relating to Spam 2.0 costs.

The existing literature has been extensively reviewed and the major problems involved with Spam 2.0 costs have been highlighted. As there was no existing work on Spam 2.0 costs, the backbone of the

studies depends on the existing works in the email spam domain. Detailed research on other cost models was carried out to develop a complete cost model for Spam 2.0.

Throughout the research, it was observed that one of the reasons why the Spam 2.0 campaign is easily proliferated is due to the users' lack of awareness, knowledge and unknown perception of Spam 2.0. Users may not be able to identify Spam 2.0 and thus end up falling for the campaign. They could also be promoting those campaigns without realizing it. Besides, the users' decision on handling Spam 2.0 could have a huge impact on spammers' revenue. As explained in Chapters 1 and 2, to the best of the author's knowledge, there was no research that focused on awareness, knowledge and perception of Spam 2.0.

The major drive to accomplish this research is twofold: (1) to develop a Spam 2.0 cost model and further identify related costs and (2) to address the public awareness, knowledge and perception of Spam 2.0. This chapter mainly focuses on providing the ideas behind carrying out this research and further contributions made out of this research. Future works that could be extended from this research are also presented in the later subsections.

7.2 Problems and Issues

This thesis identifies several problems and issues regarding cost, awareness, knowledge and perception relating to Spam 2.0. These problems are explained in Chapter 3 and are summarized according to:

- The Spam 2.0 cost model
- Public awareness, knowledge and perception of Spam 2.0

7.2.1 Issues with Spam 2.0 Cost Model

As mentioned earlier, the Spam 2.0 cost model is needed to quantify the cost involved due to spamming activities. However, there were several issues identified as delineated in Chapter 3, which are listed as follows:

- Unavailability of information on cost categories and cost parameters related to the Spam 2.0 cost model.
- Inexistence of Spam 2.0 data/unavailability of data to be used on developing the Spam 2.0 cost model.

- Inability to measure certain cost categories for Spam 2.0 without having an internal system (such as spam filtering facilities, etc.) and survey.
- Inability to measure actual time wasted on Spam 2.0 automatically.
- No specific evaluation found on existing cost models.
- Inexistence of external data on Spam 2.0 to be used as cost model input.

It is clear in the area of Spam 2.0 there were several works on the spam filtering techniques. However, there is a lack of focus on the costs involved with Spam 2.0. For this reason, the issues presented above have to be solved beforehand.

7.2.2 Issues with Public Awareness, Knowledge and Perception on Spam 2.0 Survey

As mentioned in the earlier section, the individual's awareness, knowledge and perception of Spam 2.0 will influence how they identify and handle Spam 2.0. For example, users' awareness of Spam 2.0 consequences, users' knowledge of Spam 2.0 and users' view on Spam 2.0 as a serious problem are vital in preventing Spam 2.0 propagation. Nevertheless, as revealed in Chapter 1 and Chapter 2, awareness, knowledge and perception of the Spam 2.0 problem are unknown.

Thus, as delineated in Chapter 3, the problems in doing a survey on public awareness, knowledge and perception of Spam 2.0 are as follows:

- Inexistence of exact similar studies on Spam 2.0 to be adapted into the research.
- Inexistence of validated Spam 2.0-related questions on awareness, knowledge and perception.

It is apparent that these issues have not been adequately addressed by current studies.

7.3 Dissertation Contributions

In order to analyse the problems identified in Chapter 3, the dissertation proposes to use HoneySpam 2.0 and a survey to obtain data. As a summary and to highlight its contribution, the thesis lists again the eight research questions defined in Chapter 3 and shows that the thesis has managed to answer all these research questions.

- RQ1: Can the research develop an internal system that can define all relevant cost categories and cost parameters for the Spam 2.0 cost model?

Yes, the research has managed to develop an internal system that can define all relevant cost categories and cost parameters for the Spam 2.0 cost model through the method explained in Chapter 4.

- RQ2: Can the research develop an internal system to produce enough Spam 2.0 data to be used for estimating the cost model?

Yes, the research has managed to develop an internal system to produce enough Spam 2.0 data to be used for estimating the cost model using HoneySpam 2.0 as explained in Chapter 4.

- RQ3: Can the research measure the actual time wasted on Spam 2.0 without users having to estimate the value themselves?

Yes, the research has managed to measure the actual time wasted on Spam 2.0 without users having to estimate the value themselves using the timing function as explained in Chapter 4.

- RQ4: How does the research evaluate the cost model?

Evaluation of the cost model is solely done based on available data that comes from HoneySpam 2.0 and survey. Further evaluation could be done if there exists any other data sets in the future.

- RQ6: To what extent are the public users aware of Spam 2.0?

Most of the public users are aware of Spam 2.0. More than 80% of public users reached at least the intermediate level of awareness.

- RQ7: To what extent is the knowledge of public users on Spam 2.0?

Approximately 80% of public users have a fair knowledge of Spam 2.0 but only a very small percentage were considered expert.

- RQ8: What is the perception of public users towards Spam 2.0?

Most respondents think that Spam 2.0 is a problem and it is not acceptable. Half of the respondents agreed that confessed spammers should be punished. Only a smaller percentage of respondents agreed that convicted spammers should be allowed to work in the computing field. More than half of the respondents think that they are vulnerable to spam. In terms of seriousness of crime, Spam 2.0 was ranked sixth out of seven other computer crimes. Most public users think that the spammers' motivation for spamming is to make money.

Throughout the research, several contributions made from this research are explained according to two categories: (1) cost and (2) awareness, knowledge and perception.

7.3.1 Cost

Although this research continues the concept of Spam 2.0 from several existing researches, most of the researches only focus on the method of spam detection and prevention. While this research aims to identify its cost, there is a lack of research that specifically focuses on the cost of Spam 2.0. Therefore, a detailed review on the current literature involving the topic of the cost of email spam cost models needs to be carried out. Other related cost models are also explored in order to gain additional information that may add value to the research. Cost categories and cost parameters defined in each cost model are then evaluated. The data collection method and stakeholders focused on for each cost model are also identified. The current findings from these extensive literature reviews add to a growing body of literature on the cost model studies.

Through the findings gathered in the cost model literature review, this dissertation proposes a Spam 2.0 cost model. The total cost of Spam 2.0 is specified as a sum of storage cost, loss of productivity cost, labour cost, connectivity cost and software cost. However, this thesis focuses only on two of these cost categories which are storage cost and loss of productivity cost. Nonetheless, these cost categories and related cost parameters have been identified in the proposed cost model. The cost models proposed here may be applied to estimate the costs for other spam units in other web 2.0 applications. Overall, the proposed cost model paved the way for the development of a better quantification of the cost involved due to spamming activities.

Storage cost is estimated based on the HoneySpam 2.0 data sets and a storage cost survey. Cost parameters involved in this cost calculation are size of storage used to store Spam 2.0 and current storage cost price. As defined earlier, any other related costs such as bandwidth, management and labour cost can also be included in this cost calculation. One of the contributions of this dissertation is to provide the storage cost of a spam unit itself, as there have been no prior works to quantify this cost of Spam 2.0. In addition, it proves that Spam 2.0 definitely has its cost. This result will be able to serve as a base for future studies relating to the cost of Spam 2.0. Nonetheless, it is expected that this cost will decrease in the future based on the price reduction of cost of storage packages offered by commercial companies. In addition, the storage cost is also greatly influenced by the number of spam and size of spam that managed to successfully be embedded in the content. Due to extensive research on spam filtering, this situation seems to be sided with the non-spammers. However, it may not work out that way as spammers will always find a smarter way to make their campaign work.

Similar to the storage cost, the loss of productivity cost has never been quantified previously. Cost parameters involved in this cost calculation are average time wasted on each spam unit and average spam units received in a day. Therefore, the loss of productivity cost is calculated mostly based on timing data sets obtained from a web-based survey. Technically, this will resolve the problems of users' estimation of the value of time taken to identify spam. A new type of cost category called cost of misjudgement was also introduced. This type of cost is beneficial in identifying the indirect cost due to spamming activities. Nonetheless, it was proposed to utilize the timing function available in the web-based survey system to estimate the actual time wasted in identifying Spam 2.0. The resulting value for this cost will be able to serve as a predefined value for future works on Spam 2.0 costs.

7.3.2 Awareness, Knowledge and Perception

As the focus of this dissertation is twofold which also includes the studies on awareness, knowledge and perception of Spam 2.0, a comprehensive study has been conducted on existing research relating to these topics. While the topic of awareness, knowledge and perception is widely used in public health or other social science research, the overall problems relating to the subject matter were seldom discussed in any computer security field. Thus, there was a lack of literature on this particular topic. To evaluate the awareness level, knowledge level and perception of public users towards Spam 2.0, 29 survey items have been developed that cover:

- Perceived awareness
- Actual awareness
- Perceived knowledge
- Actual knowledge
- Perception towards crime
- Perception towards crime's punishment
- Fear of crime
- Perception towards seriousness of crime
- Perception towards crime's motivation

Given that there is limited prior work in the literature on the awareness, knowledge and perception related to computer security concerns, the results of the survey presented here are the initial steps towards identifying questions that are suitable to be used in this survey, providing Spam 2.0-related

items related to awareness, knowledge and perception and developing a self-reported awareness and knowledge scale of Spam 2.0. Through modification of any prior works that are used where possible, a literature review on this work contributes to existing knowledge of both fields by providing other insights into these issues.

The current literature also lacks a comprehensive heuristics in classifying Spam 2.0. Thus, three heuristics used for spam characteristics were defined:

- Authorization issue: Does it come from a known source?
- Validation issue: Does it truly come from that known source?
- Trust issue: Does it raise suspicion or consist of suspicious content?

The heuristics identified here assist in the development of examples on Spam 2.0 identification. Hence, one of the main contributions of this study was a set of heuristics of determining Spam 2.0 to highlight the unresolved issues and, at the same time, properly understand that the researcher's view might differ from the users' view. The categories of the definitions on detecting Spam 2.0 are based on these nine categories which are as follows:

- Spam keyword
- Repetitive
- Unsolicited
- Suspicious
- Malicious and threats
- Unrelated/irrelevant
- Mistakes in written language
- Advertisement/promotion
- Access and authorization

Related terms used for each category based on their knowledge provided an insight into users' understanding on Spam 2.0. Based on users' perspective, their spam identification is highly dependent on the suspicious nature of an example provided to them. Thus, the result makes several noteworthy contributions to creating better online content to avoid being categorized as spam by public users.

Early indications of the extent of the public's awareness, knowledge and perception of Spam 2.0 problems are reported at the end of the study. The study has gone some way towards enhancing our understanding on the public awareness, knowledge and perception of Spam 2.0. These findings suggest several courses of actions to be taken such as:

- Increase users' awareness by educating them on recognizing basic spamming techniques used by spammers.
- Increase users' knowledge by reducing misconceptions on the actions considered as Spam 2.0 activities.
- Increase users' view on the seriousness of Spam 2.0 problems by conveying the danger of Spam 2.0 and how it relates to other major security concerns.

Whilst this study did not confirm the after-effects the survey had on each respondent, it did partially substantiate their awareness and knowledge on Spam 2.0 problems.

7.4 Limitations and Future Works

The research works carried out in this thesis have been published in peer-reviewed international conferences and journals. Over the course of this research, seven papers have been published. Selections of these publications are attached in Appendix II. A substantial amount of work has been devoted to complete this research. Nonetheless, there were still a number of caveats associated with the study that need to be acknowledged which might affect the results and the original objectives of the thesis. Nonetheless, in any research work, limitations are unavoidable due to the existing constraints. Still, these limitations provide an opportunity for future works as follows:

- A major limitation of this doctoral thesis was the inability to completely run the experiments to produce other related costs in Spam 2.0 cost models. Although all related cost categories and cost parameters have been properly identified, unfortunately, due to time constraints, these costs cannot be examined. Future research could also cover quantifying the indirect cost such as hatred, annoyance and trust while dealing with Spam 2.0.
- Studies on the Spam 2.0 cost model are indeed a new field of research. Therefore, any other sources of external data for comparison could not be found. Compared to email spam where email spam data sets are publicly available, Spam 2.0 data sets are not found publicly. More information on other Spam 2.0 data sets would help us establish a greater degree of accuracy on this matter. In addition, a number of possible same experimental set-ups focusing on other

web 2.0 applications could further add to the public Spam 2.0 repository to be used by other researchers.

- This research has successfully avoided using the estimation values made by users in detecting the time taken for them to identify Spam 2.0. However, the automatic measurements generated by the systems are very sensitive. Although users are made aware of this situation, they might have ignored this. Timing data sets obtained from the survey could also be influenced by external factors such as users' behaviour and attitudes while answering the survey and users' familiarity in using the web survey. Thus, it is recommended that further research be undertaken under a smaller controlled environment such as a closed lab experiment.
- In the section of 'Spam 2.0 Identification', pre-planned Spam 2.0 examples are presented to survey participants. In the future, Spam 2.0 examples could be presented in a more natural way such as pop-ups. From this research, user behaviour when dealing with spam could be observed. The factors influencing the time taken to identify Spam 2.0 could be usefully explored as intriguing issues in further research.
- Although the thesis claimed that awareness, knowledge and perception might indirectly influence the rate of spam proliferation and spammers' revenue, unfortunately the thesis did not explore these relationships. Thus, it is suggested that the association of these issues is investigated in future studies.

Bibliography

- Abrams, Lawrence. 2010. Remove the Fake Microsoft Security Essentials Alert Trojan and Antispysafeguard Series Title: Bleepingcomputer.Com. <http://www.bleepingcomputer.com/virus-removal/remove-fake-microsoft-security-essentials-alert>.
- Access Communications Pty Ltd. 2013. Adsl Overview. Accessed 26 June, http://www.accesscomms.com.au/Reference/ADSL_overview.htm.
- Acquisti, Alessandro, and Ralph Gross. 2006. "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook." In *Privacy Enhancing Technologies*, eds George Danezis and Philippe Golle, 36-58. Springer Berlin Heidelberg.
- Agho, A. O. 2001. "Correlates of Actual and Perceived Knowledge of Prostate Cancer among African Americans." *Cancer nursing* 24 (3): 165.
- Akismet. 2013. Spam Zeitgeist. Akismet. Accessed 01082013, <http://akismet.com/about/>.
- Al-Alawi, Adel Ismail, and Mohamed Fathi Abdelgadir. 2006. "An Empirical Study of Attitudes and Opinions of Computer Crimes: A Comparative Study between U.K. And the Kingdom of Bahrain." *Computer Science* 2 (3): 223-235.
- Alt-N Technologies. Spam Cost Calculator. Accessed 24 December, <http://www.altn.com/Products/SecurityPlus-Antivirus-MDaemon/SpamCostCalculator/>.
- Australian Bureau of Statistics. 2013. State and Territory Statistical Indicators, 2012 22 May 2012 Accessed 20 June, <http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by+Subject/1367.0~2012~Main+Features~Average+Weekly+Earnings~6.8#>.
- Bishop, M. 2006. "Teachers' Knowledge About Epilepsy and Attitudes toward Students with Epilepsy: Results of a National Survey." *Epilepsy & behavior* 8 (2): 397.
- Brandt, Andrew. 2010. Facebook Spam Leads to Viagra Vendor, Drive-by Download. 28 May, 2010 Accessed 3 October,
- Brucks, Merrie. 1985. "The Effects of Product Class Knowledge on Information Search Behavior." *Journal of Consumer Research* 12 (June): 1-16.
- Burstein, Frada, and Shirley Gregor. 1999. "The Systems Development or Engineering Approach to Research in Information Systems: An Action Research Perspective." In *Proceedings of the 10th Australasian Conference on Information Systems, Victoria University of Wellington, New Zealand*, edited by B. Hope and P. Yoong, 122-134.
- Chu, Zi, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. 2010. "Who Is Tweeting on Twitter: Human, Bot, or Cyborg?" In *Annual Computer Security Applications Conference (ACSAC'10), Austin, Texas, USA*, 21-30. ACM. doi: 10.1145/1920261.1920265.

- Cluley, Graham. 2012. Turn Facebook Pink, Red or Black? Don't Fall for Online Scams. 11 January 2012 Accessed 29 November, <http://nakedsecurity.sophos.com/2012/01/11/turn-facebook-pink-red-or-black-dont-fall-for-online-scams/>.
- Cobb, Stephen. 2003. The Economics of Spam. http://www.spamhelp.org/articles/economics_of_spam.pdf.
- Commtouch. 2010. *Internet Threats Trend Report Q1 2010*.
- Computer Mail Services. The Cost of Spam. Accessed 24 December, <http://www.cmsconnect.com/marketing/spamcalc.htm>.
- Crosby, R. A. 2001. "Perceived Versus Actual Knowledge About Correct Condom Use among US Adolescents: Results from a National Study." *Journal of adolescent health* 28 (5): 415.
- Davidsson, Per. 2004. *Researching Entrepreneurship*. Edited by Springer. 1st Edition ed. Boston.
- Dillman, Don A. 2007. *Mail and Internet Surveys : The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide*. Edited by 2: John Wiley & Sons, Inc.
- Dowland, P. S, S.M Furnell, H.M Illingworth, and P.L Reynolds. 1999. "Computer Crime and Abuse: A Survey of Public Attitudes and Awareness." *Computers & security* 18 (8): 715-726.
- Ellis, R. Darin, and Jason C. Allaire. 1999. "Modeling Computer Interest in Older Adults: The Role of Age, Education, Computer Knowledge, and Computer Anxiety." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 41 (3): 345-355.
- Fan, Weimiao, and Zheng Yan. 2010. "Factors Affecting Response Rates of the Web Survey: A Systematic Review." *Computers in Human Behavior* 26 (2): 132-139. doi: <http://dx.doi.org/10.1016/j.chb.2009.10.015>.
- Ferris Research. 2005. *The Global Economic Impact of Spam*.
- Ferris Research. 2009. Email Industry Statistics. <http://email-museum.com/reports/industry-statistics/>.
- Flynn, Leisa Reinecke, and Ronald E. Goldsmith. 1999. "A Short, Reliable Measure of Subjective Knowledge." *Journal of Business Research* 46 (1): 57-66. doi: [http://dx.doi.org/10.1016/S0148-2963\(98\)00057-5](http://dx.doi.org/10.1016/S0148-2963(98)00057-5).
- Freeman, R. Edward. 1984. *Strategic Management: A Stakeholder Approach*. Boston: Pitman.
- Galliers, Robert D. 1992. *Choosing Information Systems Research Approaches*. Edited by Robert D Galliers, *Information Systems Research: Issues, Methods and Practical Guidelines*: Oxford: Blackwell Scientific.
- Gartner Group. 1999. Impact of Unsolicited Commercial E-Mail. Accessed 4 April, <http://www.triumviratetechnologies.com/impact.html>.
- Gates, Bill. 2003. Toward a Spam-Free Future. <http://www.microsoft.com/mscorp/execmail/2003/06-24antispam.msp>.

- Gerrard, Meg, Frederick X. Gibbons, and Brad J. Bushman. 1996. "Relation between Perceived Vulnerability to Hiv and Precautionary Sexual Behavior." *Psychological Bulletin* 119 (3): 390-409.
- Greene, Shirley R., and Mark Kamimura. 2003. *Ties That Bind: Enhanced Social Awareness Development through Interactions with Diverse Peers*. Portland, Oregon.
- Hasan, M. R., and H. Hussin. 2010. Self Awareness before Social Networking: Exploring the User Behaviour and Information Security Vulnerability in Malaysia *Information and Communication Technology for the Muslim World (ICT4M), 2010 International Conference on*, doi: 10.1109/ict4m.2010.5971894.
- Hayati, Pedram, Kevin Chai, Vidyasagar Potdar, and Alex Talevski. 2009. Honeyspam 2.0: Profiling Web Spambot Behaviour *12th International Conference on Principles of Practise in Multi-Agent Systems (PRIMA '09), Berlin, Heidelberg: Springer-Verlag*.
- Hayati, Pedram, Kevin Chai, Alex Talevski, and Vidyasagar Potdar. 2010. "Behaviour-Based Web Spambot Detection by Utilising Action Time and Action Frequency." In *The 2010 International Conference on Computational Science and Applications (ICCSA 2010), Fukuoka, Japan*, 351-360 Springer-Verlag Berlin, Heidelberg ©2010
- Hayati, Pedram, Nazanin Firoozeh, Vidyasagar Potdar, and Kevin Chai. 2012. "How Much Money Do Spammers Make from Your Website?" In *CUBE 2012 Conference. , Pune, India*.
- Hayati, Pedram, and Vidyasagar Potdar. 2009. "Spammer and Hacker, Two Old Friends." In *3rd IEEE International Conference on Digital Ecosystems and Technologies (IEEE-DEST 2009) Istanbul, Turkey*, 290-294.
- Hayati, Pedram, Vidyasagar Potdar, Kevin Chai, and Alex Talevski. 2010. "Web Spambot Detection Based on Web Navigation Behaviour." In *24th IEEE International Conference on Advanced Information Networking and Applications (AINA 2010), Perth, Western Australia*, 797-803. doi: 10.1109/aina.2010.92.
- Hayati, Pedram, Vidyasagar Potdar, William F. Smyth, and Alex Talevski. 2010. "Rule-Based Web Spambot Detection Using Action Strings." In *The Seventh Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference (CEAS 2010), Redmond, Washington*.
- Hayati, Pedram, Vidyasagar Potdar, Alex Talevski, and Kevin Chai. 2010. "Web Spambot Characterising Using Self Organising Maps." *International Journal of Computer Systems Science and Engineering*.
- Hevner, Alan R., Salvatore T. March, Jinsoo Park, and Sudha Ram. 2004. "Design Science in Information Systems Research." *MIS Q.* 28 (1): 75-105.
- Hinde, S. 2002. "Spam, Scams, Chains, Hoaxes and Other Junk Mail." *Computer Security* 21 (7): 592-606.
- Ilger, Michael, Jurgen Straus, Wilfried Gansterer, and Christian Proschinger. 2006. *The Economy of Spam*.

- Indermaur, David. 1987. "Public Perception of Sentencing in Perth, Western Australia." *Australian & New Zealand Journal of Criminology* 20 (3): 163-183.
- iPermitMail. Spam Calculator. Accessed 24 December, <http://www.ipermittmail.com/resource/spamcalc.shtml>.
- Jeffries, Paul C. 2008. Application Spam. Accessed 29 November <https://blog.facebook.com/blog.php?post=10199482130>.
- Jennings, Richi. 2009. Cost of Spam Is Flattening - Our 2009 Predictions. <http://email-museum.com/2009/01/28/cost-of-spam-is-flattening-our-2009-predictions/>.
- Joshi, Shrijit, and Meghana Joshi. 2012. "Web Applications & Operating Cost of Data Center." *International Journal in Foundations of Computer Science & Technology (IJFCST)* 2 (5).
- Judge, Paul, Dmitri Alperovitch, and Weilai Yang. 2005. "Understanding and Reversing the Profit Model of Spam." In *Workshop on Economics of Information Security 2005 (WEIS 2005)*, Boston, MA, USA.
- Karidis, John, Jose E. Moreira, and Jaime Moreno. 2009. True Value: Assessing and Optimizing the Cost of Computing at the Data Center Level *New York, NY, USA*: ACM.
- Kim, Yeonbae, Yuri Park, Jeong-Dong Lee, and Jongsu Lee. 2006. "Using Stated-Preference Data to Measure the Inconvenience Cost of Spam among Korean E-Mail Users." *Applied Economics Letters* 13 (12): 795-800. doi: 10.1080/13504850500425287.
- Kshetri, Nir. 2006. "The Simple Economics of Cybercrimes." *Security & Privacy, IEEE* 4 (1): 33-39.
- Lang, Micheal, Jonathan Devitt, Sean Kelly, Andrew Kinneen, John O'Malley, and Darren Prunty. 2009. Social Networking and Personal Data Security: A Study of Attitudes and Public Awareness in Ireland *Management of e-Commerce and e-Government, 2009. ICMECG '09. International Conference on*, doi: 10.1109/ICMeCG.2009.105.
- Levine, Tamar, and Smadar Donitsa-Schmidt. 1998. "Computer Use, Confidence, Attitudes, and Knowledge: A Causal Analysis." *Computers in Human Behavior* 14 (1): 125-146. doi: [http://dx.doi.org/10.1016/S0747-5632\(97\)00036-8](http://dx.doi.org/10.1016/S0747-5632(97)00036-8).
- Li, Xinhui, Ying Li, Tiancheng Liu, Jie Qiu, and Fengchun Wang. 2009. The Method and Tool of Cost Analysis for Cloud Computing *IEEE International Conference on Cloud Computing (CLOUD '09)* doi: 10.1109/cloud.2009.84.
- Liu, Yiqun, Rongwei Cen, Min Zhang, Shaoping Ma, and Liyun Ru. 2008. Identifying Web Spam with User Behavior Analysis *4th international workshop on Adversarial information retrieval on the web (AIRWeb '08)*, Beijing, China: ACM.
- March, Salvatore T., and Gerald F. Smith. 1995. "Design and Natural Science Research on Information Technology." *Decision Support Systems* 15 (4): 251-266. doi: [http://dx.doi.org/10.1016/0167-9236\(94\)00041-2](http://dx.doi.org/10.1016/0167-9236(94)00041-2).
- Markines, Benjamin, Ciro Cattuto, and Filippo Menczer. 2009. Social Spam Detection *New York, NY, USA*: ACM.

- McAfee. 2009. *The Carbon Footprint of Email Spam Report*.
- McLaughlin, Anne Marie, James B Canavan, Emma J Adams, Ruth McDonagh, Harp Breet, Gerard J Fitzpatrick, and Maria B Donnelly. 2008. "A Survey of Mrsa Awareness and Knowledge among the General Public and Patients' Visitors." *Journal of Infection Prevention* 9 (5): 18-23.
- MessageLabs Intelligence. 2010. Messagelabs Intelligence: 2010 Annual Security Report. http://www.symanteccloud.com/mlireport/MessageLabsIntelligence_2010_Annual_Report_FINAL.pdf.
- Mihut, Marius, and Nicolae Tomai. 2010. "A Cost Model for the It Department." *Journal of Applied Quantitative Methods* 5 (2): 358-364.
- Mingers, John. 2003. "The Paucity of Multimethod Research: A Review of the Information Systems Literature." *Information Systems Journal* 13 (3): 233-249. doi: 10.1046/j.1365-2575.2003.00143.x.
- Modest Software. Cost of Spam Email Calculator. Accessed 24 December, <http://jspamfilter.com/costcalculator.cfm>.
- Morrissey, Brian. 2003. Report : Spam Cost Corporate America \$9b Last Year. <http://www.internetnews.com/IAR/article.php/1564761/Report+Spam+Cost+Corporate+America+9B+Last+Year.htm>.
- Net-security. 2008. Latest Spam Statistics Series Title: Help Net Security. <http://www.net-security.org/secworld.php?id=6056>.
- NetworkWorld.com. Spam Calculator. Accessed 24 December, <http://www.networkworld.com/spam/index.jsp>.
- Nowack, Marlene. 1997. "The Impact of the Internet on Statistical Organisations." *Statistical Journal of the UN Economic Commission for Europe* 14 (4): 345.
- Nucleus Research. 2003. *Spam: The Silent Roi Killer*.
- Nucleus Research. 2004. 2nd Annual Spam Report: Cost of Spam More Than Doubled in Past Year to \$1,934 Annually Per Employee. <http://nucleusresearch.com/news/press-releases/2nd-annual-spam-report-cost-of-spam-more-than-doubled-in-past-year-to-1934-annually-per-employee/>.
- Nucleus Research. 2007. *Spam : The Repeat Offender*.
- Numion. Download Time Calculator. Accessed 09/01, <http://www.numion.com/calculators/time.html>.
- Nunamaker, Jay F., Jr., and Minder Chen. 1990. Systems Development in Information Systems Research *System Sciences, 1990, Proceedings of the Twenty-Third Annual Hawaii International Conference on*, doi: 10.1109/hicss.1990.205401.
- Omar, Adnan, and Alfred Samman. 2011. Cost of Spam to Academic Institution *Association of Business Information Systems (ABIS 2011) Refereed Proceedings, Houston, Texas*,

- Patel, Chandrakant D., and Amip J. Shah. 2005. "Cost Model for Planning, Development and Operation of a Data Center." In *Technical Report: Hewlett Packard Lab*.
- PC Magazine. 2013. Botnet Definition. PCMag.com. Accessed 28 May, http://www.pcmag.com/encyclopedia_term/0,2542,t=botnet&i=38866,00.asp.
- Pfleeger, Shari Lawrence, and Gabrielle Bloom. 2005. "Canning Spam: Proposed Solutions to Unwanted Email." *Security & Privacy, IEEE* 3 (2): 40-47.
- Potdar, Vidyasagar, Farida Ridzuan, Pedram Hayati, Alex Talevski, Elham Afsari Yeganeh, Nazanin Firuzeh, and Saeed Sarencheh. 2010. "Spam 2.0: The Problem Ahead." In *Computational Science and Its Applications – Iccsa 2010*, eds David Taniar, Osvaldo Gervasi, Beniamino Murgante, Eric Pardede and Bernady O. Apduhan, 400-411. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Quinn, KJ Spike. 2006. *2006 New Zealand Computer Crime and Security Survey*. New Zealand.
- Ridzuan, Farida, and Vidyasagar Potdar. 2012. "Spam 2.0." In *Proceedings of the CUBE International Information Technology Conference, Pune, India*, 724-731. ACM.
- Ridzuan, Farida, Vidyasagar Potdar, and Wendy Hui. 2012. A Survey of Awareness, Knowledge and Perception of Online Spam *Computing and Convergence Technology (ICCCT), 2012 7th International Conference on*,
- Ridzuan, Farida, Vidyasagar Potdar, and Jaipal Singh. 2011. "Storage Cost of Spam 2.0 in a Web Discussion Forum." In *Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference, Perth, Australia*, 200-209. ACM. doi: 10.1145/2030376.2030400.
- Rockbridge Associates Inc. 2005. *2004 National Technology Readiness Survey*.
- Rockbridge Associates Inc. 2009. *2009 National Technology Readiness Survey Spam Report*.
- Rowland, Robin. 2003. Spam, Spam, Spam : The Cyberspace Wars Series Editor: C. N. Online. <http://www.cbc.ca/news/background/spam/>.
- Sara Radicati. 2009. *Email Statistics Report, 2009-2013*.
- Savage, Scott J., and Donald Waldman. 2005. "Broadband Internet Access, Awareness, and Use: Analysis of United States Household Data." *Telecommunications Policy* 29 (8): 615-633. doi: <http://dx.doi.org/10.1016/j.telpol.2005.06.001>.
- Scam Sniper. 2011. Phishing Scam Alert: Comment Spam Leads to Facebook Phishing Scam. Accessed 3 October, <http://scamsniper.blogspot.com.au/2011/06/phishing-scam-alert-commenting-spam.html>.
- Schadler, Ted. 2009. *Should Your Email Live in the Cloud? A Comparative Cost Analysis*.
- SecureMX Mail Scrubbing. Spam Cost Calculator. Accessed 24 December, <http://securemx.in/services/spamcalculator.html>.
- Shin, Youngsang, Minaxi Gupta, and Steven Myers. 2011. "Prevalence and Mitigation of Forum Spamming."

- Siponen, Mikko T. 2001. "Five Dimensions of Information Security Awareness." *SIGCAS Computers and Society* 31 (2): 24-29. doi: 10.1145/503345.503348.
- Sonnenreich, Wes, Jason Albanese, and Bruce Stout. 2006. "Return on Security Investment (Rosi) – a Practical Quantitative Model." *Journal of Research and Practice in Information Technology* 38 (1): 55-66.
- Sophos. 2008. *Security Threat Report 2008*.
- . 2010. <http://www.sophos.com/sophos/docs/eng/papers/sophos-security-threat-report-jan-2010-wpna.pdf>.
- . 2011. *Security Threat Report 2011*.
- SpamEater.com. Spam Cost Calculator. Accessed 24 December, <https://spameater.com/spam.calculator.html>.
- Spamfighter. Use Our Spam Calculator to See How Much Spam Is Costing Your Company Each Year. Accessed 24 December, http://www.spamfighter.com/SPAMfighter/Spam_Calculator.asp.
- Spamhaus. 2012. The Definition of Spam. Accessed 2 April, <http://www.spamhaus.org/consumer/definition/>.
- Stander, A., A. Dunnet, and J. Rizzo. 2009. A Survey of Computer Crime and Security in South Africa *Information Security South Africa (ISSA), University of Johannesburg, South Africa*.
- Stanton, Jeffrey M. 1998. "An Empirical Assessment of Data Collection Using the Internet." *Personnel Psychology* 51 (3): 709-725.
- Stringhini, Gianluca, Christopher Kruegel, and Giovanni Vigna. 2010. Detecting Spammers on Social Networks *26th Annual Computer Security Applications Conference (ACSAC '10), Austin, Texas*, doi: 10.1145/1920261.1920263.
- Sureka, Ashish. 2011. "Mining User Comment Activity for Detecting Forum Spammers in Youtube." *1103.5044*.
- Symantec Intelligence. 2012a. Symantec Intelligence Report: February 2012. http://www.symantec.com/theme.jsp?themeid=state_of_spam.
- Symantec Intelligence. 2012b. Symantec Intelligence Report: January 2012. http://www.symantec.com/theme.jsp?themeid=state_of_spam.
- Takemura, Toshihiko, and Hiroyuki Ebara. 2008. Spam Mail Reduces Economic Effects *2008 Second International Conference on the Digital Society*,
- Takemura, Toshihiko, and Atsush Umino. 2009. "A Quantitative Study on Japanese Internet Users' Awareness to Information Security: Necessity and Importance of Education and Policy." *World Academy of Science, Engineering and Technology* 60: 638-644.
- Ukai, Yasuharu, and Toshihiko Takemura. 2007. "Spam Mails Impede Economic Growth." *The Review of Socionetwork Strategies* 1 (1): 14-22.

- Urizar Jr , Guido G. , and Marilyn A. Winkleby. 2003. "Aids Knowledge among Latinos: Findings from a Community and Agricultural Labor Camp Survey." *Hispanic Journal of Behavioral Sciences* 25 (3): 295-311. doi: 10.1177/0739986303256911.
- Veen, Ytje JJ van der, Hélène ACM Voeten, Onno de Zwart, and Jan H Richardus. 2010. "Awareness, Knowledge and Self-Reported Test Rates Regarding Hepatitis B in Turkish-Dutch: A Survey." *BMC Public Health* 10:512.
- VicomSoft Ltd. Find out How Much Intergate Email Security Could Be Saving Your Organisation. Accessed 24 December, <http://www.vicomsoft.com/services/email-security/roi-calculator/>.
- WebpageFX Team. 2011. Spam: More Than an Annoyance? <http://www.webpagefx.com/blog/internet/spam-more-than-an-annoyance-infographic/>.
- Weible, R., and J. Wallace. 1998. "Cyber Research: The Impact of the Internet on Data Collection." *Marketing Research* 10 (3): 19-25.
- Wild, T. Cameron, Riley Hinson, John Cunningham, and Jason Bacchiochi. 2001. "Perceived Vulnerability to Alcohol-Related Harm in Young Adults: Independent Effects of Risky Alcohol Use and Drinking Motives." *Experimental and Clinical Psychopharmacology* 9 (1): 117-125.
- Windows & .NET Magazine. 2003. *The Secret Cost of Spam*.
- Wood, Jane, and G. Tendayi Viki. 2004. "Public Perceptions of Crime and Punishment." In *Forensic Psychology: Debates, Concepts and Practice*, ed. J. Adler. Cullompton, UK: Willan.
- Wygant, Steve, Danny Olsen, Vaughn Call, and Joseph Curtin. 2005. "Comparative Analyses of Parallel Paper, Phone, and Web Surveys: Some Effects of Reminder, Incentive and Mode." Conferences, workshops, tutorials, presentations Brigham Young University.
- Yan, Yuk Yee. 2009. "Awareness and Knowledge of Andropause among Chinese Males in Hong Kong." *American Journal of Men's Health* 4 (3): 231-236.

Appendices

Appendix I Web Survey Questionnaire

Public Awareness, Knowledge and Perception of Online Spam

This survey is created to:

- Assess public awareness, knowledge and perception of online spam.
- Evaluate the time in identifying online spam.

For this research purpose, we define online spam as the propagation of unsolicited, anonymous and mass content to infiltrate legitimate Web 2.0 applications such as forums, blogs, internet messaging services, social networking sites, wikis and video sharing sites. To fit the research purpose, please limit your participation to one time. This survey is anonymous. Thank you for taking the time to participate in this survey. Please certify that you are 18 years old or above.

- Yes
- No

Q1 On average, how many hours do you spent on using Internet in a day?

- Less than 1 hour
- 1-5 hours
- 6-9 hours
- More than 9 hours

Q2 What activities do you normally engage in when you use the Internet?

- Searching for information
- Gaming
- Chatting & Social Networking
- Email
- Other _____

Q3 Does your work/study relate to a technology field? (e.g: computing, communication, engineering)

- Yes
- No

Q4 Have you ever heard of online spam?

- Yes
- No

Q5 Have you had any of these experiences while browsing the Internet?

	Yes	No	Don't Know
Found pages that are only full of repeated keywords	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Redirected to an unrelated page from what is expected.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Found pages with repetitive links.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Received message considered as unwanted/suspicious/annoying on a web 2.0 application.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Account being hacked/hijacked.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Found pages with unrelated links.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Attacked by virus.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Received/Seen suspicious link on a web 2.0 application.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Found pages that are only advertising with very little content.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being asked username, passwords and/or credit card details by a trustworthy-looking entity in the Internet communication.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Received unwanted postings on your social network account .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Received unwanted friend requests on your social network account.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being tagged by unwanted parties on a web 2.0 application.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q6 My overall knowledge of online spam can best be described as:

- None
- Poor
- Fair
- Good
- Expert

Q7 Which of these actions do you think is considered as online spam?

	Yes	No	Don't Know
Found pages that are only full of repeated keywords	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Redirected to an unrelated page from what is expected.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Found pages with repetitive links.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Received message considered as unwanted/suspicious/annoying on a web 2.0 application.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Account being hacked/hijacked.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Found pages with unrelated links.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Attacked by virus.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Received/Seen suspicious link on a web 2.0 application.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Found pages that are only advertising with very little content.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being asked username, passwords and/or credit card details by a trustworthy-looking entity in the Internet communication.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Received unwanted postings on your social network account .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Received unwanted friend requests on your social network account.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being tagged by unwanted parties on a web 2.0 application.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q8 Assess the statement below and please choose the appropriate response for each item.

	True	Not True	Don't Know
Online spam can be used as a part of phishing attack.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online spam can be used to disseminate malware.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online spam can be used to promote affiliate websites.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
There is no difference between online spam and email spam.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
All online spam can be detected and treated using existing anti-spam techniques that are designed for email spam.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online spam has lower viewer impact than email spam.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Auto registration software can be used to register spam accounts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online spam can be found on legitimate websites.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online spam can be used to provide false information to users.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online spam can lead to other crimes such as fraud.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q9 Do you think that online spam is a problem?

- Yes
- Maybe
- No

Q10 For statements below, on a scale of 1 to 7, where 1 means strongly disagree and 7 means strongly agree, use the slider to indicate how strongly you disagree or agree.

_____ Virus is a serious problem.

_____ Hacking is a serious problem.

_____ Phishing is a serious problem.

_____ Computer fraud is a serious problem.

_____ Online spam is a serious problem.

_____ Email spam is a serious problem.

_____ Spyware is a serious problem.

_____ Worm is a serious problem.

Q11 Do you think that spamming is acceptable?

- Yes
- Maybe
- No

Q12 Do you think that confessed spammers should be punished?

- Yes
- Maybe
- No

Q13 Do you think that convicted spammers should be allowed to work in computing field?

- Yes
- Maybe
- No

Q14 Why do you think people spam? Please rank the statements below from 1-6, where 1 means the most relevant answer and 6 means the most irrelevant answer.

_____ To make money.

_____ For malicious reasons.

_____ To obtain a higher rank in search engine.

_____ To promote their product and services.

_____ For religious purposes.

_____ For fun.

Q15 How vulnerable are you to spam?

- Very vulnerable.
- Somewhat vulnerable.
- Indifferent.
- Not very vulnerable.
- Not at all vulnerable.

Q16 What is the likelihood that you will be spammed?

- Very Unlikely
- Unlikely
- Undecided
- Likely
- Very Likely

Q17 Do you think the screen below contains any spam?



- Yes
- No
- Don't Know

Please justify your answer:

Q18 Do you think the screen below contains any spam?



- Yes
- No
- Don't Know

Please justify your answer:

Q19 Do you think the screen below contains any spam?

 **Firefly Airlines** 9 May
20% off "Bran New Kiss" album showcase tickets!
Can't get enough of K-pop?

Here's your chance to join UKISS in celebrating the launch of their new album "Bran New Kiss" on June 10, 2011 at 8pm! It will be held at Dewan Wawasan, Menara PGRM in Cheras, Kuala Lumpur, Malaysia

Firefly Facebook page Fans will receive 20% OFF the tickets for three categories: Rock Zone (RM153), The Pit (RM330) and Package480 (RM480).

Those with a Package480 ticket are entitled to snap a group pix with the band itself!

Just cut and paste this link in your browser and key in the code to enjoy:

LINK: <https://www.fatdeal.com.my/index.php?referer=fyfb>
CODE: fy999

Attachment unavailable
The attachment source was deleted or the privacy settings on this attachment do not allow you to view it.

 **Firefly Airlines** 9 May
Can't get enough of K-pop?

Here's your chance to join UKISS in celebrating the launch of their new album "Bran New Kiss" on June 10, 2011 at 8pm! It will be held at Dewan Wawasan, Menara PGRM in Cheras, Kuala Lumpur, Malaysia

Firefly Facebook page Fans will receive 20% OFF the tickets for three categories: Rock Zone (RM153), The Pit (RM330) and Package480 (RM480).

Those with a Package480 ticket are entitled to snap a group pix with the band itself!

Just cut and paste this link in your browser and key in the code to enjoy:

LINK: <https://www.fatdeal.com.my/index.php?referer=fyfb>
CODE: fy999

Attachment unavailable
The attachment source was deleted or the privacy settings on this attachment do not allow you to view it.

- Yes
- No
- Don't Know

Please justify your answer:

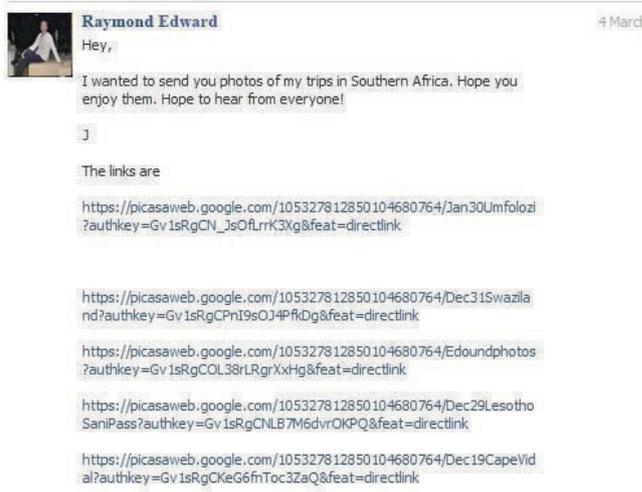
Q20 Do you think the screen below contains any spam?



- Yes
- No
- Don't Know

Please justify your answer:

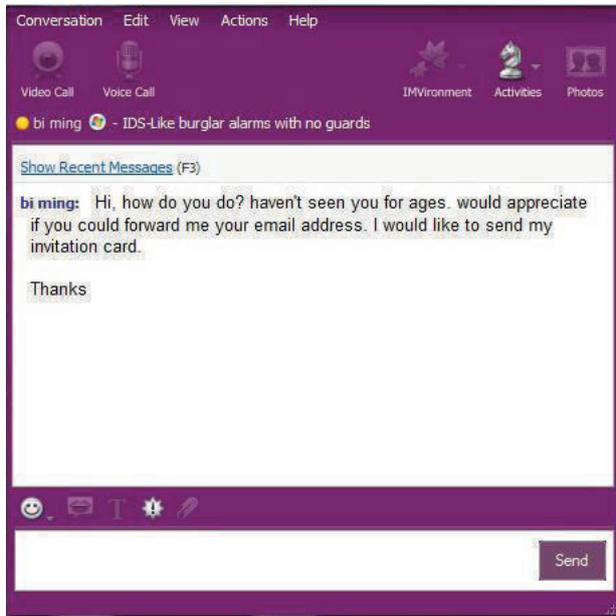
Q21 Do you think the screen below contains any spam?



- Yes
- No
- Don't Know

Please justify your answer:

Q22 Do you think the screen below contains any spam?



- Yes
- No
- Don't Know

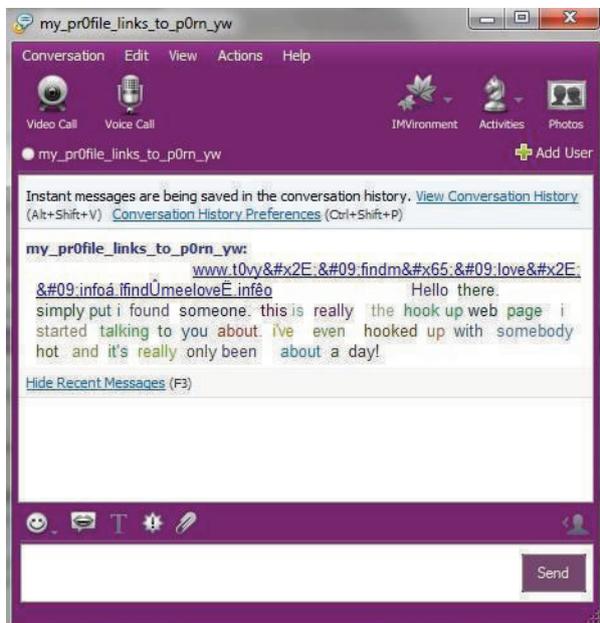
Q23 Do you think the screen below contains any spam?



- Yes
- No
- Don't Know

Please justify your answer:

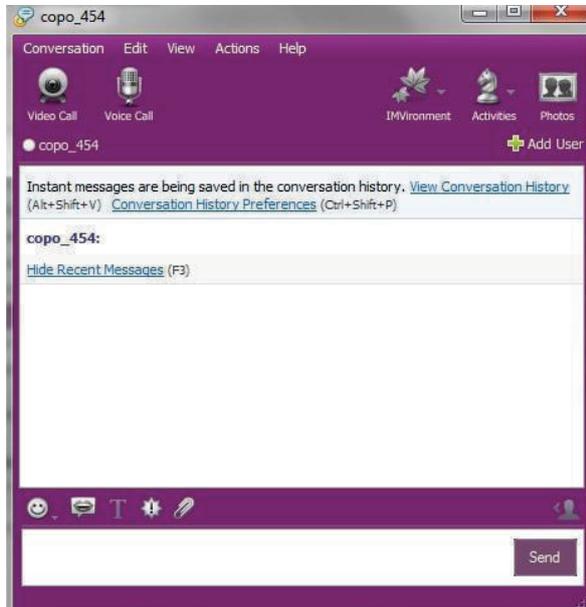
Q24 Do you think the screen below contains any spam?



- Yes
- No
- Don't Know

Please justify your answer:

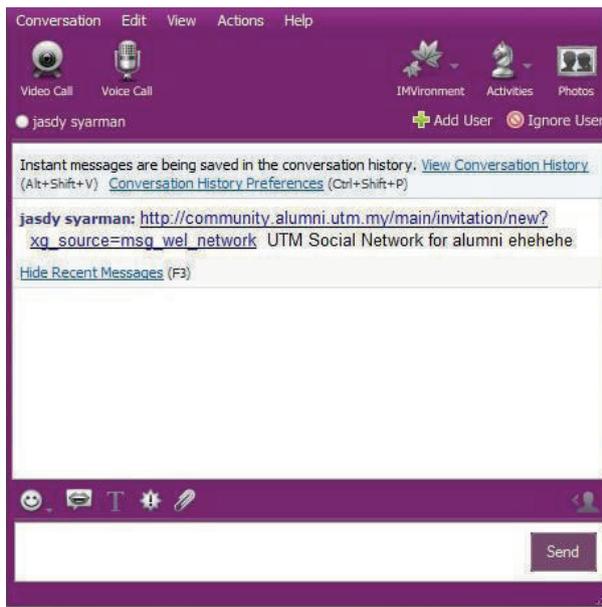
Q25 Do you think the screen below contains any spam?



- Yes
- No
- Don't Know

Please justify your answer:

Q26 Do you think the screen below contains any spam?



- Yes
- No
- Don't Know

Please justify your answer:

Q27 Please select your age group

- 18-24
- 25-34
- 35-49
- 50-64
- >64
- Refused to specify

Q28 Please select your gender

- Male
- Female
- Refused to specify

Q29 Please select your highest education level:

- Primary
- Secondary
- Certification
- Diploma/Advanced Diploma
- Undergraduate
- Postgraduate
- Refused to specify

Appendix II Selected Publications

- Farida Ridzuan, Vidyasagar Potdar and Wendy Hui. 2013. "Awareness, Knowledge and Perception of Online Spam." *Journal of Next Generation Information Technology (JNIT)* 4(3): 9-22.
- Farida Ridzuan, Vidyasagar Potdar and Wendy Hui. 2012. "A Survey of Awareness, Knowledge and Perception of Online Spam." In the 7th International Conference on Computing and Convergence Technology (ICCCT), Seoul, Korea, 1106-1110. IEEE Explore.
- Farida Ridzuan and Vidyasagar Potdar, 2012. "Spam 2.0." In Proceedings of the CUBE International Information Technology Conference, Pune, India, 724-731. ACM.
- Farida Ridzuan, Vidyasagar Potdar, and Jaipal Singh. 2011. "Storage Cost of Spam 2.0 in a Web Discussion Forum." In Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference, Perth, Australia, 200-209. ACM.
- Farida Ridzuan, Vidyasagar Potdar, Alex Talevski, and William F. Smyth. 2010. "Key Parameters in Identifying Cost of Spam 2.0." In 24th IEEE International Conference on Advanced Information Networking and Applications (AINA 2010), Perth, Western Australia, 789-796. IEEE Computer Society.

Awareness, Knowledge and Perception of Online Spam

¹Farida Ridzuan, ²Vidyasagar Potdar, ³Wendy Hui

^{1, First Author} *Anti Spam Research Lab, School of Information Systems, Curtin University,
faridahazwaniridzuan@gmail.com*

^{2, Corresponding Author} *Anti Spam Research Lab, School of Information Systems, Curtin
University, v.potdar@cbs.curtin.edu.au*

³*School of Information Systems, Curtin University, wendy.hui@curtin.edu.au*

Abstract

Online spam is a new way of spamming, using Web 2.0 applications as platforms. It can easily proliferate in spite of the first layer of security being in place, such as detection and prevention software, because of lack of awareness and knowledge on the part of the Internet users. It not only creates nuisance for the Internet users, it may also lead to bigger problems, like cybercrime involving hacking, phishing, etc. This paper presents the descriptive analysis of a web-based survey, conducted on 368 Internet users on their awareness, knowledge and perception of online spam. The purpose of the survey was to gauge the Internet users' awareness and knowledge of online spam, and investigate their perception of different aspects of the problem. To the best of our knowledge, it was the first survey conducted to highlight and investigate the issues involving online spam and, as such, the paper is a unique and pioneering contribution in the field.

Keywords: *Online Spam, Spam 2.0, Awareness, Knowledge, Perception.*

1. Introduction

Web 2.0 applications, such as forums, blogs, Internet messaging services, social networking sites, Wikis and video sharing sites, are much in use these days for personal or business purposes. Most Web 2.0 applications are supported by a platform that allows Internet users to create and share their own content [1]. However, spammers can also use this tool for their own benefit. This new threat of spam, posed by spammers to Web 2.0 applications, and to Internet users in general, is called online spam, or Spam 2.0.

Online spam is defined as the propagation of unsolicited, anonymous and mass content to infiltrate legitimate Web 2.0 applications [2, 3]. Spammers create online spam and share it through the network in the same way as various types of content created and shared by normal users. However, while normal users simply want to share information with, or obtain information from others, online spammers have different, sinister motives. They create spam in order to direct higher traffic to their sites and to advertise their products and services to generate revenues, or provide false information to, and steal valuable information from users [4, 5].

Online spam is a far more serious threat to online communities than email spam. Email spam, if it bypasses the filters, can still be read only by the owner of the inbox. The owner, in turn, can delete it by themselves. On the other hand, online spam can reach a large number of targeted and domain-specific users. Further, once an online spam is posted on a Web 2.0 application, it can typically be removed only by the administrators. The administrators may need some time before identifying and removing online spam posted on the application. Before that happens, however, its contents can be read by a large number of users and many of them may be seriously affected, especially when the online spam is used for phishing or fraud, or as a medium to spread spyware and malware [6].

Although recent works on both email and online spam are moving towards more behaviour-based solutions [7-11], currently online spam is mainly detected and prevented by using filters, in the manner of email spam [3]. However, there is no guarantee that these filters will work every time. If an online spam bypasses the filters, and if the spam campaign is attractive enough, gullible users can easily be

deceived. Online spam basically proliferates because of the users' lack of knowledge and awareness, and erroneous perceptions. This paper seeks to assess Internet users' level of knowledge and awareness about online spam, find out their perception of different aspects of the problem, and gain an understanding of their views on the threats they are exposed to while using the Internet.

The paper is organized as follows. Section 2 surveys the existing studies on public awareness, knowledge and perception of cyber-crime in general, as studies on online spam are rare, and leads to the derivation of survey questions. Section 3 describes the methodology used in the survey, including the survey design, sampling method, and the respondents' demographics. Section 4 presents the survey findings, while Section 5, after a comprehensive discussion, leads to the conclusion.

2. Literature Review

To the best of the authors' knowledge, there is no existing study that focuses on awareness, knowledge and perception issues in the context of online spam. To support this, further detailed explanation on each issue is being given below, divided into three subsections:

- Public Awareness of Online Spam
- Knowledge of Online Spam
- Perception of Online Spam

In these subsections, we highlight the importance of each issue. Any related definitions or terms have also been explained.

2.1. Public Awareness of Online Spam

A certain level of awareness of online spam is important not only for information technology or computer professionals but also for public users. Online spam campaigns are easily proliferated on the Internet because of the users' inability to distinguish between spam and non-spam. Awareness of at least simple and easy to identify online spams on the part of public users can help reduce successful online spam campaigns, or at least prevent them from being proliferated to other, easy-to-deceive Internet users. In addition, awareness of online spam can enable the users to report it to the administrators of the application who, in turn, can remove online spam posts more quickly. However, there is a lack of precise information on the extent of public awareness of online spam. This is probably because the concept of awareness as a tool to counter online spam has so far been taken lightly, as it does not seem to be in line with the scope of traditional engineering and hard computer science [12]. Greene and Kamimura define public or social awareness as "*naming the problem, speaking out, consciousness raising and researching*" [13]. As raising awareness could play a crucial role in raising consciousness of the issue, it is of vital importance to make the users aware of the issues of online spam, and commit them to tackling the problem.

Focusing on organizational aspects, most researchers emphasize the importance of examining whether organizations are conducting awareness training with regard to an issue, and further, investigating its costs, standards, policies and procedures [14, 15]. However, focusing on public users, the objective can only be to determine whether the users are aware of the issue. This type of research works similarly as in the fields of health and medicine. In the field of computers, several awareness studies have been undertaken to address a number of issues, such as computer crime and abuse [16] and information security [17-20].

A quantitative study on the Japanese Internet users' awareness of information security brings out the importance of providing education and evolving a policy based on individual attributes, by analysing variance based on non-parametric methods [17]. Awareness of information security in this study was investigated on the basis of 5 different indices, including recognition concerning individual information, recognition concerning illegal copying, recognition concerning countermeasures, recognition concerning the Internet, and awareness of the moral issues [17]. In this research, several

hypotheses were initially drawn up and tested, based on a sample size of 1483 [17]. However, the purpose of our research is not to test hypotheses.

Another study was conducted on the importance of having awareness before joining social networks, with regard to the user behaviour in using social network applications [18, 20]. The result was based on the data gathered through a survey involving close-ended questions aimed at 119 students [18]. Users' awareness of potential threats in using social networks was evaluated on the basis of their behaviour during the sign-up process, and information revealed on the social network. The study showed that lack of awareness while using social networks resulted in higher possible risks for the users to disclose their personal information, which leads to information security vulnerability [18].

Unlike these researches, which evaluated the awareness of the respondents indirectly, several other researchers simply asked direct questions to assess the respondents' awareness. For example, in a research on awareness, information sharing and privacy on Facebook, the respondents were questioned on Facebook rules and profile visibility through question like, "*Do you know whether Facebook provides for any tool to manage who can search for and find your profile?*" and "*Do you know what Facebook privacy settings are?*" [20]. Similarly, in another research on awareness of social networking and personal data security, the respondents were asked questions like, "*Do you know that Bluetooth devices/CDs/DVDs/Flash-drives can carry viruses?*", "*Are you aware of Trojan/worm/malware?*" and "*Have you ever experienced a virus, worm or other intruder on your computer?*" [19]. In another research on computer crime and abuse, questions such as, "*Have you ever heard of [certain Acts related to the Internet use in the UK]?*", were asked to assess the awareness of relevant legislation [16].

2.2. Knowledge of Online Spam

In this paper, we also examine the public knowledge of online spam. Initially, public users need to be aware of the issues involving online spam. This will prompt individuals to learn and acquire further knowledge about that particular issue. However, to the best of the authors' knowledge, there is no specific study in computer security carried out to address this issue. Hence, knowledge related questions were pooled from other fields, most importantly public health. In public health research, knowledge studies are commonly carried out, based on a certain existing knowledge scale, to examine people's understanding of certain sicknesses. However, in a field where no knowledge scale exists, symptoms, causes, suitable solutions and related information on that particular issue, can be used to evaluate knowledge.

2.3. Perception of Online Spam

In this paper, we took interest in exploring the perception towards crime and punishment, and fear of crime, crime in this context specifically referring to online spam. The detailed explanation of these aspects was covered in a study by J. Wood and G. T. Viki [21]. However, although not mentioned clearly, these aspects have also been explored in several works in the field of computer security, such as those of A. Stander, A. Dunnet, and J. Rizzo, and A. I. Al-Alawi and M. F. Abdelgadir [15, 22]. The questions asked in these surveys helped us determine how seriously online spam was viewed by the public.

2.3.1. Perception towards crime

Perception towards crime was explored by examining the respondents' attitude towards, and opinion about, this particular crime. Most of the researchers have tried to evaluate the relationship between the attitudes towards crime with socio-demographic victimization and fear of crime. However, they were typically focused on general crime [21, 30]. As for crime involving the Internet, the closest survey was on the respondents' perception of computer crime and abuse, such as spreading virus, viewing or altering someone else's data, theft of computer equipment, unauthorized copying of software, unauthorized copying of data, computer fraud and sabotage. Further questions were then asked, based on hacking activities as, according to the authors, "*Computer hackers represent the most 'hyped' forms of abuse in the mass media*" [16]. This survey was reportedly based on 175 public responses.

Another survey focusing on 712 college students was carried out with a different objective, that of comparing attitudes regarding information security ethics, and tested three hypotheses [23]. The respondents in this survey were given scenarios, and their responses were recorded, based on a five measure scale, that included ethical, acceptable, questionable, unethical and computer crime [23].

2.3.2. Perception towards punishment of crime

Perception towards the punishment of crime is understood to be people's punitive attitudes towards a particular crime. From the point of view of general crime, several researches have focused on finding relationship between punitive attitudes with different socio-demographic factors, such as gender and racial differences [21]. The results also vary, with some researchers managing to show that men have more punitive attitudes towards crime, even though similar results have not been achieved in other cases [21]. However, the purpose of our research was purely to gain an insight into a specific crime, i.e. online spamming.

There have not been many studies to gather public opinion towards a crime involving computers. The one coming closest was done by Dowland et al. [16], focusing on computer crime and abuse.

2.3.3. Fear of crime

As borne out by the research on general crime, fear of certain crimes may lead to a negative impact on individual behaviours and quality of life [21], as the residents will be afraid of certain neighbourhoods linked to that particular crime and this will further diminish the sense of community [21]. Similarly for the crime involving the Internet, the users might be afraid to use certain services provided through the Internet. The research by Al-Alawi and Abdelgadir did not assess the fear of computer related crime directly, but it did lead to a different viewpoint, i.e. to assess the perceived level of safety that the respondents had while doing online transactions. As reported by them, the perceived level of safety is certainly a factor in the willingness of the public to conduct online transactions [22]. Although a certain level of awareness is required for the users to be able to exercise caution and avoid being victimized, it should not go to the extent where the trust between business providers and the users as customers is broken. Previous researches on general crime have also examined the relationship between the fear of crime and attitudes towards crime and punishment, with the outcomes being both negative and positive [21]. Nonetheless, the objective of our research was to understand the public opinion on the fear of online spam.

This aspect has also been evaluated in some researches on awareness, information sharing and privacy on Facebook. In this paper, questions such as “*Specifically, how worried would you be if a [certain scenario] took place?*” were asked, and the respondents were required to rank their worries on a 7-point Likert scale [20].

3. Methodology

This paper aims to study and report the level of public awareness, knowledge and perception of online spam. In order to gain an in-depth understanding of this topic, we chose web-survey as our research instrument. This allowed us to reach the target group of the research, i.e. the Internet users. In addition, the web-survey also has the advantage of lower costs [24] and faster feedback [25]. The option provided in the software allowed the respondents to opt for compulsory questions that had to be answered by them, hence decreasing the quantity of missing data [26].

Data were collected from 368 Internet users. Only respondents who were at least 18 years of age were allowed to participate in the survey. For this study, we used Qualtrics Survey which is a surveying tool available to Curtin University students. The survey was personally administered by the authors.

3.1. Survey Design

Items were developed through a comprehensive review of literature, search of unpublished reports, and input from expert advisory panels. The questions were developed in accordance with Dilman's design method for internet surveys [27]. Questions were also modified iteratively through expert consensus, following which item analysis was used to reduce the item pool. Based on the experts' feedback, suitable minor modifications to the wording of questions were made.

The web survey questionnaire consisted of 29 questions. The survey was divided into three major sections shown in Figure 1. Section A consisted of basic demographic questions, followed by Section B which covered awareness, knowledge and perception of online spam. Finally, Section C consisted of questions about spam identification. The paper focuses on data from Section B of the survey.

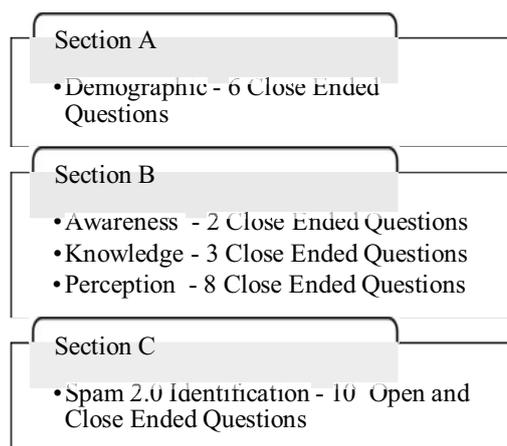


Figure 1. Survey Structure

On the knowledge part of online spam, we explored the perceived and actual knowledge. Perceived knowledge was directly inquired into, through the statements like “*My overall knowledge of online spam can best be described as ...*”. This statement allowed the users to rate themselves as having no knowledge at all, having poor knowledge, fair knowledge, good knowledge, or being an expert.

On the other hand, actual knowledge questions were derived from public health research [28]. We followed the practices of this research by measuring the respondents' knowledge using twenty statements, to which the respondents could respond as “True”, “Not true”, or “Don't know. Furthermore, statements for assessing knowledge were derived from a detailed literature review on, and definition of, Spam 2.0 [2-9, 29-35]. We asked the respondents “*Which of these actions do you think is considered as online spam?*” and presented ten statements given below:

- Finding pages that are only full of repeated keywords
- Beings redirected to an unrelated page from what was expected
- Finding pages with repetitive links
- Receiving messages considered as unwanted/suspicious/annoying on a Web 2.0 application
- Finding pages with unrelated links
- Receiving/seeing suspicious links on a Web 2.0 application
- Finding pages that only advertise with very little content
- Receiving unwanted postings on one's social network account
- Receiving unwanted friend requests on one's social network account
- Being tagged by unwanted parties on a Web 2.0 application

In addition, we also tested the respondents' knowledge by requesting, “*Assess the statement below and choose the appropriate response for each item*”, and presenting another ten statements to which

the respondents could answer as “True”, “Not true” or “Don’t know”. These statements were derived from our previous research, and are listed as follows:

- Online spam can be used as a part of phishing attack
- Online spam can be used to spread malware
- Online spam can be used to promote affiliate websites
- There is no difference between online spam and email spam
- All online spam can be detected and treated using existing anti-spam techniques that are designed for email spam
- Online spam has lower viewer impact than email spam
- Auto registration software can be used to register spam accounts
- Online spam can be found on legitimate websites
- Online spam can be used to provide false information to users
- Online spam can lead to other crimes such as fraud

As mentioned earlier, our research purpose was not to test any hypotheses, but purely to determine the perception of respondents towards online spam. Hence, we followed the research by Dowland et al. [16] and asked, “*Do you think online spam is a problem?*” and “*Do you think spamming is acceptable?*” to determine the respondents’ perception of online spam.

To determine attitudes towards suitable punishment, we again followed the research by Dowland et al. [16] by asking ‘*Do you think confessed spammers should be punished?*’ and “*Do you think convicted spammers should be allowed to work in the field of computing?*”. From these two questions, we were able to determine the public opinion towards spammers.

To assess the public perception on the fear of online spam, we modified the questions based on the research of M. Lang et. al [20] and asked questions such as, “*How vulnerable are you to spam?*” and “*What is the likelihood that you will be spammed?*”.

3.2. Sampling Method

The survey was conducted from 17th February 2012 to 7th April 2012. Participants were enlisted through link advertising. Personal invitations to participate in the survey were also sent through personal email lists. A personal URL linked to the survey was embedded in the invitation email message, also asking the participants to distribute the link to their contacts. Enlisting was also done through link advertising on Facebook using several personal accounts. In order to achieve more responses and reach a broader response age group, the instrument was distributed and publicized through invitation linked to the survey in a few community groups’ Facebook wall. Consequently, we were not able to determine the response rate.

3.3. Respondent Demographics

Figure 2 summarizes our self-reported respondent demographics. We considered gender, age-group and education level as sensitive questions, and hence the respondents were allowed to choose “Refuse to specify” as their answer. A total of 368 respondents completed the survey.

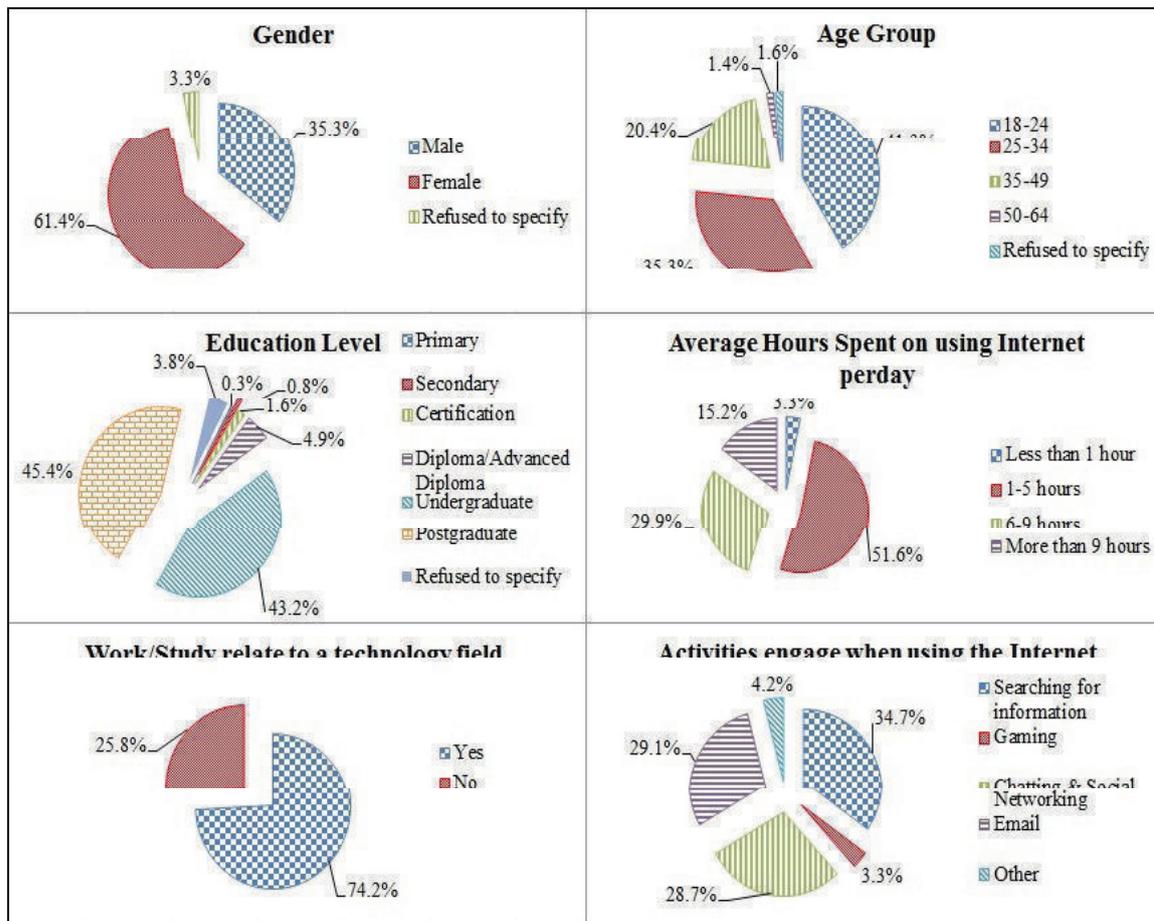


Figure 2. Summary of Respondent Demographics

4. Descriptive Survey Findings

In this section, we present our findings on the Internet users' awareness, knowledge and perception of online spam.

4.1. Awareness of Online Spam

Participants were asked if they had ever heard of online spam. This question was used as a measure of online spam awareness. It was followed by another question to determine actual awareness. In this particular question, we considered the participants as having awareness if they had encountered any of the cases considered as online spam.

- Have you ever heard of online spam? (Yes/No)

91.6% (337 out of 368) of respondents stated that they had heard of online spam while only 8.4% stated that they had never heard of online spam. This indicates that the public's perceived awareness is quite high.

- Have you had any of these experiences while browsing the Internet? (Yes/No/Don't know)

In order to make it consistent with the subsequent questions in the knowledge section, we listed some suspicious activities. If the respondents had encountered any of the activities considered spam, we considered them as having awareness. The summary of the results is represented in Figure 3.

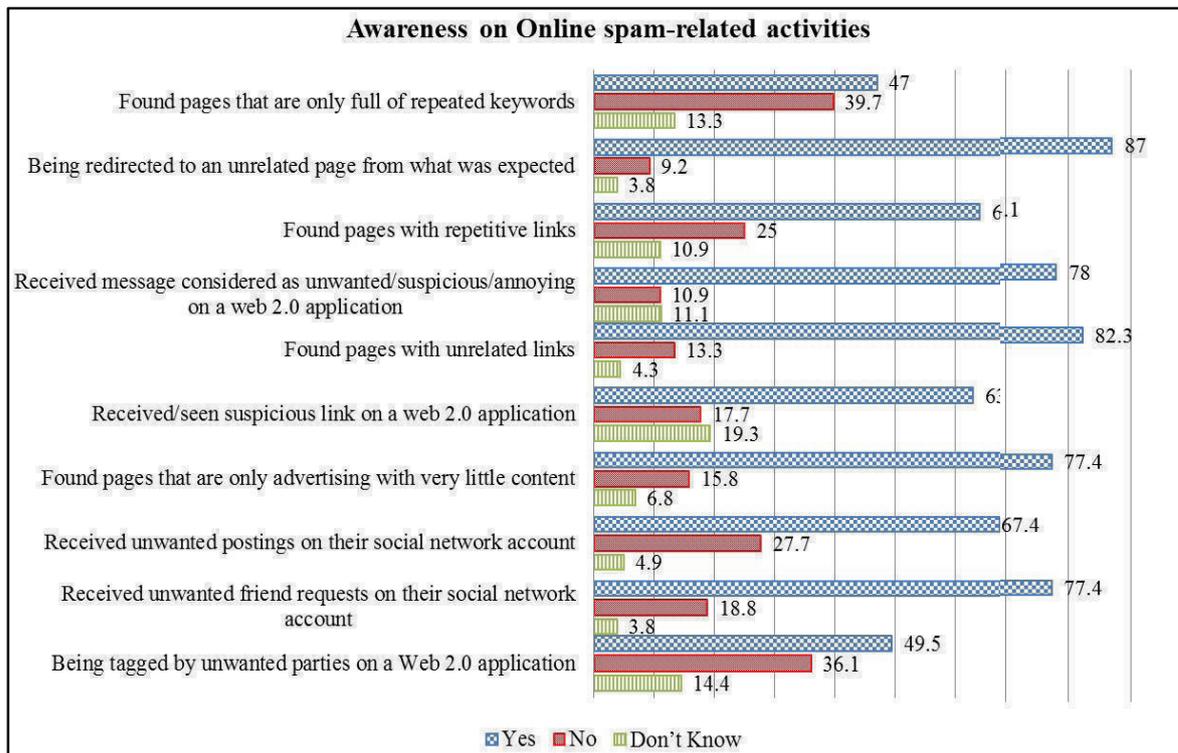


Figure 3. Respondents' Awareness of Online Spam-related Activities

- Knowledge of Online Spam

Participants were asked to rate their overall knowledge of online spam. This item measured the respondents' perceived knowledge, followed by questions to determine their actual knowledge.

- My overall knowledge of online spam can best be described as (None/Poor/Fair/Good/Expert).

Only 1.6% considered themselves as having no knowledge at all about online spam. 26.4% of the respondents considered themselves as having poor knowledge. More than half of the total respondents (53%) considered themselves to have a fair knowledge of online spam while 15.8% rated their knowledge of online spam as good. There were only 3.3% of the respondents who considered themselves as experts on online spam.

- Which of these actions do you think is considered online spam?

We listed several questions containing suspicious activities and asked the respondents to identify whether they considered these activities as online spam or not. Again, we allowed the respondents to choose "Don't know" as their answer. Figure 4 shows the percentages of the respondents who answered in "Yes", "No" and "Don't know".

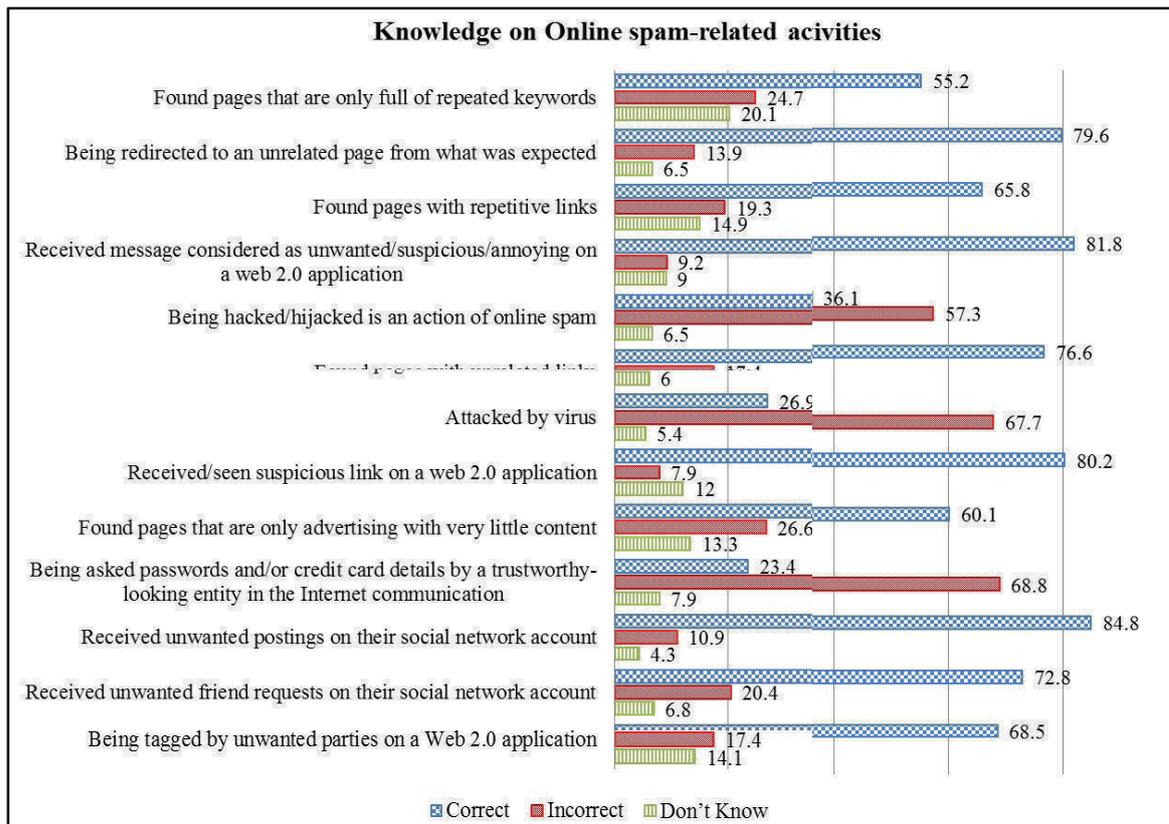


Figure 4. Respondents' Knowledge of Online Spam-related Activities

- Assess the statement below and choose the appropriate response for each item. (True/Not True/Don't Know)

For this question, we listed several statements regarding online spam and asked the respondents to decide whether the statements were true or false. We also provided the "Don't know" option. Most of the questions were basic and easy, but we also purposely included two technical and difficult questions. Percentages of the respondents who gave "True", "False" and "Don't know" answers are given in Figure 5.

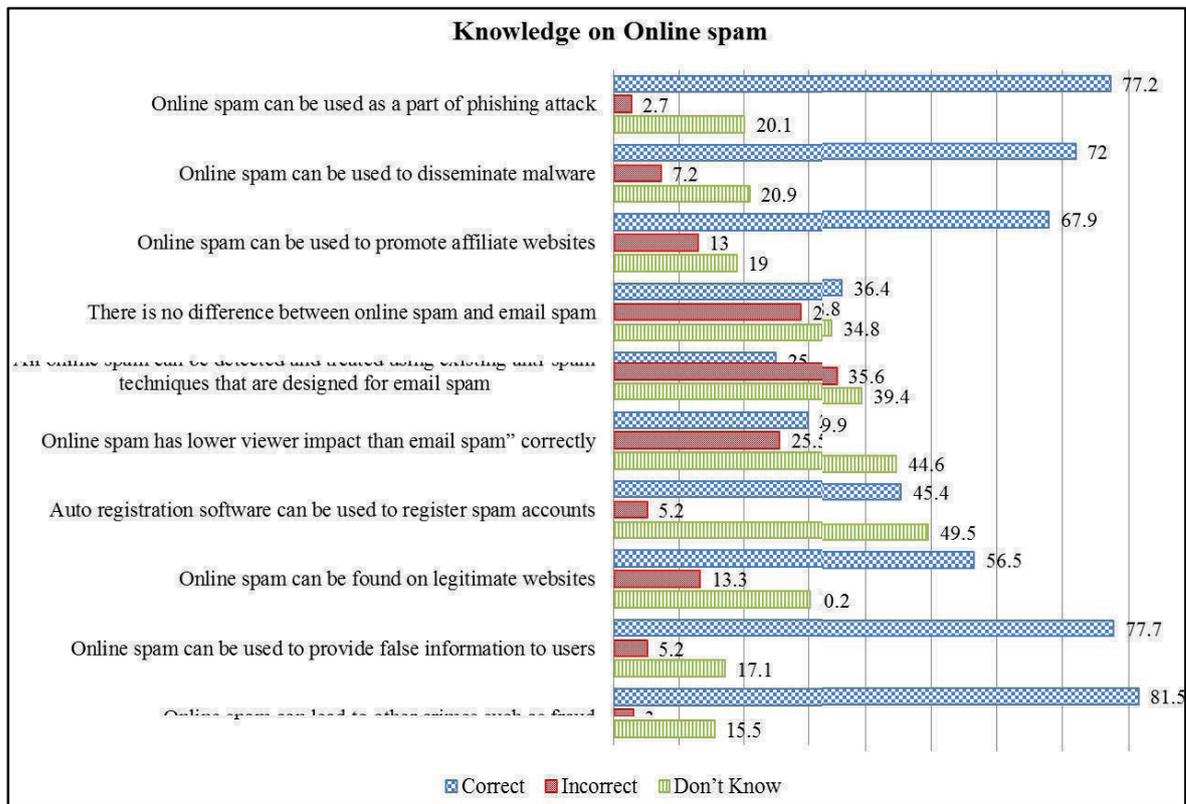


Figure 5. Respondents' Knowledge of Online Spam

4.2. Perception of Online Spam

In this section, we present the descriptive analysis of the responses to the questions related to the perception of online spam. The questions include the respondents' perception of spamming activity, their attitudes towards spammers, and their perception of the fear of crime. Figure 6 shows the respondents' answers to the questions "Do you think that online spam is a problem?" and "Do you think that spamming is acceptable?". 77.2% of the respondents thought that online spam was a problem, 2.4% thought that online spam was not a problem, while 20.4% chose "Maybe" as their answer. 72.8% of the respondents thought that spamming was unacceptable, 6.3% thought that online spam was acceptable, while 20.9% chose "Maybe" as their answer.

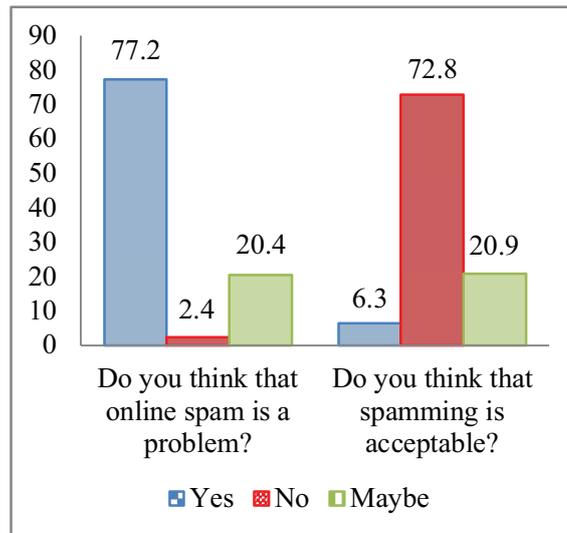


Figure 6. Respondents' Perception of Online Spam

Figure 7 shows the results for public perception regarding punishment to online spammers. 57.1% of the respondents thought that confessed spammers should be punished, 4.9% said that confessed spammers should not be punished, while 37.8% chose "Maybe" as their answer. When asked whether they thought that convicted spammers should be allowed to work in the field of computing, 24.5% of the respondents chose "Yes", 35.5% chose "No", and 39.9% chose "Maybe" as their answers.

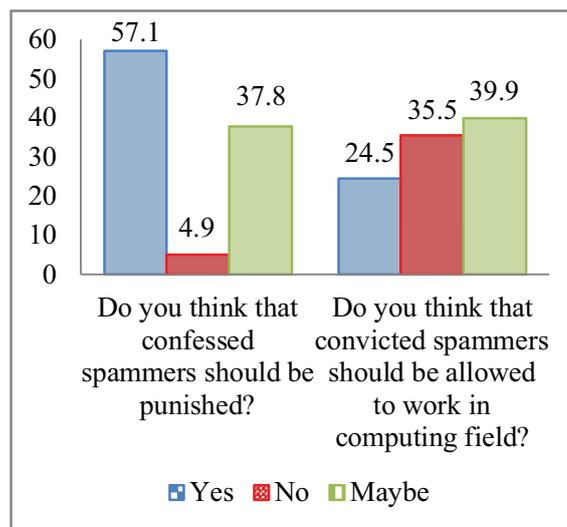


Figure 7. Respondents' Perception of Punishment for Online Spam

Figure 8 shows the results on the respondents' perceived vulnerability and their fear of crime. 12.2% of the respondents stated that they were very vulnerable to spam. 41.3% said that they were somewhat vulnerable, 21.2% said that they were indifferent, 22.6% said that they were not very vulnerable, and another 2.7% stated that they were not at all vulnerable. For the second question on the fear of crime, 11.4% of the respondents thought that it was very unlikely for them to be spammed, 21.7% thought that it was unlikely for them to be spammed, 32.9% stated that they were undecided, while 27.4% said that it was likely that they could be spammed. Only 6.5% stated that they were very likely to be spammed.

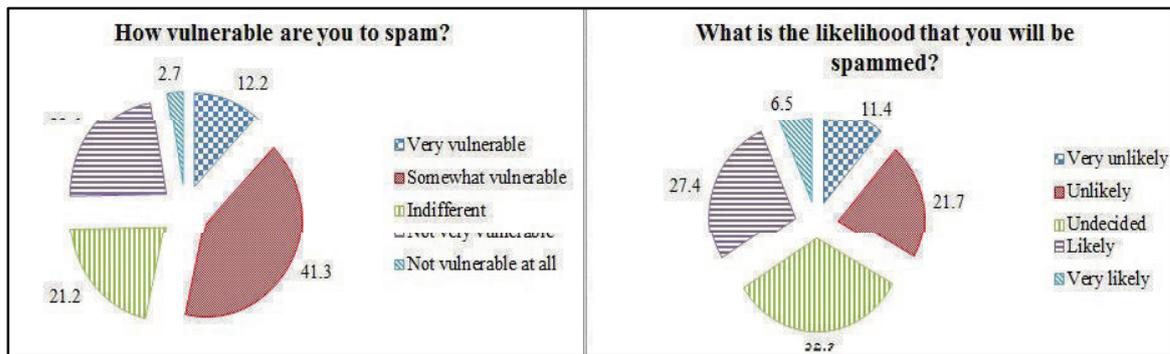


Figure 8. Respondents' fear of Online Spam

5. Discussion and Conclusion

As can be seen from the survey findings, perceived awareness of online spam seems to be quite high. However, it does not translate to actual awareness. Certain spamming activities are recognized by most of the Internet users as the respondents either answered "Yes" or "No", and the percentage of the respondents who answered "Don't know" was quite low. However, there certainly are spamming activities that are not recognized by many Internet users, causing most of the respondents to choose "Don't know" as their answer. We found that the spamming techniques that fall under this category are mostly those that occur in social networking sites, using links and tags to embed spam content.

Based on the response to the questions on actual awareness, it is evident that most of the respondents were aware, could identify and had had experiences with basic spamming techniques used by spammers, such as "Redirecting to an unrelated page from what is expected" and "Finding pages with unrelated links". The percentages of respondents choosing "Don't know" were low for both of these questions (3.8% and 4.3% respectively). However, there are a few spamming techniques, such as "Finding pages that are only full of repeated keywords", "Being tagged by unwanted parties" and "Receiving/seeing suspicious links on a Web 2.0 application", that seem to be unfamiliar to the Internet using public. It is possible that the respondents had never encountered these activities, but it is also likely that, even if they had encountered these, they did not know that it was online spam. The percentages of respondents who answered "Don't know" to these questions were relatively high (13.3%, 14.4% and 19.3% respectively), compared to other questions.

Most respondents perceived themselves as having a fair amount of knowledge about online spam. Not many categorized themselves as having no knowledge about online spam, nor did many claim to be an expert in the field. However, the questions on actual knowledge reveal that some of the non-spam activities are mistakenly viewed as online spam. These include "Account being hacked/hijacked", "Attacked by virus" and "Being asked username, password and/or credit card details by a trustworthy-looking entity in Internet communication".

Aside from the percentages of wrong answers, the percentages of respondents choosing "Don't know" are also important. As expected, the questions on the technical side of online spam, and the ones that we considered difficult, had the highest percentages of respondents choosing "Don't know". These questions were "Online spam has lower viewer impact than email spam" and "Auto registration software can be used to register spam accounts" (44.6% and 49.5% respectively).

Considering those who answered the questions incorrectly, and those who chose "Don't know" as not knowledgeable, it can be seen that there are some facts which are not well-understood by the general public. 75 % of the respondents had a misconception that all online spam can be detected and treated using existing anti-spam techniques designed for email spam. 63.6% of the respondents did not

know that there is a difference between online spam and email spam. Other facts about online spam were known to at least more than 50% of the respondents.

Most respondents agreed that online spam is a problem. The seriousness of online spam is reflected by the fact that the majority of the respondents thought that spamming was unacceptable. However, it is also astounding that 6.3% of them considered spamming as acceptable. The response to the next question also shows the seriousness of online spam. More than half of the respondents thought that confessed spammers should be punished. Nevertheless, 24.5% of the respondents were still of the view that convicted spammers should be allowed to work in the field of computing.

We used the questions on perceived vulnerability to divide the respondents into those who agreed that they were vulnerable to spam and those who thought they were not. More than half of the respondents (53.5%) thought that they were vulnerable to spam. However, there still were 25.3% who considered themselves as invulnerable. The response to the next question shows a similar pattern, with 33.1% of the respondents thinking that there was a low likelihood for them to be spammed while 33.9% thinking that they were highly likely to be spammed. There could be two reasons for low perceived vulnerability. The respondents might be thinking that the Web 2.0 applications provide sufficient protection against online spam. They may also be under the impression that they have enough knowledge to handle online spam.

To conclude, even though most of the Internet users are aware of online spam, it is also apparent that many have inadequate basic knowledge about it. Therefore, making the Internet users aware of these aspects may reduce the severity of the problems posed by online spam. Further, although most Internet users identify online spam as a serious problem, their attitude towards punishing spammers severely is unclear, further giving credence to the necessity of spreading awareness of online spam.

6. References

- [1] T. O'Reilly. (2005, 3 October). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1008839
- [2] P. Hayati, V. Potdar, A. Talevski, N. Firoozeh, S. Sarenche, and E. A. Yeganeh, "Definition of spam 2.0: New spamming boom", In the IEEE International Conference on Digital Ecosystems and Technologies (DEST 2010), Dubai, UAE, pp. 580-584, 2010.
- [3] P. Hayati and V. Potdar, "Toward Spam 2.0: An evaluation of Web 2.0 anti-spam methods", In the 7th IEEE International Conference on Industrial Informatics Cardiff, Wales, UK, pp. 875-880, 2009.
- [4] V. Potdar, F. Ridzuan, P. Hayati, A. Talevski, E. A. Yeganeh, N. Firuzeh, and S. Sarencheh, "Spam 2.0: The Problem Ahead", In The 2010 International Conference on Computational Science and Applications (ICCSA 2010), Fukuoka, Japan, pp. 400-411, 2010.
- [5] P. Hayati, and V. Potdar, "Evaluation of spam detection and prevention frameworks for email and image spam: a state of art", In Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services (iiWAS '08), Linz, Austria, pp. 520-527, 2008.
- [6] F. Ridzuan, V. Potdar, and J. Singh, "Storage cost of spam 2.0 in a web discussion forum", In Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference, Perth, Australia, pp. 202-209, 2011.
- [7] P. Hayati, K. Chai, V. Potdar, and A. Talevski, "HoneySpam 2.0: Profiling Web Spambot Behaviour", In 12th International Conference on Principles of Practise in Multi-Agent Systems (PRIMA '09), Springer/Berlin Heidelberg, pp. 335-344, 2009.
- [8] P. Hayati, K. Chai, A. Talevski, and V. Potdar, "Behaviour-Based Web Spambot Detection by Utilising Action Time and Action Frequency", In The 2010 International Conference on Computational Science and Applications (ICCSA 2010), Fukuoka, Japan, pp. 351-360, 2010.
- [9] P. Hayati, V. Potdar, K. Chai, and A. Talevski, "Web Spambot Detection Based on Web Navigation Behaviour", In the 24th IEEE International Conference on Advanced Information Networking and Applications (AINA 2010), Perth, Western Australia, pp. 797-803, 2010.
- [10] Z. Zhuangzheng, "Research on The E-mail Worm Propagation Model Based on User Behaviors," JDCTA: International Journal of Digital Content Technology and its Applications, AICIT, vol. 6, no. 22, pp. 239-245, 2012.
- [11] L. Hongtao, Y. Hongfeng, L. Zhaoyu, and W. Yu, "Classification and Detection of the Micro-blogging Users," IJACT: International Journal of Advancements in Computing Technology, AICIT, vol. 4, no. 20, pp. 432-441, 2012.

- [12] M. T. Siponen, "Five dimensions of information security awareness", SIGCAS Computers and Society, ACM, vol. 31, no. 2, pp. 24-29, 2001.
- [13] S. R. Greene and M. Kamimura, "Ties that Bind: Enhanced Social Awareness Development Through Interactions with Diverse Peers", Annual meeting of the Association for the Study of Higher Education, Portland, Oregon, pp. 213-228, 2003.
- [14] K. S. Quinn, "2006 New Zealand Computer Crime and Security Survey", vol. 2, Alpha Omega Group Publications, 2006.
- [15] A. Stander, A. Dunnet, and J. Rizzo, "A Survey of Computer Crime and Security in South Africa", In Proceeding of Information Security South Africa (ISSA) Conference, University of Johannesburg, South Africa, pp. 217-228, 2009.
- [16] P. S. Dowland, S. M. Furnell, H. M. Illingworth, and P. L. Reynolds, "Computer crime and abuse: A survey of public attitudes and awareness," Computers & security, Elsevier, vol. 18, no. 8, pp. 715-726, 1999.
- [17] T. Takemura and A. Umino, "A quantitative study on Japanese internet users' awareness to information security: Necessity and importance of education and policy", World Academy of Science, Engineering and Technology, vol. 60, pp. 638-644, 2009.
- [18] M. R. Hasan and H. Hussin, "Self awareness before social networking: Exploring the user behaviour and information security vulnerability in Malaysia", In the 2010 International Conference on Information and Communication Technology for the Muslim World (ICT4M), pp. C-7 – C-12, 2010.
- [19] M. Lang, J. Devitt, S. Kelly, A. Kinneen, J. O'Malley, and D. Prunty, "Social Networking and Personal Data Security: A Study of Attitudes and Public Awareness in Ireland", In ICMECG '09 Proceedings of the 2009 International Conference on Management of e-Commerce and e-Government, pp. 486-490, 2009.
- [20] A. Acquisti and R. Gross, "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook," In Privacy Enhancing Technologies. Springer Berlin/Heidelberg, vol. 4258, pp. 36-58, 2006.
- [21] J. Wood and G. T. Viki, "Public perceptions of crime and punishment", In Forensic psychology: Debates, concepts and practice, Cullompton, UK: Willan Publishing, 2004.
- [22] A. I. Al-Alawi and M. F. Abdelgadir, "An Empirical Study of Attitudes and Opinions of Computer Crimes: A Comparative Study between U.K. and the Kingdom of Bahrain", Journal of Computer Science, Science Publications, vol. 2, pp. 223-235, 2006.
- [23] A. L. Harris, "IS ethical attitudes among college students: A comparative study", In the Proceedings of the Information Systems Education Conference 2000 (ISECON 2000), Philadelphia, pp. 801-807, 2000.
- [24] R. Weible and J. Wallace, "Cyber Research: The Impact of the Internet on Data Collection", Marketing Research, American Marketing Association, vol. 10, no. 3, pp. 19-31, 1998.
- [25] M. Nowack, "The impact of the Internet on statistical organisations," Statistical Journal of the UN Economic Commission for Europe, IOS Press, vol. 14, no. 4, pp. 345, 1997.
- [26] J. M. Stanton, "An Empirical Assessment of Data Collection Using the Internet", Personnel Psychology, Wiley Online Library, vol. 51, no. 3, pp. 709-725, 1998.
- [27] D. A. Dillman, Mail and Internet Surveys : The Tailored Design Method -- 2007 Update with New Internet, Visual, and Mixed-Mode Guide, John Wiley & Sons, Inc., 2007.
- [28] Y. J. v. d. Veen, H. A. Voeten, O. d. Zwart, and J. H. Richardus, "Awareness, knowledge and self-reported test rates regarding Hepatitis B in Turkish-Dutch: a survey", BMC Public Health, BioMed Central, vol. 10, pp. 512, 2010.
- [29] V. Potdar, Y. Like, N. Firoozeh, D. Mukhopadhyay, F. Ridzuan, and D. Tejani, "The changing nature of Spam 2.0," In the Proceedings of the CUBE International Information Technology Conference, Pune, India, pp. 826-831, 2012.
- [30] P. Hayati, and V. Potdar, "Spam 2.0 State of the Art", International Journal of Digital Crime and Forensics (IJDCF), vol. 4, no. 1, pp.17-36, 2011.
- [31] P. Hayati, V. Potdar, A. Talevski, and K. Chai, "Characterisation of web spambots using self organising maps", International Journal of Computer Systems Science & Engineering (IJCSSE), vol. 26, no. 2, pp.87-96, 2011.
- [32] P. Hayati, N. Firoozeh, V. Potdar and K. Chai, "How much money do spammers make from your website?", In Proceedings of the CUBE International Information Technology Conference, Pune, India, pp. 732-739, 2012.
- [33] P. Hayati, V. Potdar, A. Talevski, and W. F. Smyth, "Rule-based on-the-fly web spambot detection using action strings", In Proceedings of the Collaboration, Electronic messaging, Anti-Abuse and Spam Conference (CEAS 2010), Redmond, USA, pp. 13-14, 2010.
- [34] F. Ridzuan, V. Potdar, and A. Talevski, "Factors involved in estimating cost of email spam", In Proceedings of the International Conference on Computational Science and Its Applications (ICCSA 2010), Fukuoka, Japan, March 23-26, pp 383-399, 2010.
- [35] F. Ridzuan, V. Potdar, A. Talevski, and W. F. Smyth, "Key Parameters in Identifying Cost of Spam 2.0", In Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications (AINA), Perth, Australia, April 20-23, pp. 789-796, 2010.

A Survey of Awareness, Knowledge and Perception of Online Spam

Farida Ridzuan

Anti-Spam Research Lab
School of Information Systems
Curtin University
Perth, Western Australia
farida.mohdridzuan@postgrad.curtin.edu.au

Vidyasagar Potdar

Anti-Spam Research Lab
School of Information Systems
Curtin University
Perth, Western Australia
v.potdar@cbs.curtin.edu.au

Wendy Hui

School of Information Systems
Curtin University
Perth, Western Australia
wendy.hui@curtin.edu.au

Abstract— Online spam does not only create nuisance for the Internet users, they could also lead to further and bigger problems such as hacking, phishing, etc. It is well known that online spam are currently handled using detection and prevention methods. Despite the effectiveness of these methods in reducing the proliferation of online spam, it is also known that spammers will continue to find ways to promote spam. They deceive users who may lack awareness and knowledge about this crime. Therefore, it is important to investigate Internet users' awareness of online spam, how much knowledge they actually have about online spam, and their perception of online spam. This paper provides a descriptive analysis of public awareness, knowledge and perception drawn out from a web based survey involving 368 Internet users. To the best of our knowledge, this is the first such survey that has ever been conducted to highlight these issues.

Keywords—component; Online spam, Spam 2.0, survey, awareness, knowledge, perception.

I. INTRODUCTION

Web 2.0 applications such as forums, blogs, Internet messaging services, social networking sites, wikis and video sharing sites are often used either for personal or business purposes. Most Web 2.0 applications are supported by a platform that allows Internet users to create and share their own contents [1]. However, spammers can use this freedom for their own benefits. This new threat imposed by spammers is called online spam.

Online spam is defined as the propagation of unsolicited, anonymous and mass content to infiltrate legitimate Web 2.0 applications [2, 3]¹. Spammers create online spam and share it through the network the same way as normal users share their contents. However, their underlying motive is different from that of normal users who simply want to obtain and share information with others. Spammers create spam in order to generate revenue, obtain higher traffic to their sites, advertise

their products and services, provide false information to users, and steal valuable information from users [4].

Online spam is a threat to the online communities. Compared to email spam attack, online spam can reach many targeted and domain-specific users. Furthermore, email spam can only be read by the inbox's owner. Once an email spam bypasses the filter, the users are able to delete it by themselves. On the other hand, the contents of online spam can be read by many users [5]². If an online spam was posted on a Web 2.0 application, it typically can be removed only by an administrator. An administrator may require a bit of time before removing online spam that is found on a Web 2.0 application. Before that happens, many users may be affected, especially when the online spam is used as a medium to spread spyware, malware, phishing and fraud.

Currently, online spam is prevented and detected using filters [3]. However, there is no guarantee that these filters work 100% of the time. If an online spam bypasses the filters and if the spam campaign is attractive, unknowledgeable users may be deceived. Thus, it can be seen that online spam proliferates because of user's lack of awareness, lack of knowledge, and erroneous perceptions that influence how they handle online spam. This paper seeks to assess Internet users' awareness, knowledge and perception of online spam, and gain the understanding of Internet users' views about the threats that they are exposed to while using the Internet.

To the best of authors' knowledge, there is no existing study which focuses on awareness, knowledge and perception issues in the context of online spam. An exploratory study was done to assess the level of understanding of South African organisations on general computer crime. However, they only reported about awareness training, respondents' behaviours and attitudes in reporting computer crime while focusing on other issues such as costs, policies and procedures [6].

¹ We took the definition of online spam similarly as Spam 2.0. However, in this research, we do not use the term Spam 2.0 as it sounds too technical.

² In our previous paper, we described these relationship as 1:1 for email spam and 1:M for online spam. This is the reason why online spam has higher viewer impact.

Similarly, in research [7], only awareness training are reported while focusing on other issues such as cost, standards, policies and procedures.

The closest study was done in 1999 in order to determine public attitude and awareness on cybercrimes [8]. However, computer crimes listed in this research covers on virus, data privacy, theft of computer equipment, software piracy, fraud and sabotage. Another research is done focusing on attitudes and opinions of computer crimes [9], but they also focused only on computer crimes listed in [8]. In this study, they made a comparative study between U.K. and the Kingdom of Bahrain to test the hypothesis whether the perceived level of safety is a factor in the willingness of the public to conduct online transactions. Nonetheless, our objectives is to focus on online spam, includes the awareness, knowledge and perception of Internet community towards online spam.

This paper is organized as follows. Section 2 presents the method used in this survey. This section will describe the survey design, sampling method and respondent demographics. A descriptive of our survey findings will be further presented in Section 3. Finally, Section 4 presents discussion and conclusion.

II. METHOD

This paper aims to study and report the public awareness, knowledge and perception of online spam. In order to gain an in-depth understanding of this topic, we chose web survey as the research instrument. This allows us reach the research target group, which are the Internet users. In addition, using web survey has the advantage of lower cost [10] and faster feedback [11]. The option provided in the software allows researchers to opt for compulsory questions that have to be answered by users hence decreasing missing data [12].

Data were collected from 368 Internet users. Only respondents who are at least 18 years old were allowed to participate in the survey. For this study, we used Qualtrics Survey, which a surveying tool available to Curtin University students and personally administered by the author.

A. Survey Design

The web survey questionnaire consists of 29 questions. The survey is divided into four major sections. The first section consists of basic demographic questions, followed by Section 2, which covers awareness, knowledge and perception of online spam. Section 3 consists of questions about spam identification. Finally, Section 4 covers the sensitive questions such as age and income. This paper focuses on data from Section 2 of the survey.

This survey consists mainly of closed-ended questions developed from the existing literature. Some of the questions require the survey participants to assess their awareness and knowledge directly, but some of the questions measures these variables indirectly. Basically, the participants were asked to answer the following questions:

- Have you ever heard of online spam?
- Have you had any experience being spammed when using the Internet?
- How do you rate your overall knowledge of online spam?
- How much do you actually know about online spam?
- Do you think that online spam is a problem?
- Do you think that spamming is acceptable?
- Do you think that confessed spammers should be punished?
- Do you think that convicted spammers should be allowed to work in computing field?
- How vulnerable you are to spam?
- What is the likelihood that you will be spammed?

Most of the questions were given a choice “Maybe” which may be considered as “Do not know”, “Not sure of the context”, “Not sure of the term”, “Refused to answer” or “Not sure of the answer”.

B. Sampling Method

Data from this survey were collected from 17 February 2012 till April 2012. Participants were recruited through link advertising. Personal invitations to respond to the survey are given through personal email lists. A personal URL linked to the survey was embedded in the invitation email message, also ask the participants to distribute the link to their contacts. Recruitment was also done through link advertising on Facebook using several personal accounts. In order to achieve more responses and reach a broader response age group, the instrument was distributed and publicized through invitation linked to the survey in a few community groups Facebook wall. Consequently, we were not able to determine the response rate.

C. Respondent Demographics

Table 1 summarizes our self-reported respondent demographics. We consider gender, age group, education level as sensitive questions hence respondents are allowed to choose “Refused to specify” as their answer. There were a total of 368 respondents who completed the survey.

III. DESCRIPTIVE SURVEY FINDINGS

In this section, we present our findings on Internet users’ awareness, knowledge and perception of online spam.

A. Awareness of Online Spam

Participants were asked if they have ever heard of online spam; this is used as a measure of online spam awareness. This question is followed by another question that determines actual awareness. For this particular question, we consider the participants as having awareness if they have encountered any of the cases considered as online spam.

TABLE I. SUMMARY OF RESPONDENT DEMOGRAPHICS

Items	Categories	Percentage
Gender	Male	35.3
	Female	61.4
	Refused to specify	3.3
Age group	18-24	41.3
	25-34	35.3
	35-49	20.4
	50-64	1.4
	Refused to specify	1.6
Education level	Primary	0.3
	Secondary	0.8
	Certification	1.6
	Diploma/Advanced	4.9
	Diploma	43.2
	Undergraduate	45.4
	Postgraduate	3.8
Average hours spent on using Internet per day	Less than 1 hour	3.3
	1-5 hours	51.6
	6-9 hours	29.9
	More than 9 hours	15.2
Work/study relate to a technology field	Yes	74.2
	No	25.8
Activities engage when using the Internet	Searching for information	88.3
	Gaming	8.4
	Chatting & Social Networking	73.1
	Email	73.9
	Other	10.6

TABLE II. PERCENTAGE OF RESPONDENTS' AWARENESS ON ONLINE SPAM-RELATED ACTIVITIES

Items	Yes	No	Don't Know
Found pages that are only full of repeated keywords	47	39.7	13.3
Being redirected to an unrelated page from what was expected	87	9.2	3.8
Found pages with repetitive links	64.1	25	10.9
Received message considered as unwanted/suspicious/annoying on a web 2.0 application	78	10.9	11.1
Found pages with unrelated links	82.3	13.3	4.3
Received/seen suspicious link on a web 2.0 application	63	17.7	19.3
Found pages that are only advertising with very little content	77.4	15.8	6.8
Received unwanted postings on their social network account	67.4	27.7	4.9
Received unwanted friend requests on their social network account	77.4	18.8	3.8
Being tagged by unwanted parties on a Web 2.0 application	49.5	36.1	14.4

- Have you ever heard of online spam? (Yes/No)

91.6% (337 out of 368) of respondents stated that they have heard of online spam and 8.4% stated that they have never heard of online spam. This indicates that the respondents' perceived awareness is quite high.

- Have you had any of these experiences while browsing the Internet? (Yes/No/Don't Know)

In order to make it consistent with the subsequent questions in the knowledge section, we listed some

suspicious activities; if they have encountered any of the activities that are considered spam, then we consider them as having awareness. The summary of the results are represented in Table 2.

- Knowledge of Online Spam

Participants were asked to rate their overall knowledge of online spam; this question measures their perceived knowledge followed by questions to determine respondent's actual knowledge.

- My overall knowledge of online spam can best be described as (None/Poor/Fair/Good/Expert).

Only 1.6% rated themselves as having no knowledge at all about online spam. 26.4% of respondents considered themselves as having poor knowledge about online spam. More than half of total respondents (53%) considered themselves to have fair knowledge about online spam. Another 15.8% of the respondents rated their knowledge of online spam as good. There are only 3.3% of the respondents that considered themselves as experts.

- Which of these actions do you think is considered as online spam?

We listed several questions contain suspicious activities. We asked respondents to identify whether they considered these activities as online spam or not. Again, we allowed respondents to choose "don't know" as their answer. Table 3 represents the summary of respondent's percentage of answering the questions correctly, incorrectly or they chose "don't know".

TABLE III. PERCENTAGE OF RESPONDENTS' KNOWLEDGE ON ONLINE SPAM-RELATED ACTIVITIES

Items	Correct	Incorrect	Don't Know
Found pages that are only full of repeated keywords	55.2	24.7	20.1
Being redirected to an unrelated page from what was expected	79.6	13.9	6.5
Found pages with repetitive links	65.8	19.3	14.9
Received message considered as unwanted/suspicious/annoying on a web 2.0 application	81.8	9.2	9.0
Being hacked/hijacked is an action of online spam	36.1	57.3	6.5
Found pages with unrelated links	76.6	17.4	6
Attacked by virus	26.9	67.7	5.4
Received/seen suspicious link on a web 2.0 application	80.2	7.9	12
Found pages that are only advertising with very little content	60.1	26.6	13.3
Being asked passwords and/or credit card details by a trustworthy-looking entity in the Internet communication	23.4	68.8	7.9
Received unwanted postings on their social network account	84.8	10.9	4.3
Received unwanted friend requests on their social network account	72.8	20.4	6.8
Being tagged by unwanted parties on a Web 2.0 application	68.5	17.4	14.1

TABLE IV. PERCENTAGE OF RESPONDENTS' KNOWLEDGE ON ONLINE SPAM.

Items	Correct	Incorrect	Don't Know
Online spam can be used as a part of phishing attack	77.2	2.7	20.1
Online spam can be used to disseminate malware	72	7.2	20.9
Online spam can be used to promote affiliate websites	67.9	13	19
There is no difference between online spam and email spam	36.4	28.8	34.8
All online spam can be detected and treated using existing anti-spam techniques that are designed for email spam	25	35.6	39.4
Online spam has lower viewer impact than email spam" correctly	29.9	25.5	44.6
Auto registration software can be used to register spam accounts	45.4	5.2	49.5
Online spam can be found on legitimate websites	56.5	13.3	30.2
Online spam can be used to provide false information to users	77.7	5.2	17.1
Online spam can lead to other crimes such as fraud	81.5	3	15.5

- Assess the statement below and please choose the appropriate response for each item. (True/Not True/Don't Know)

For this question, we listed several statements regarding online spam and asked the respondents to decide whether the statement were true or false. We provided the "don't know" option. Most of the questions were basic and easy but we also purposely included two technical and hard questions. Percentages of respondents that have given correct, incorrect and don't know answers are summarized in Table 4.

B. Perception of Online Spam

In this section, we present the descriptive analysis for questions related to perception of online spam. The questions include the respondents' perception about spamming activity, their attitudes towards confessed and convicted spammers, their perception on spammers' motives behind spamming, and their perceived vulnerability towards online spam attacks. The results for some perception questions with yes/no/maybe answer are represented in Table 5. The rest of the results for perceptions of online spam are presented in the later paragraph.

TABLE V. PERCENTAGE OF RESPONDENTS' PERCEPTION ON ONLINE SPAM.

Items	Yes	No	Maybe
Do you think that online spam is a problem?	77.2	2.4	20.4
Do you think that spamming is acceptable?	72.8	6.3	20.9
Do you think that confessed spammers should be punished?	57.1	4.9	37.8
Do you think that convicted spammers should be allowed to work in computing field?	24.5	35.5	39.9

- How vulnerable are you to spam?

12.2% of respondents stated that they are very vulnerable to spam. 41.3% said that they are somewhat vulnerable, 21.2% said that they are indifferent, 22.6% said that they are not very vulnerable and another 2.7% stated that they are not at all vulnerable.

- What is the likelihood that you will be spammed?

11.4% of respondents corresponds that it is very unlikely for them to be spammed. 21.7% stated that it is unlikely for them to be spammed. 32.9% stated that it is undecided, while 27.4% said that it is likely that they will be spammed. Only 6.5% stated that they are very likely to be spammed.

IV. DISCUSSION AND CONCLUSION

Perceived awareness seems to be quite high; however, this does not translate to actual awareness. It can be seen that certain spamming activities are recognized by most of the respondents, as they either answered "yes" or "no"; the percentage of respondents who answered "don't know" is quite low. However, there are also spamming activities that are not recognized by most of the respondents, causing them to choose "don't know" as their answer. We found that the spamming techniques that fall under this category are mostly those used in social networking sites that use links and tags to embed spam contents.

Based on the questions on actual awareness, it is evident that most of the respondents are aware, can identify and have experiences with basic spamming techniques used by spammers such as "redirect to an unrelated page from what is expected" and "found pages with unrelated links". The percentages of respondents choosing "don't know" are low for both of these questions (3.8% and 4.3% respectively). However, there are a few spamming techniques such as "found pages that are only full of repeated keywords", "being tagged by unwanted parties" and "received/seen suspicious link on a Web 2.0 application" that seem to be unfamiliar to the respondents. It is possible that the respondents have never encountered them, but it is also likely that, even if they encountered them, they do not know that it is considered as online spam. The percentages of respondents that have answered "don't know" for these questions are relatively high (13.3%, 14.4% and 19.3% respectively), compared to the other activities.

Most respondents perceived themselves to have a fair amount of knowledge about online spam. Not many categorized themselves as having no knowledge about or being an expert in online spam. The questions on actual knowledge reveal some of the non-spam activities are mistakenly viewed as online spam. These include "account being hacked/hijacked", "attacked by virus" and "being asked username, passwords and/or credit card details by a trustworthy-looking entity in Internet communication".

REFERENCES

Aside from the percentages of wrong answers, the percentages of respondents choosing “don’t know” as answers are also important. As expected, the questions about the technical side of online spam and the ones that we considered difficult have the highest percentages of respondents choosing “don’t know”. These questions are “online spam has lower viewer impact than email spam” and “auto registration software can be used to register spam accounts” (44.6% and 49.5% respectively).

Considering those who answered the questions incorrectly and those who chose “don’t know” as not knowledgeable, it can be seen that there are some facts which are not well-understood by the general public. 75 % of the respondents had a misconception that all online spam can be detected and treated using existing anti-spam techniques that are designed for email spam. 63.6% of the respondents do not know that there is a difference between online spam and email spam. Other facts about online spam were known by at least more than 50% of respondents.

Most respondents agreed that online spam is a problem. The perceived seriousness of online spam is reflected by the fact that majority of them think that spamming is unacceptable. However, it is also astounding that 6.3% of the respondents that think spamming is acceptable. The next question also shows the perceived seriousness of online spam. More than half of the respondents thinking that confessed spammers should be punished. Nevertheless, 24.5% of the respondents think that convicted spammers should be allowed to work in computing field.

We use the questions on perceived vulnerability to divide the respondents into those who agreed that they are vulnerable to spam and those who think that they are not vulnerable to spam. More than half of respondents (53.5%) think that they are vulnerable to spam. However, there are 25.3% who think that they are not vulnerable to spam. The next question shows a similar pattern, with 33.1% of the respondents who think that there is a low likelihood for them to be spammed, and 33.9% who think that there is a high likelihood for them to be spammed. There could be two reasons for low perceived vulnerability. They may think that the applications provide sufficient protection against online spam. They may also think that they have enough knowledge to handle online spam.

As a conclusion, even though most respondents are aware of online spam, it is also apparent that respondents have inadequate basic knowledge about online spam. Although most respondents perceive online spam as a serious problem, their attitude towards punishing spammers severely are unclear.

- [1] T. Oreilly. (2005, 3 October). *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1008839
- [2] P. Hayati, V. Potdar, A. Talevski, N. Firoozeh, S. Sarenche, and E. A. Yeganeh, "Definition of spam 2.0: New spamming boom," presented at the IEEE International Conference on Digital Ecosystems and Technologies (DEST 2010), Dubai, UAE, 2010.
- [3] P. Hayati and V. Potdar, "Toward Spam 2.0: An evaluation of Web 2.0 anti-spam methods," in *7th IEEE International Conference on Industrial Informatics* Cardiff, Wales, UK, 2009, pp. 875-880.
- [4] V. Potdar, F. Ridzuan, P. Hayati, A. Talevski, E. A. Yeganeh, N. Firuzeh, and S. Sarencheh, "Spam 2.0: The Problem Ahead," in *Computational Science and Its Applications – ICCSA 2010*. vol. 6017, D. Taniar, O. Gervasi, B. Murgante, E. Pardede, and B. O. Apduhan, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 400-411.
- [5] F. Ridzuan, V. Potdar, and J. Singh, "Storage cost of spam 2.0 in a web discussion forum," presented at the Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference, Perth, Australia, 2011.
- [6] A. Stander, A. Dunnet, and J. Rizzo, "A Survey of Computer Crime and Security in South Africa," in *Information Security South Africa (ISSA)*, University of Johannesburg, South Africa, 2009, pp. 217-228.
- [7] K. S. Quinn, "2006 New Zealand Computer Crime and Security Survey," New Zealand2006.
- [8] P. S. Dowland, S. M. Furnell, H. M. Illingworth, and P. L. Reynolds, "Computer crime and abuse: A survey of public attitudes and awareness," *Computers & security*, vol. 18, p. 715, 1999.
- [9] A. I. Al-Alawi and M. F. Abdelgadir, "An Empirical Study of Attitudes and Opinions of Computer Crimes: A Comparative Study between U.K. and the Kingdom of Bahrain," *Computer Science*, vol. 2, pp. 223-235, 2006.
- [10] R. Weible and J. Wallace, "Cyber Research: The Impact of the Internet on Data Collection," *Marketing Research*, vol. 10, pp. 19-25, 1998.
- [11] M. Nowack, "The impact of the Internet on statistical organisations," *Statistical Journal of the UN Economic Commission for Europe*, vol. 14, p. 345, 1997.
- [12] J. M. Stanton, "An Empirical Assessment of Data Collection Using the Internet," *Personal Psychology*, vol. 51, pp. 709-725, 1998.

Spam 2.0

Vidyasagar Potdar
Anti-Spam Research Lab
School of Information Systems
Curtin University
Perth, Western Australia

v.potdar@cbs.curtin.edu.au

Farida Ridzuan
Anti-Spam Research Lab
School of Information Systems
Curtin University,
Perth, Western Australia

farida.mohdridzuan@postgrad
@curtin.edu.au

Jaipal Singh
Dept of Electrical & Computer
Engineering
Curtin University
Perth, Western Australia

j.singh@cbs.curtin.edu.au

ABSTRACT

In this paper, we provide a high level overview of Spam 2.0, how it works, its impacts and its categorizations (which are annoying, tricky, deceiving and evil). We also describe the existing approaches taken to combat Spam 2.0, including the detection approach, the prevention approach, and the early detection approach. Three techniques based on the detection approach presented in this paper include: content based, metadata based and user flagging based. We also explore several open issues/problems in this area. These include problems regarding tools and technologies, awareness and responsibility, and spam and spammers. Issues discussed regarding awareness and responsibility are users' lack of awareness, governments' inaction in tackling Spam 2.0, companies' apathy in combating it, lack of collaboration between countries, and unclear accountabilities in this regard. The paper also identifies future trends for both anti-spammers and spammers. Anti-spammers will likely focus their efforts more on behaviour based techniques and produce more language independent tools. Implementation of dynamic forms and forcing every user to actually go through the registration form will be good ways to control spam. From a monetary perspective, estimating intangible costs associated with Spam 2.0 will help raise the awareness of public users regarding spamming. On the other hand, the spammers will predictably continue to find methods to decrease the filters' efficiency by imitating real users' behaviours and finding other spamming opportunities.

Categories and Subject Descriptors

K.4.1 [Public Policy Issues]: Abuse and Crime Involving Computers; H.3.5 [Online Information Services]: Web-based services

General Terms

Security

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CUBE 2012, September 3–5, 2012, Pune, Maharashtra, India.
Copyright 2012 ACM 978-1-4503-1185-4/12/09...\$10.00.

Keywords

Spam 2.0, Web 2.0, spam

1. INTRODUCTION

Web 2.0 applications are the latest domain being manipulated by spammers, although spams can also be found in other domains such as in email, text messaging (SMS) and Internet telephony. Web 2.0 allows users to use the web as a platform for information sharing. This means that the developers trust the users as active content contributors instead of previously having one way interaction [1, 38-40]. This freedom given to users to provide information, however, has also given the spammers a way to manipulate this opportunity.

Some examples of Web 2.0 applications are blogging, social bookmarking, social networking, tagging, audio and video sharing, etc. These applications have made it much easier for Internet users to own personal sites and create their own communities. The need for programming skills has been discarded since there are many user friendly packages available on the net.

According to a recent research, the Americans spend almost 25% of their time on social networking sites and blogs [2]. With this amount of time spent on social networks and blogs, it is quite possible for these users to encounter spam. Without sufficient knowledge, they become prime targets for spammers.

This paper begins with a brief outline of Spam 2.0 and description of how it operates. The rest of the paper is organized as follows. Section 2 discusses the existing anti-spam approaches. Section 3 describes the problems/open issues that arise in Spam 2.0. In Section 4, the future trends of Spam 2.0 for both spammers and anti-spammers have been discussed, while Section 5 presents the conclusions of this paper.

2. SPAM 2.0

In this section, we will provide a high level overview of Spam 2.0. We will define Spam 2.0 and demonstrate how serious this problem is. We will then describe how it runs, its impacts and its categorization.

2.1 Spam 2.0 Definition

With over 2 billion internet users in the world today [3], thousands of visits are made to countless websites every day.

Many of these websites are vulnerable to a new threat called Spam 2.0. Spam 2.0 is defined as “propagation of unsolicited, anonymous, mass content to infiltrate legitimate Web 2.0 applications” [4-6].

In 2007, it was estimated that 75% of Google’s blogspot blogs were spam [7]. Akismet detected approximately 40 million spam comments per day from their clients’ sites [8]. Mollom claimed that their services identified 84% of total messages from 49,691 active websites as spam [9]. Each day, more and more Web 2.0 applications are reported to be attacked by spammers [10-11].

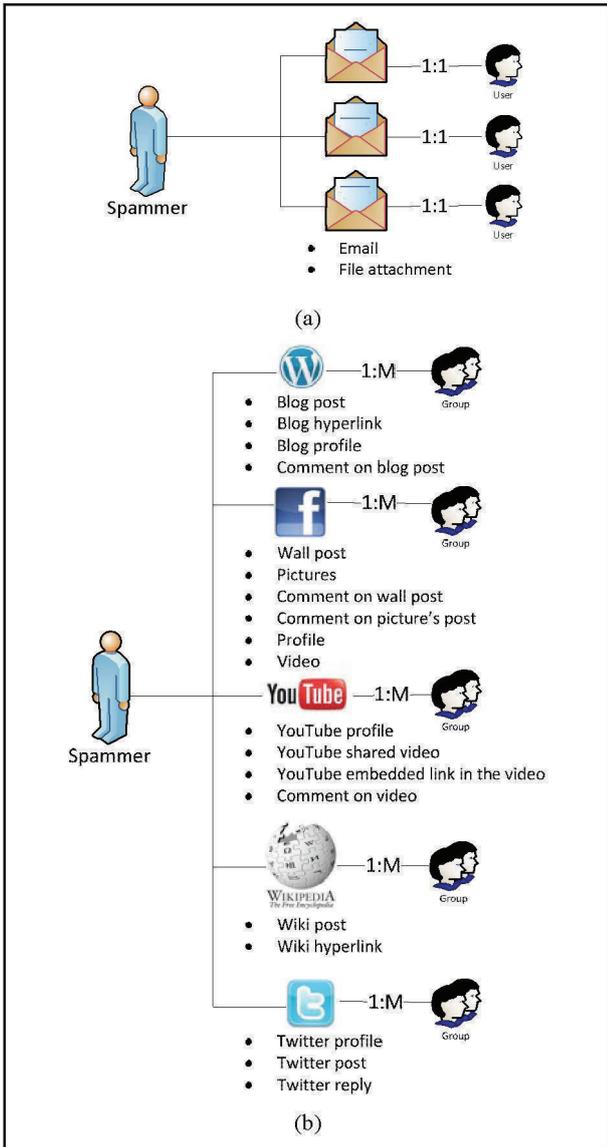


Figure 1: (a) 1:1 relationship; (b) 1:M relationship.

2.2 How Spam 2.0 Operates

In Web 2.0 applications, spammers (either automated bots or human spammers) attempt to operate as real users posting real comments/posts. They add specific information and usually

embed a link or attachment to advertise their page. This information may contain texts, images, hyperlinks, sounds, videos, file attachments etc. Such information (spam) neither increases the quality nor the value of a page [12]. The underlying motives of spamming in Web 2.0 are similar to email spamming, which is to generate revenues and increase traffic to the spammers’ websites [13]. Nevertheless, Spam 2.0 is more devastating than email spam or any other form of spam.



Figure 2: Example of annoying spam.



Figure 3: Example of tricky spam.



Figure 4: Example of deceiving spam.

2.3 Impacts of Spam 2.0

Spam 2.0 not only annoys the users, it has a higher viewer impact, compared to email spam. To briefly describe how Spam 2.0 has a higher viewer impact than email spam, we first explain the spam unit.

The spam unit is defined as an “attribute that can be manipulated by spammers to embed their spam content” [14]. In the email domain, spam can be embedded in the email content. Therefore, the spam unit in the email domain is the email itself and its attachment.

There are also a few spam units involved explicitly in different Web 2.0 applications. Examples of spam units in Web 2.0 applications are the user profile, post, poll and personal message.

Theoretically, email spam has a 1:1 relationship. This means that an email is read just by one user, but Web 2.0 applications provide sharing services and are community based. Spam 2.0 has 1:M relationship whereby 1 spam unit in any Web 2.0 application can be read by many users, thus having a higher viewer impact. Figure 1 depicts these relationships for some of the most popular Web 2.0 applications. For instance, profiles, wall posts, comments on wall posts, photos, comments on photos and personal messages are all spam units in Facebook. Features built-into the Web 2.0 applications allow the spammers to manipulate them to reach out to wider groups of viewers/readers, faster and more easily.

2.4 Spam 2.0 Categorization

Spam 2.0 advertisements can be categorized into four types, which are annoying, tricky, deceiving and evil [8]. The first type is a typical annoying spam containing simple, repeated and obvious texts that already exist in the spam keyword database, along with obvious links. This type of spam can be easily detected. Figure 2 shows an example of a Facebook spam posted from a linked email. This spam originated from an email and, since Facebook allows users to reply to comments from an email, this causes the spam to be posted on the user's wall.

Spams can also be tricky. This type of spam is not very easy to detect and may contain hidden links, texts and hyperlinks. Figure 3 shows an example of a similar window from a chat session. It comes from a contact and the content is hidden. Figure 3(a) shows the windows of a chat session that look like having no content at all. But after being highlighted (refer to Figure 3(b)), this chat window shows texts which are hyperlinked to a webpage. It is possible that the hidden link is targeted at those who click randomly on a window.

The third category of spam is deceiving. This type of spam is crafted with bad intentions, such as to do scam or fraud, or launch phishing attacks. Figure 4 shows an example of a deceiving spam which is an application in Facebook promoting to switch to a better interface. Once clicked, the users will be asked to verify, giving their Facebook account access to a third party which, then, can use it for scamming purposes [15]. This particular fake application promotes an online survey which would profit a third party [15]. More examples of deceiving spams can be seen in [15-17].

The fourth type of spam is evil spam. The contents of this type of spam are crafted to create a way to spread viruses, malware, worms, Trojans and other such tools that raise security threats in the community. Figure 5(a) shows an example of a spam linked between an email and Facebook. The email recipient would think that it is a valid email from the Facebook team and would not hesitate to click on the link. However, once clicked, the user will be redirected to another browser and the user's PC will get infected by malware. This works in a similar way to the example in Figure 5(b). The link is embedded in a Youtube video and the users are asked to install security software which provides an open door to malware. More examples are as given in [17-19].

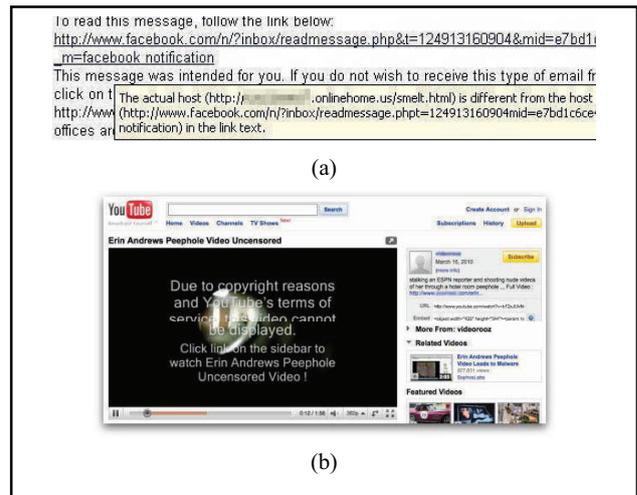


Figure 5: Example of evil spam.

Overall, Spam 2.0 not only wastes resources, such as storage and bandwidth [14], it can also initiate a greater online security threat such as scam, hacking and malware attack [15, 17-20]. Spam 2.0 also affects users with intangible costs, such as causing them annoyance, breaching their trust, and tarnishing the website's reputation.

3. EXISTING APPROACHES

In this section, we will describe three approaches to manage Spam 2.0, which are detection, prevention and early detection [13]. We will also explain the advantages and disadvantages of these approaches.

3.1 Detection Approach

To the best of our knowledge, there are three techniques using the detection approach, which are content based, metadata based and user flagging based.

3.1.1 Content based

The content based technique has been used widely in the early research on spam detection. This technique relies heavily on the content. Techniques used to detect Spam 2.0 are primarily taken from the email spam domain. As a result, their efficiency works similarly as in the email domain. In this technique, a filter is created, based on the analysis of several features, such as keyword search and keyword frequency. Training classifiers are needed to differentiate between spam and non-spam content. Often, an n-grams language model is used in this technique.

However, there are several drawbacks in this technique, including exhaustive computation, the need for regular filter updates, language dependency and the requirement of either a labelled or unlabelled training data set. In addition, it does not provide a real time result. This technique also allows spammers to frequently fool the filter. In time, if such fooling continues, it will slowly corrupt the filter's performance. This technique has been implemented in blogs [21], forums [22], tags [23] and Twitter [24].

3.1.2 Metadata based

This technique works in a similar way as the content based techniques, except that the analysis of spam patterns is done

throughout the process of mining selected features from the metadata. The metadata based techniques, such as link detection, generally work faster than content based detection and are language independent. However, this technique requires training and does not provide real time results. In addition, it does not work effectively when used single-handedly. There has been, considerable work focusing on this technique in forums [22], tags [23] and Twitter [25].

3.1.3 User flagging based

The user flagging based techniques work by receiving input from the users in a Web 2.0 application. For instance, forum users can report spam if it is found in a thread; or Facebook users can mark a post as spam if they detect it on their wall. This technique allows the system to receive assistance from active users for spam detection. Nevertheless, the disadvantages of this technique are that it requires the end user's cooperation for flagging possible spam, and this function can also be manipulated by the spammers. In most cases, knowledgeable users rather take a passive approach by not reporting or flagging spam, even though they are aware of the spam content. In addition, a human workforce is needed to identify, check and delete the flagged spam content manually.

Overall, one of major drawbacks of the detection approach is that it allows the miscreants to place spam on the Web 2.0 applications in the first place. Someone has to spend resources, including time and money to detect and remove it. Furthermore, detecting spam entirely depends on the efficiency of the technique. Nevertheless, the detection approach can handle all types of spammers, either bots or humans.

3.2 Prevention Approach

This section focuses on the second approach, which is the prevention approach. One of the most popular techniques that have adopted the prevention approach is CAPTCHA. As the name suggests, this approach prevents spammers from placing spam on Web 2.0 applications and prohibits spammers from entering the system at all. CAPTCHA is now implemented in most Web 2.0 applications [26], especially for registration. CAPTCHA is designed to differentiate between human and non-human users to protect Web 2.0 applications from automated bots, based on the response. To meet this goal, CAPTCHA texts or images have to be answered correctly, which can be done by humans but not by bots. Nevertheless, smarter bots have now even implemented OCR techniques to overcome this approach. Therefore, images used in CAPTCHA are created using text with more noises and curves, and are scattered, distorted and crossed with lines. Thus, these images rely heavily on human visibility and capability, thus providing a trade-off between human users and automated spammers. However, this approach is still unsuitable to control human spammers.

Moreover, CAPTCHA has shown to be defeated by a tool called X-Rumer. Therefore, in order to create a challenge which can easily be tackled by humans, and yet is hard enough for bots to crack, many methods have been developed to craft a better CAPTCHA. These methods include animated CAPTCHA, skill testing CAPTCHA (usually simple mathematical problems in the form of simple numbers and image-based), and audio CAPTCHA. Some examples of CAPTCHA can be seen in Figure 6 below.

CAPTCHA is now used by most popular websites during registration. Although it works well to prevent bots from registering Web 2.0 applications, it is sometimes also difficult for

human users to solve it. Some of the CAPTCHA text cannot be seen clearly and is eye straining. The users have to repeat it until they get it right. Some questions provided in CAPTCHA cannot be solved by all users because they are too hard, language dependent, or require some specific knowledge. Some of the hard-to-solve CAPTCHA examples are provided in Figure 7.

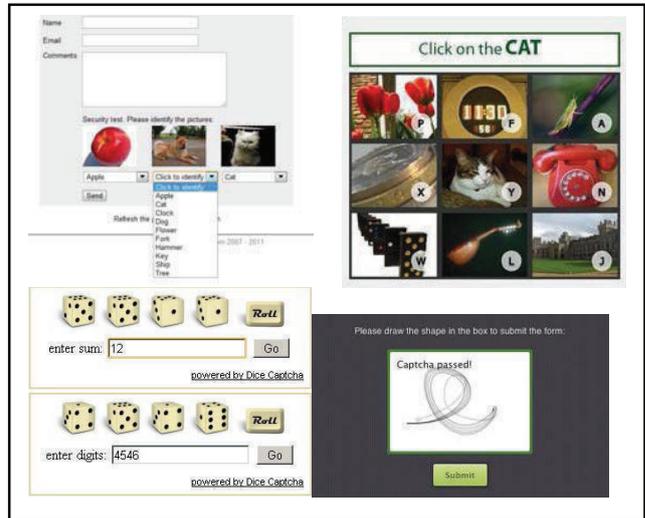


Figure 6: Examples of CAPTCHA.

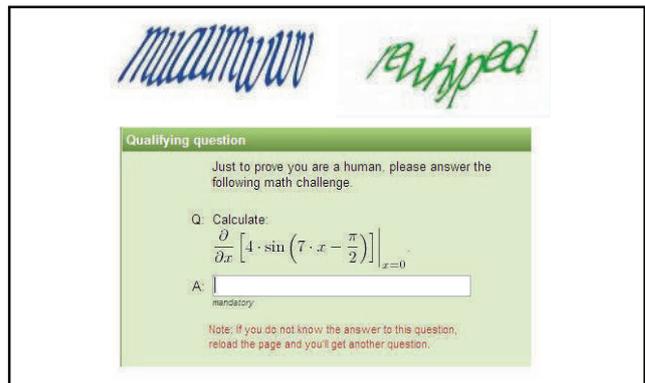


Figure 7: Hard-to-solve CAPTCHA

3.3 Early Detection Approach

The two previous approaches are active approaches whereby the user or the website administrator has to actively take steps for managing spam. The third approach is a detection approach in which the user's participation in detecting spam is discarded, so users are not required to solve CAPTCHA to access the system. This approach relies on behaviour based techniques to detect spammers.

This technique is able to detect spam before it enters Web 2.0 applications, and overcomes the disadvantages of other approaches. It can also detect patterns defined earlier based on several features, and classify them. It is language independent and provides a more real time solution than the detection and prevention approaches.

However, one of the limitations with this technique is that it is unclear how well it will be able to adapt to new patterns. This

technique has been implemented to detect Spam 2.0 in Twitter [24], web spam [27] and Youtube [28]. There have also been successful efforts to implement it to detect spam profiles in social networking websites [29].

4. PROBLEMS/OPEN ISSUES

All of the studies in the previous section have presented existing approaches for combating spam. However, there are several problems/open issues in this area. This section is divided into three sub-sections concerning problems/open issues related to tools and technologies, awareness and responsibility, and spam and spammers.

4.1 Tools and Technologies

Tools and technologies grow rapidly to fulfil human needs. In order to solve the spam problem, researchers are working intensively to develop anti-spam tools and technologies. However, the existing technologies are lacking in some important features. Despite having an automated filter, they still require a human workforce to distinguish, decide and manually eliminate the filtered spam. Anti-spam technologies have not successfully developed a full real time solution for controlling spam. From the users' side, there is a need for comments/posts to be seen immediately after posting. However, such real time posting is beneficial for spammers too, as they can reach readers more easily and quickly. Also, the existing tools and technologies provide a trade-off between the users and spammers. For instance, simple CAPTCHA can be solved easily by both users and bots, while a more complex CAPTCHA might be too hard to be solved by users, but still could be cracked by bots.

Web 2.0 applications are created on a platform to enhance the users' interactions, but still keeping them simple and easy to be used. Unfortunately, this opportunity has also been taken by spammers, in order to manipulate this platform to run their spamming activities. Any information provided on the web can be posted by both real users and spammers. With the growth of spam in this platform, more problems have arisen for the users, such as justifying the quality of information provided on the web, and trusting the information and the website. Due to these, many websites' reputation is being questioned. Making a transaction on the web seems a high risk task, especially when it involves financial and privacy issues. There have been numerous security issues in dealing with spam. Spammers go beyond just posting an annoying post or advertising their links to boost up their websites in search rankings. They are also drawn in to get involved with security threats, such as propagation of malware and viruses, hijacking accounts, identity thefts, phishing attacks, frauds, etc. Hence, there is a need for a framework for spam resistant Web 2.0 applications.

4.2 Awareness and Responsibility

We will discuss the awareness and responsibility issues from different perspectives, including those of the users, governments, companies and countries.

4.2.1 Users' lack of awareness

The function of developing anti-spam tools and technologies is to assist people in dealing with spam. In the end, it is the people's responsibility to eliminate spam. Spam can manually be eliminated by the web administrator. It can also be eliminated if the users are involved in identifying spam, and marking and reporting it as spam. Hence, it is very important that users play an

active role in order to combat spam. Furthermore, the spammers' business models rely heavily on the users' clicks in order to generate revenues. If there are no users clicking on the link, their spamming will not work at all. Nevertheless, the issue is that not all users are aware of this problem. Users of Web 2.0 applications are increasing each day. This is also opening up many loop-holes for spammers as most users use the applications without sufficient knowledge.

4.2.2 Lack of government action

We cannot solely rely on the law enforcers to catch these spammers. In fact, in Australia, there is no law against web spammers. On the other hand, there is a monetary implication of spam for the National Broadband. Although it is unclear how much the government spends on filtering internet traffic to combat spam, there are definitely some costs involved which are lost due to spamming activities.

4.2.3 Unclear accountability

When Spam 2.0 is found on the web, it is also unclear whose responsibility it is to take action against them. The Web 2.0 platform is built to share information collectively. As a result, it is very hard to determine who is responsible for what is posted on the net. Should it be the website owner who is running the website, or should it be the users' responsibility? Catching spammers is a nuisance. If the responsibility is placed on the content provider, there would be a need to come up with better ways to track them, as tracking just the IP address is currently not enough to determine the real user.

4.2.4 Company's lack of efforts

On the social side, the companies' inputs and efforts in fighting spam are also questionable. The companies may not think that they are involved with this issue or they hold any responsibility regarding it. They might also be thinking that there are no costs imposed on them by spams.

4.2.5 Lack of collaboration between countries

This issue becomes larger and more important if viewed from a wider perspective. The level of collaboration among countries to reach zero spam is very low. Several developed and developing countries have passed spam laws. However, spammers can originate in country A, but use country B's resources, and obtain revenues from a different bank in country C. Therefore, a collaborative effort is required, if we want to deal with spammers.

4.3 Spam & Spammers

The root of the problem is spam itself. Spams are created in a way that they get through the filter, which is usually the first inline layer of security in Web 2.0 applications. The spammers want the readers who read spam posts or comments to be unable to distinguish whether it is spam or not and believe in its content. So, it all comes back to the basic problem that spam for one may not be spam for others. There is a grey area between spam and non-spam. Since there is no clear border line between these two, it creates the possibility for the spam content to be flagged as genuine content.

Furthermore, spam that is sent to Web 2.0 applications also has its own innovation. It does not only contain obvious images, texts and links, it may also contain personal touches, hidden links, and be embedded in various types of files and formats. The existence

of URL shorteners also gives the advantage to spammers, since the original link is not visible until it is clicked [30]. Spam content is crafted to make the users believe that it is a very good deal, and increases their interest in clicking them. Changing the nature of spam makes it harder for the anti-spammer teams to keep updating their algorithms and databases, and to further produce new ways of combating spam.

Further, spam created by manual spammers is also difficult to be detected. Even though manual spammers cannot spam to the extent bots can, they can target the high traffic and popular websites, like FB/CNET, etc. Since they are humans, their spams are crafted in a smarter way with a personal touch. As a result, their success rate is higher even though it is not certain how much they profit from this type of spamming. These are some of the most important issues that need to be addressed in tackling spamming.

5. FUTURE TRENDS

Section 3 has covered the problems and open issues that need to be addressed with regard to Spam 2.0. In this section, we are going to predict future trends for both spamming and anti-spamming.

5.1 Anti-Spammers

Anti-spammers have developed tools and techniques, such as those discussed in Section 2. Nevertheless, it is an ongoing effort to keep on combating Spam 2.0. It is possible to control spam by focusing on developing more behaviour based techniques, creating more language independent tools, using dynamic forms, and forcing the user to go through a registration form and take into account other intangible costs.

5.1.1 Behaviour based techniques

A future trend in developing anti-spam technologies is not only to depend on content filters but also to switch to behaviour-based approaches. Currently, there are only a few researchers focusing in this line in order to develop cutting edge solution [31-34]. Behaviour based anti-spam tools are better because they are based on the behaviour of the spammers and do not depend on the spam content. Spam content can be created neatly by the spammers to follow the real users' actions. The smarter the spam content, the harder for the content based anti-spam tools to detect spam. Therefore, using behaviour based techniques is a better solution. It is just a matter of searching for new ways to track the spammers' behaviour.

5.1.2 Language independent

The research in developing tools and techniques of language independent spam is also moving forward. Since there are thousands of languages and lettering/scripting, and most of the anti-spam tools are in English, spam contents are also crafted in different languages. Therefore, it is now a trend for anti-spam commercial companies to provide their anti-spam solutions in different languages. This customization will increase the marketability of their products and ensure that they can cater to a wider group of users. As the awareness of spamming issues is growing in several countries, there is also a prospect that local companies will develop anti-spam solutions, focusing on local markets. Therefore, there will also be specific anti-spam solutions that are language specialized, and only cater to certain countries and certain lettering/scripting.

5.1.3 Dynamic form

In order to prevent spammers from continuing to fill in the static sign up form so easily, it is recommended that the sign up HTML forms be created dynamically, which will allow customization in terms of sequence of data that needs to be filled up. However, from the perspective of a developer, the spammers will just need to be able to read/detect the programmers' fieldname in order to solve a dynamic form.

5.1.4 Registration form

It has also been found that when spammers use auto-submitter tools, they do not even have to go through the registration page to make a successful registration [35]. Therefore, in order to prevent spammers from registering and acting like real users, and posting spam posts in a Web 2.0 application, there is a need to force them to go through registration forms.

5.1.5 Intangible costs

With the cloud services increasing in popularity, the cost of storage for keeping spam will decrease significantly. Nevertheless, the security threat caused by spam is still a matter of concern. Furthermore, the intangible costs such as the feelings of annoyance and emotional devastation faced by users when dealing with spam is a larger issue. Hence, the motivation for combating spam will be greater once the intangible costs are taken into account.

5.2 Spammers

Having discussed the anti-spammers' side in Section 4.1, we will now predict the spammers' trends in the future. The innovative nature of the spammers is also not static. It is like an arms race between spammers and anti-spammers. Predictably, spammers will continue to decrease the filter efficiency, operate like real users, and exploit the factors that give them greater opportunity to continue their spamming activity.

5.2.1 Decrease filter efficiency

Similar to what has happened in the email domain, spammers will predictably continue messing with the filtering systems. Research shows that they are following the acts of real users to mess with the filters, a technique called vote gaming attack [36]. If this is done using the millions of pieces of content posted by real users, it is possible for the system to identify a genuine content as spam and vice versa.

5.2.2 Operate like real user

Spammers will also continue to try behaving like real users and operate by posting spam content that looks legitimate. In Web 2.0 applications, they want to be seen as real people writing real comments so that their comments are approved and attract more readers. In order to do this, they need to craft smarter posts/comments. Using an auto-submitter called X-Rumer, a post/comment is created using several key words which make it seem logical and reasonable in accordance with the original topics or posts, thus avoiding the spam filter. They are also likely to continue using current popular news and events.

Creating new profiles which look like real users takes time, but spammers may copy or download real profiles from any network and post them on another one. However, they also have to bypass CAPTCHA to successfully make a registration. Therefore, they may take an easier way out by hijacking the real users' online

accounts. They would not change the passwords of the real users; their purpose will be to use the accounts to reach other profiles linked to this account, as a friend's post/comments are more trustworthy than those of an unknown person. There is also a proof of collaboration between virus/Trojan authors with other cyber thieves [37]. Even though it is unclear how big the market is, the effectiveness of this collaboration is giving an advantage to the spammers.

5.2.3 Spammers' opportunity

The openness of the Internet has allowed personal items to be shared more than ever before. These include audios, photos and videos. At this point, any item that can be shared online can also be used by spammers to embed their spam content in it. All emerging tools and technologies that can be used to hide their spam content, or even hide themselves from being detected as spammers, are being fully utilized. Hence, it is just a matter of time before the spammers come up with new items of spamming or new ways to spam.

6. CONCLUSION

This paper gave an account of Spam 2.0. It demonstrated that Spam 2.0 is the latest trend of spamming that can pose a serious threat to online communities. Even though researchers have developed several solutions based on detection, prevention and early detection approaches, it is just a matter of time before the spammers defeat them. The paper also discussed future possibilities for both spamming and anti-spamming sides. To conclude, it indicated that there is considerable room for improvement in order to eliminate Spam 2.0.

7. REFERENCES

- [1] O'Reilly, T., 2005. What is Web 2.0. <http://oreilly.com/web2/archive/what-is-web-20.html>
- [2] Nielsenwire, 2010. What Americans Do Online: Social Media And Games Dominate Activity. http://blog.nielsen.com/nielsenwire/online_mobile/what-americans-do-online-social-media-and-games-dominate-activity
- [3] Internet World Stats, 2011. World Internet Usage Statistics News and World Population Stats. <http://www.internetworldstats.com/stats.htm>
- [4] Hayati, P., Potdar, V., Talevski, A., Firoozeh, N., Sarenche, S., Yeganeh, E. 2010. Definition of spam 2.0: New spamming boom. In *Digital Ecosystem and Technologies (DEST)*, Dubai, UAE, 2010. IEEE Computer Society.
- [5] Hayati, P., Potdar, V. 2009. Toward Spam 2.0: An Evaluation of Web 2.0 Anti-Spam Methods. In *7th IEEE International Conference on Industrial Informatics*, Cardiff, Wales, 2009.
- [6] Hayati, P., Chai, K., Potdar, V., Talevski, A. 2009. HoneySpam 2.0: Profiling Web Spambot Behaviour. In *12th International Conference on Principles of Practise in Multi-Agent Systems*, Nagoya, Japan, 2009, pp. 335-344.
- [7] Roundtable, 2007. 75% of Google's Blogspot Blogs are Spam. <http://www.seroundtable.com/archives/012778.html>
- [8] Akismet, 2011. 25 Billion Pieces of Spam. <http://blog.akismet.com/>
- [9] Mollom, 2011. Scorecard | Mollom. <http://mollom.com/scorecard>
- [10] SC Magazine, 2009. Twitter, Facebook and LiveJournal Attacked. <http://www.scmagazine.com.au/News/152328,twitter-facebook-and-livejournal-attacked.aspx>
- [11] SC Magazine, 2008. Facebook user profiles hacked. <http://www.scmagazine.com.au/News/107015,facebook-user-profiles-hacked-wall-feature-relaying-spam.aspx>
- [12] Chai, K., Hayati, P., Potdar, V., Wu, C., Talevski, A. 2010. Assessing Post Usage for Measuring the Quality of Forum Posts. In *Proceedings of the 4th IEEE International Conference on Digital Ecosystems and Technologies (DEST)*.
- [13] Hayati, P., Potdar, V. 2008. Evaluation of spam detection and prevention frameworks for email and image spam: a state of art. In *Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services*, ACM, Linz.
- [14] Ridzuan, F., Potdar, V., Talevski, A., Smyth, W.F. 2010. Key Parameters in Identifying Cost of Spam 2.0. In *Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications*. IEEE Computer Society, 2010, 789-796.
- [15] Scam Sniper, 2011. Phishing Scam Alert: Comment Spam Leads to Facebook Phishing Scam. <http://scamsniper.blogspot.com.au/2011/06/phishing-scam-alert-commenting-spam.html>
- [16] Naked Security Sophos, 2012. The Pink Facebook rogue application and survey scam. <http://nakedsecurity.sophos.com/2012/02/27/pink-facebook-survey-scam/>
- [17] Webroot, 2010. Facebook spam Leads to Viagra Vendor, Drive-by Download. <http://blog.webroot.com/2010/05/28/facebook-spam-leads-to-viagra-vendor-drive-by-download/>
- [18] Naked Security Sophos, 2010. Erin Andrews Peephole Video maker jailed, as hackers take advantage. <http://nakedsecurity.sophos.com/2010/03/16/erin-andrews-peephole-video-maker-jailed-hackers-advantage/>
- [19] Sean, 2010. CPAlead Spam on YouTube. <http://www.f-secure.com/weblog/archives/00002019.html>
- [20] Hayati, P., Potdar, V. 2009. Spammer and Hacker, Two Old Friends. In *3rd IEEE International Conference on Digital Ecosystems and Technologies (IEEE-DEST 2009)* Istanbul, Turkey, 2009.
- [21] Thomason, A. 2007. Blog Spam: A Review. In *Conference on Email and Anti-Spam* (Mountain View, California, August 2-3, 2007).CEAS2007.
- [22] Shin, Y., Gupta, M., Myers, S. 2011. Prevalence and mitigation of forum spamming. In *the 30th IEEE International Conference on Computer Communications*. (Shanghai, China, April 12-14, 2011) IEEE INFOCOM 2011. IEEE Computer Society, Shanghai, China.
- [23] Markines, B., Cattuto, C., Menczer, F. 2009. Social Spam Detection. In *Fifth International Workshop on Adversarial Information Retrieval on the Web* (Madrid, Spain, April 21, 2009). AIRWeb'09. ACM.
- [24] Chu, Z., Gianvecchio, S., Haining, W., Jajodia, S. 2010. Who is Tweeting on Twitter: Human, Bot or Cyborg? In *Annual Computer Security Applications Conference* (Austin, Texas, USA, December 6-10, 2010). ACSAC'10. ACM.
- [25] Grier, C., Thomas, K., Paxson, V., Zhang, M. 2010. The Underground on 140 Characters or Less. In *17th ACM*

- Conference on Computer and Communications Security* (Chicago, Illinois, USA, October 4-8, 2010). CCS'10. ACM.
- [26] Egele, M., Bilge, L., Kirda, E., Kruegel, C. 2010. CAPTCHA Smuggling: Hijacking Web Browsing Sessions to Create CAPTCHA Farms. In the *25th Symposium on Applied Computing* (Sierre, Switzerland, March 22-26, 2010). ACS'10. ACM.
- [27] Liu, Y., Cen, R., Zhang, M., Ma, S., Ru, L. Identifying Web Spam With User Behavior Analysis. In *Fourth International Workshop on Adversarial Information Retrieval on the Web* (Beijing, China, April 22, 2008). AIRWeb'08. ACM.
- [28] Sureka, A. Mining User Comment Activity for Detecting Forum Spammers in Youtube. In *1st International Workshop on Usage Analysis and the Web of Data in the 20th International World Wide Web Conference* (Hyderabad, India, March 28, 2011). WWW2011.
- [29] Stringhini, G., Kruegel, C., Vigna, G. 2010. Detecting Spammers on Social Networks. In *Annual Computer Security Applications Conference* (Austin, Texas, USA, December 6-10, 2010). ACSAC'10. ACM.
- [30] Weiss, D. 2009. The Security Implications of URL Shortening Services. <http://unweary.com/2009/04/the-security-implications-of-url-shortening-services.html>
- [31] Hayati, P., Potdar, V., Chai, K., Talevski, A. 2010. "Web Spambot Detection Based on Web Navigation Behaviour. In *24th IEEE International Conference on Advanced Information Networking and Applications (AINA 2010)*, Perth, Western Australia, 2010.
- [32] Hayati, P., Chai, K., Talevski, A., Potdar, V. 2010. Behaviour-Based Web Spambot Detection by Utilising Action Time and Action Frequency. In *The 2010 International Conference on Computational Science and Applications (ICCSA 2010)*, Fukuoka, Japan, 2010.
- [33] Hayati, P., Potdar, V., Talevski, A., Chai, K. 2010. Web Spambot Characterising using Self Organising Maps. In *International Journal of Computer Systems Science and Engineering*, 2010.
- [34] Hayati, P., Potdar, V., Smyth, W. F., Talevski, A. 2010. Rule-Based Web Spambot Detection Using Action Strings. In *The Seventh Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference (CEAS 2010)*, Redmond, Washington, 2010.
- [35] Shin, Youngsang, Minaxi Gupta, and Steven Myers. "The Nuts and Bolts of a Forum Spam Automator." In the *LEET'11 Proceedings of the 4th USENIX conference on Large-scale exploits and emergent threats*, Berkeley, CA, USA, 2011.
- [36] Ramachandran, A., Dasgupta, A., Feamster, N., Weinberger, K. 2011. Spam or Ham? Characterizing and Detecting Fraudulent "Not Spam" Reports in Web Mail Systems. In the *8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference* (Perth, Western Australia, September 1-2, 2011). CEAS2011.
- [37] Brian Krebs, 2012. Zeus Trojan author in spam kingpins. <http://www.theage.com.au/it-pro/security-it/zeus-trojan-author-in-with-spam-kingpins-20120222-1tmqp.html>
- [38] Sukanta Sinha, Rana Dattagupta, Debajyoti Mukhopadhyay; Identify Web-page Content meaning using Knowledge based System for Dual Meaning Words; *International Journal of Engineering Research and Applications*; Vol.2, N0.4; July-August 2012; pp.877-880; ISSN 2248-9622.
- [39] Ruma Dutta, Anirban Kundu, Debajyoti Mukhopadhyay; Clustering based Web Page Prediction; *International Journal of Knowledge and Web Intelligence*; Inderscience Publishers; UK; Vol.2, No.4, 2011; pp.257-271; ISSN 1755-8255.
- [40] Debajyoti Mukhopadhyay, Debasis Giri, Sanasam Ranbir Singh; An Approach to Confidence Based Page Ranking for User Oriented Web Search; *ACM SIGMOD Record*, ACM Press, New York, USA; Vol.32, No.2, June 2003; pp.28-33; ISSN 0163-5808.

Storage Cost of Spam 2.0 in a Web Discussion Forum

Farida Ridzuan, Vidyasagar Potdar, Jaipal Singh
Anti-Spam Research Lab
Digital Ecosystems and Business Intelligence Institute
Curtin University, Perth, Western Australia
{farida.mohdrizuan@postgrad.,v.potdar@, j.singh@cbs.} curtin.edu.au

ABSTRACT

This paper presents an empirical research that identifies cost of Spam 2.0. This experiment is a part of ongoing research for identifying the cost of Spam 2.0 and focuses only on storage cost. The data is collected via a honeypot setup using a discussion forum for a period of 13 months. Forum provides a good place for the spammers to continue their spamming activities. Spamming give both direct and indirect cost towards forum owner and forum users. In this paper, we present a method to measure direct cost focusing only on storage cost. The main observation of the experiment is done towards 450,772 posts, 141 personal messages and 62,798 profiles. It uses 2.69 GB storage space. We first define our cost formula. We then set up a web based discussion forum and collect the information posted on the forum. This data is pre-processed to discover information that can be used in our formula. In order to identify the storage used for spam, we define related attributes based on maximum storage and impact factor features named as spam unit, and measure the storage taken by all these spam units. We evaluate the cost of storage based on three sources which are our real self-hosted server, commercial web hosting package and cloud hosting package. The experiment resulted that the storage cost for our research forum are AUD 23.66 based on self-hosted server, AUD133.90 for commercial web hosting, and AUD11.53 for cloud hosting. The highest storage cost for 10,000 spam posts, profiles and personal messages is AUD2.963, AUD0.068 and AUD0.056.

Categories and Subject Descriptors

C.4 [Performance of systems]: Measurement techniques, D.2.9 [Software Engineering]: Management---Cost estimation, K.4.2 [Computers and Society]: Social Issues---Abuse and crime involving computers

General Terms

Experimentation, Measurement.

Keywords

Cost, Spam 2.0, Storage Cost, Discussion Forum, Web 2.0.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CEAS '11, September 1–2, 2011, Perth, Western Australia, Australia.
Copyright 2011 ACM 978-1-4503-0788-8/11/09...\$10.00.

1. INTRODUCTION

Web 2.0 is a platform offering openness and freedom of information sharing. While it gives many benefits to the Internet users, this freedom have opened up/provide door to problems such as spamming activities and online harassment [1]. Spam 2.0 is defined as “propagation of unsolicited, anonymous, mass content to infiltrate legitimate web 2.0 applications” [2].

A discussion forum, similarly like other web 2.0 applications, is “a website where visitor can contribute content” [3-5]. It is one of the most popular types of web 2.0 application used by spammers for their spamming activities. A discussion forum’s popularity or how active a forum is greatly depends on the user, as they are the ones who generate the content. To fit the purpose of a discussion forum where users could simply interact with each other, a registered or a non-registered user (depends on the forum setting) does not have any limitations to fully utilize the functions available in a forum. This includes posting any unrelated content like images and hyperlinks in the forum.

Forum spamming is regularly associated with link spamming as they share the same motivations. Spammers want to acquire higher traffics for their websites [3,4]. Putting their targeted link in the forum is a way of advertising their links. One click from forum users is what they are expecting. More clicks to the targeted website could also help them to generate revenue. Concurrently, more visitors to the targeted websites will help to boost the websites rank in search engine [3,4] and associated revenue from online ads displayed on the pages.

Indirectly, a legitimate forum could be down ranked by the users if they found spam posts in it. Users tend not to trust the value of the information posted in that forum hence lowering that forum’s reputation [5,6]. Thus, it becomes the forum administrator’s work to clean their forum from spam post, which will adds more work to the forum administrator and requires time and human’s expertise to identify spam. Additionally, network resources are drained and wasted specifically storage used for keeping spam in the forum. Study on the effect of forum spamming has been covered in [4,7].

Despite of having all the harmful consequences of this situation, it is not easy to eliminate this problem in forum. The owner of the email could be blacklisted and the email could be reported as spam. While in the social networking websites, fake profile social networking users could be blacklisted, deleted and identified as spammers. Unfortunately, we cannot simply blacklist a legitimate forum with valuable information though it contains spam [3]. It is more likely the task of a forum user to report the spam to the forum administrator and the task of the forum administrator to remove the spam themselves. Eventually, it depends on the forum administrator to take the action.

Most of the research in this area is focusing on the method for detecting spam either content based or behavior based. From an economic view, a study has been done by [8] to identify whether Captcha is really beneficial to be used in order to protect the web applications. They believe that it fits the purpose to make it harder for the spammer's to attack a website but it tends to work less efficient in the future. Even though [9] has raised the question "who is harmed by spammers who generate fake content?" and despite knowing that there are some financial cost involved when dealing with forum spamming, no one has tried to identify the cost involved. Thus, the focus of this paper is to investigate and produce the real cost when dealing with Spam 2.0. This will be done by estimating cost of storage for a honeypot implemented as a forum.

Towards reaching our goal, we need to identify on how to estimate the cost of storage in our forum. We first define the concept of *spam unit* in a forum and *forum attributes* related to each spam unit in Section 2. We then aggregated forum attributes which contribute to the highest storage size and giving the higher impact view towards others. Next, we generate the cost of storage. Basics of this process are to measure the amount of storage used to store spam unit.

The rest of this paper is organized as follow. Section 2 starts by introducing cost of spam, the concept of spam unit including factors considered when evaluating spam unit and further focus on the definition of storage cost. In Section 3, we present the methodology used in this experiment. Section 4 covers on the detailed experiment setup followed by the result in Section 5. Section 6 explains on the analysis on the cost of this experiment. Finally, Section 7 concludes the paper and describes some potential future work.

2. THE COST OF SPAM 2.0

Spam 2.0 has its own effects either directly or indirectly. Directly, it will cause storage and bandwidth drainage, human resource efforts in eliminating spam and power consumption in running network resources. Hence, there are costs caused by Spam 2.0. Indirectly, Spam 2.0 can cause annoyance to the users, tarnish websites' reputation and decrease trust level towards services provided by the websites.

According to the economic view of stakeholder's business, minimising the costs particularly the direct cost is mainly crucial in achieving better financial expenditure as this will later affect company's profit. Choosing storage and bandwidth are two of the main factors to be considered in setting IT facilities. The question of cost is often raised in deciding and choosing the best packages provided by commercial business. Cost of these two can be affecting one's business as it will have a long run attachment towards the company's financial.

Hence, estimating the cost of Spam 2.0 becomes very crucial. This type of spam occurs only in web 2.0 environment which involve web 2.0 applications. Web 2.0 applications include forum, blog, social network application and wiki. Nevertheless, in this paper, we will focus on estimating the storage cost of Spam 2.0 anchored in a subset of web 2.0 applications which is the online discussion forum.

In this section, we first introduce the concept of *spam unit* followed by listing all attributes in a forum. We then define the forum attributes contained in each spam unit. Next, we define a general cost formulae for storage applicable for this experiment.

2.1 Spam Unit in Forum

Spam unit is defined as a group of attribute that can be manipulated by the spammers to embed their spam content. In a forum, all functions that are available to the users are equally accessible to the spammers. They could manipulate functions to fulfill their spamming purpose such as registering new profile, casting vote for a poll, create a poll, sending personal messages to other members and posting messages in a forum. There are well defined attributes for each spam unit. Hence, we now come to describe which attributes in a forum could be spammed by spammers. These attributes are shown in Table 1.

Table 1. Spam units' maximum storage and impact factor.

No	Spam Units	Storage	Max. Storage	Impact Factor	
				Possible No of Viewers	Attribute Creation
Profile					
1	Signature	64KB	H	H	L
2	Url	256B	M	H	L
3	Title	256B	M	H	L
Poll					
4	Poll Question	256B	M	H	H
5	Choice	3B	L	H	L
Personal Message(PM)					
6	PM content	64KB	H	L	H
7	PM Subject	256B	M	L	L
Post					
8	Post content	64KB	H	H	H
9	Post subject	256B	M	H	H

H: High, M: Medium, L: Low

Spam units are evaluated based on two characteristics, which are maximum storage and impact factor. Impact factors of these attributes are determined by two features which are the possible viewers' rate and attribute creation flexibility.

Maximum storage factor refers to the storage defined by the administrator for each attributes. Each attribute can be classified into three categories based on storage size factor that is high, medium and low. We categorize an attribute as high if the maximum storage is 64KB, medium if it is 256B and low category if they consumed small storage (3B-80B).

The possible viewers' rate represents the likelihood of each attributes being viewed by users. High possible number of viewers means the attribute is viewed by more than one user. On the other hand, an attribute is considered under low category if the number of viewers is one or none.

Another listed feature for the impact factor is the flexibility in attribute creation that explains the way each attribute can be created and manipulated by spammer. An attribute is classified as high if the attribute is easy to be created and manipulated. On the other hand, an attribute is categorized under low category if the attribute is easy to be created but hard to be manipulated.

Even though there are more spam units involved in a forum process, we decided to list only those attributes which we believe will give the highest contribution towards spam storage.

2.1.1 Spam Profile

Most forums require their users to register before allowing them to use the forum. In order to register, users have to fill up the registration form which involves filling up information such as username, password, email address, gender, birth date, website title, website url, location, AIM, MSN, YM, signature, etc. As we have mentioned earlier, we are not going to consider all the attributes created during registration. Instead, we are just going to consider attributes that contribute to the highest storage size. According to the forum setting, the highest storage size in our forum are users' website title, website url and signature. Therefore, we define spam profile as:

$$profile = title + url + signature \dots\dots\dots(Eq. 1)$$

Registering a new profile will allow the spammers to post spam in the forum. In order to successfully register a new profile, users need to pass the Captcha test. Unfortunately, with the use of forum spam automator, most forum that applied Captcha can be defeated [3,7]. Thus, it is back to the forum administrator to clean the forum from these spammers.

2.1.2 Spam Poll

A forum user normally could create a question poll and allow other users to cast a vote for each poll. Even though a poll question can be very long, but when user cast a vote, the choice has to be prepared by the poll creator and thus, each choice casted by a user can be either only a character or a numeric. We define spam poll as:

$$poll = poll\ question + \sum choice \dots\dots\dots(Eq.2)$$

2.1.3 Spam Personal Message

One of the functions available when using a forum is that each forum users can communicate with each other privately through personal messages. When sending a personal message to other user, this forum user could send a personal message subject with its personal content that could contains text and links. Personal messages then will be kept in the server until the forum user who received it decided to delete this personal message. Thus, we define spam personal message as:

$$personal\ messages = pm\ subject + pm\ content \dots\dots\dots(Eq. 3)$$

2.1.4 Spam Post

From a spammer's point of view, instead of using personal message to spam thousands of users, it is easier to just use post messages as this spam unit will have a higher viewer impact. If a forum is an open forum where it can be viewed by non-registered users, then their spam post will even have a higher viewer impact. Furthermore, when they are posting a post or a reply to an original post, there is no limit on how long a message can be posted. Hence, targeting to spam using post is very beneficial to spammers. We define spam post as:

$$post = post\ subject + post\ content \dots\dots\dots(Eq. 4)$$

Looking at how each forum attributes can be manipulated by spammer to become a spam unit, it is essential for us to see how this will affect the consumption of network resources such as the storage.

Setting up a forum is not a hard task. What a forum administrator needs is just a domain name, web hosting packages such as phpBB or SMF and a forum is ready. The cost of setting up a forum varies depending on the web hosting packages which includes storage and bandwidth quota. It is important to define the total cost in this experiment as the cost of storage and any additional costs related to setup a discussion forum.

$$Total\ Cost, TC = Storage\ Cost + Related\ Cost^1 \dots\dots\dots(Eq. 5)$$

2.2 Storage Cost

Storage cost is the cost paid by forum owner for server storage to store forum-related content [10]. Each function provided in the forum will contribute to expanding the storage. This cost is incurred every time a new user is registered, or when a user is sending personal message to others, or when a new post is posted in the forum. While each function seems to contribute only to a very small portion of increasing the storage size, it only happens if the user is a real user. If it is a spammer, they could be posting thousands of posts at one time hence increasing storage rapidly. We define the cost of storage as:

$$Storage\ Cost, SC = amount\ of\ storage * cost\ per\ GB\ of\ storage^2 \dots\dots\dots(Eq. 6)$$

3. METHODOLOGY

In order to achieve the objectives of this experiment, we present the methodology as in Figure 1. This methodology comprises of five steps as shown below.

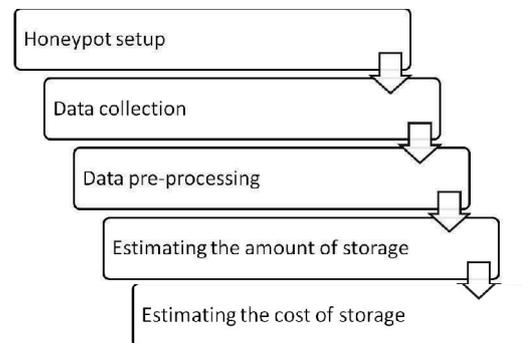


Figure 1. Methodology for estimating the cost of storage.

3.1 Honeypot Setup

In this experiment, we would like to gather spam posts naturally as what has been posted in real-world forums. Instead of downloading a real forum and classifying post as spam or not, we decided to construct our own dataset by setting up a forum.

¹ Related costs are any associated cost that can be explicitly defined for all cost sources such as bandwidth cost, filtering services cost, redundancy data cost, and content delivery cost.

² 1GB= 1024MB, 1MB= 1024KB and 1KB=1024B.

3.2 Data Collection

Forum data has been collected from our previous work [11] from June 2010 to June 2011 for a total duration of 13 months. It is noted that this forum started just like a new forum taking some time before the advertisement could reach the spammers and spammers start spamming this forum. Forum is left running as usual without any moderation towards any activities happened in the forum. The forum went offline in December 2010 and January 2011. Hence, there is no data collected in these months. Maintenance for our server has been done on 18-20 April 2011; therefore data for this duration are also omitted.

3.3 Data Pre-processing

Data pre-processing stage includes extracting related data from the forum such as post messages, profile, personal messages and poll. We eliminated other irrelevant data such as the tracking and log data. Cleaning the data involves taking only data that we used in this experiment such as what we have defined in Section 2.1.

3.4 Estimating the Amount of Storage

In this stage, we need to estimate the amount of storage. Figure 2 outline the details of what we want to achieve from the experiment. The end result will be the total cost consisting storage cost and other related costs. The input for storage cost is the size used for each spam unit.

All content in the forum is stored in as HTML. Total size of spam units will be determined by the number of characters from each spam units. Each spam units will have its own smaller forum attributes such as what we have defined in subsection 2.1. In order to determine this value, we need to calculate the number of characters in each spam units, then calculating total number of characters for each forum attributes. Each ASCII character will contribute to 8 bits or 1 byte of size. This process will produce the total size used for each spam unit. Therefore, we define the size to be used in order to determine the amount of storage as:

$$size = \sum_{profile} + \sum_{poll} + \sum_{personal\ message} + \sum_{post} \dots (Eq. 7)$$

3.5 Estimating the Storage Cost of Forum Spam

Estimating the cost of forum requires the detailed definition of all costs involved in setting up and running a forum. This section will further explain costing elements taken into account for measuring total cost of forum spam. The cost of forum spam can be estimated based on three sources, which are self-hosted server, commercial web hosting and cloud hosting.

4. EXPERIMENTAL SETUP

This section will explicate on the experiments done in order to estimate the storage cost for each spam unit. We divided this section into honeypot setup and storage cost survey.

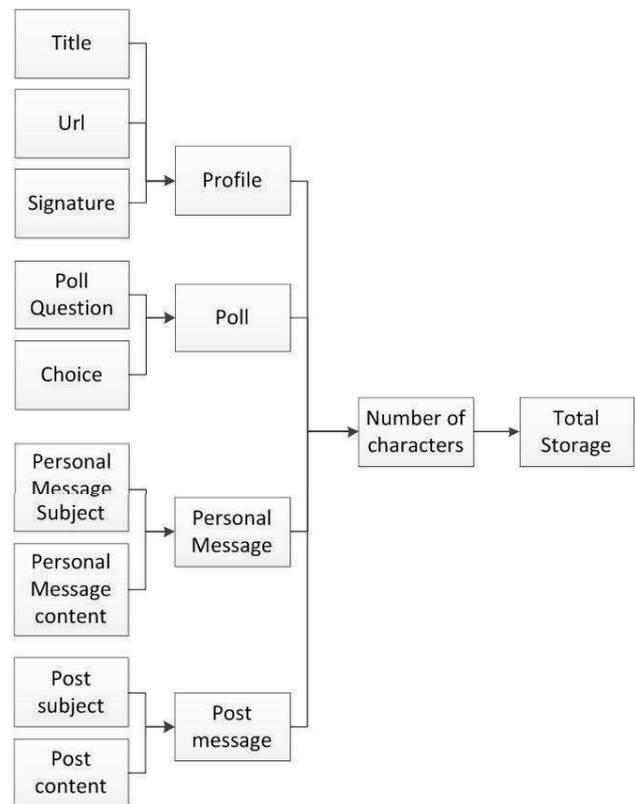


Figure 2. Estimating the amount of storage.

4.1 Honeypot Setup

In order to get the amount of storage used in this experiment, we first need to setup a honeypot. We set up an SMF forum and advertise it by listing the URLs in Pligg sites. It has been recognized that users of this forum are spammers which created fake profile and posted spam [12]. We set up this forum similarly as normal forum where the users need to login before they are allowed to use the forum. We also applied a change of setting in this forum. Since April 2011 onwards, we implemented Captcha for forum registration as the first layer of security.

The forum goes online for the whole duration and continues collecting spam data except when server maintenance is done. We then continue the process as stated in Section 3.2 and 3.3. The objective of this experiment is to collect and measure the amount of storage consumed for all four spam units.

4.2 Storage Cost Survey

As we have stated in Section 3.5, we will estimate the amount of storage based on three sources, which are self-hosted server, commercial web hosting and cloud hosting. We now explain on how to obtain the storage cost for each source.

4.2.1 Self-hosted server

This type of cost will only consider our real cost which includes bandwidth and storage costs. Our self-hosted server cost is AUD8000 for 146GB hard disk drive with 5 years lifetime. We were given 200GB on peak and 200GB off peak quota for the ADSL connection for AUD160 per month. We also have to pay for AUD50 per year for all services including the domain name. From this specification, we define the related cost as follow:

- cost per GB of storage for a month lifetime $_{source 1} = (8000/146)/60 = \text{AUD}0.91$
- cost per GB of bandwidth $_{source 1} = \text{AUD}160/400 = \text{AUD}0.40$ per GB per month
- additional cost $_{source 1} = \text{AUD}50$

It is unlikely that all forum administrators have their own server; hence in order to run a forum, they usually opt for the web hosting packages provided by commercial company, we now describe those packages.

4.2.2 Commercial web-hosting

Estimating the cost of storage based on commercial web-hosting sources would be the hardest task because we cannot split up the charge of storage and bandwidth specified by the company. Even in commercial business, there are two ways of charging customers [13], which are to charge the customers based on their usage, or customer beforehand will choose a package consists of a certain amount of storage and bandwidth quota. If they exceeded this quota, they will have to pay extra charges for their usage.

Instead of just relying on our source, we are going to investigate the top 5 recent commercial companies. These companies are providing several packages with different prices to cater all type of customers. We consider 21 basic web hosting packages from them. These packages usually provide the ratio of storage space to bandwidth quota ranges from 1:20 to 1:2 with both the median and average ratio is 1:3. The price per month for all these packages is also in a wide range depending on the storage and bandwidth quota given. Since these costs cannot be explicitly identified such as CPU memory, storage, bandwidth and maintenance, we assume that 25% of the cost provided by the packages goes toward storage and backup. Another 75% would go to other related costs such as bandwidth, server maintenance, human resource cost, dedicated IP addresses, database, email and FTP accounts. Hence, we took the average of storage prices per GB for these 21 packages which is AUD5.15 per GB per month.

4.2.3 Cloud-hosting

Cloud hosting packages provided by commercial companies also includes several important costs. They allow their customers to choose a preferred operating system, database, resource management software, web hosting software and application development servers. We noticed that the prices specified in most of the packages are almost similar. Hence, we took the average price between three packages (namely Amazon, Microsoft and Ninefold) and define the cloud hosting costs as follows:

- Storage per GB permonth = AUD0.13
- Redundancy Storage per GB permonth = AUD0.118
- Content delivery per GB permonth = AUD0.195

5. RESULT

This section refers to the stage of estimating the amount of storage. From June 2010 to June 2011, a total of 62,798 spam profiles were created, 141 spam personal messages were sent between the forum users, and 450,772 spam posts messages were posted in the forum. No polls were created in this duration, therefore there is no storage cost for poll spam unit.

We first present basic spam statistics from the dataset in order to find out which spam units preferred to be manipulated by spammers and spam units that consumed more storage. We also

investigate spamming volume for each spammer. This section further explains each spam unit statistics including spam profile, spam personal message and spam post explicitly in Section 5.1.2, 5.1.3 and 5.1.4.

5.1.1 Statistics from Dataset

Total spam collected in our discussion forum for each month is shown in Table 2. Storage size in this table is generated based on Equation 7 which is the summation of spam profiles, spam personal messages and spam posts. We are going to divide the time frame into two phases. The first phase includes June 2010 to November 2010 and the second phase includes data collected from February 2011 to April 2011.

Table 2. Total spam in the discussion forum.

Month-Year	Total spam	No of characters	Size (MB)	Accumulated Size (MB)
Jun-10	78	257,429	0.2455	0.2455
Jul-10	12,181	9,444,9617	90.0742	90.3197
Aug-10	31,880	224,943,266	214.5226	304.8423
Sep-10	55,258	364,397,424	347.5165	652.3588
Oct-10	74,350	670,342,690	639.2886	1291.6474
Nov-10	59,458	656,084,770	625.6912	1917.3386
Feb-11	30,113	57,827,566	55.1487	1972.4872
Mar-11	113,281	271,096,021	258.5373	2231.0245
Apr-11	62,318	223,881,765	213.5103	2444.5348
May-11	41,024	181,946,302	173.5175	2618.0523
Jun-11	33,770	145,683,223	138.9343	2756.9867
Total	513,711	2,891,363,782	2756.9867	

It is obvious that the highest spam we obtained in the first phase is in October 2010 followed by November 2010 which is reflected by the storage size consumed in those months. June 2010 as expected, is having the lowest total number of spam because we just started the forum in June 2010.

In second phase which includes the duration of February 2011 to June 2011, the forum has received the highest number of spam in March 2011 followed by April 2011. Comparing the first two months in first phase and second phase, we could say that second phase has made a sudden increment for the total number of spam indicated that spammers are faster in reaching out to forum where they could send more spam.

We will now break down this section to further show the spam unit preferred to be manipulated by spammers, spam unit that consumed more storage and spammer's spamming rate.

5.1.1.1 Spam unit preferred to be manipulated by spammers

Previous section showed the total number of spam. Obviously, spammers are sending spam using personal messages, posts and through profile creation. They only do not initiate a poll. We now further see the spam units preferred to be manipulated by spammers as shown in Figure 3.

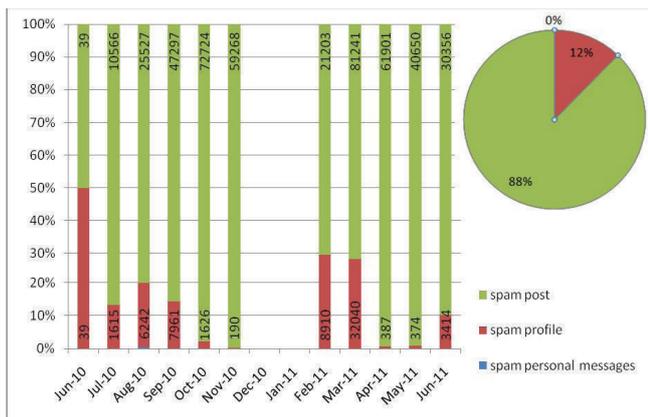


Figure 3. Total number of spam units with its percentage according to month.

Based on Figure 3 above, it is clear that spam units preferred to be manipulated by spammers is spam post as it is accounted for 88% from all spam units followed by profile 12%. Eventhough there are 111 and 30 spam personal messages sent in August 2010 and April 2011, personal messages accounted the lowest spam units received compared to the overall number of spam making it close to only 0%. It is also obvious that the trend using post as their spam units started from the earliest month, June 2010 where this value accounted for 50% and this continued to grow.

Despite of still receiving spam profile in October 2010 and November 2010, the percentage of using spam post as their main medium for spamming accounted almost 100% from overall spam units. From February to April 2011, spammers posted more than 70% of spam posts compared to other spam units. Spam post is the most preferred unit by spammers as they could be send in massive numbers using forum spam automator [3] and the impact view is high compared to using other spam units. Spam profile is obviously very low in April 2011 and May 2011 because of the Captcha implementation. Still, the forum is not 100% free of spam, as there are still successful registration made by spammers.

Considering that spam post is the most preferred spam unit to be used by spammers, we need to consider creating a higher security layer before enabling a post being posted in the forum but this will also burden real users.

5.1.1.2 Spam unit consumed more storage

Eventhough spammers are using all spam units for spamming and all these spam units could probably have the same maximum storage, it should be noted that spammers are spamming all these spam units at a different spamming rate. Hence, we now present the spam unit that consumed more storage in Figure 4.

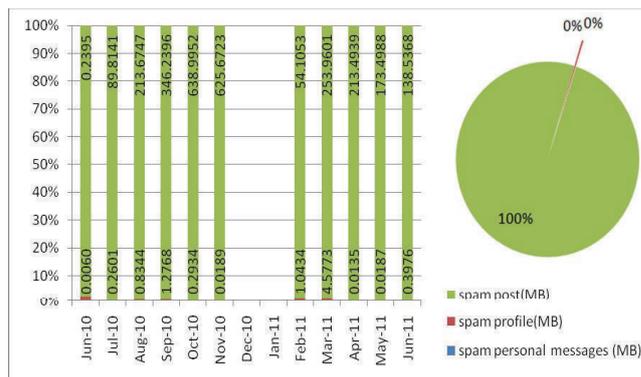


Figure 4. Size for each spam units with its percentage according to month.

Figure 4 reveals that spam posts accounted for the highest spam unit that consumed storage with almost 100%. It is remarkable to see the percentage for spam post. Although the forum received quite a number of spam profile and personal messages, this graph obviously show that spammers prefer to use spam post and spam with a lot of characters making it the highest spam unit that consumed storage.

5.1.1.3 Spammer's spamming rate

It is apparent that spammers are choosing spam post as their main spamming medium and therefore, spam posts consumed more storage. Nevertheless, it is also noted that each spammers has different spamming volume. Details of spammer's spamming rate are now shown in Figure 5.

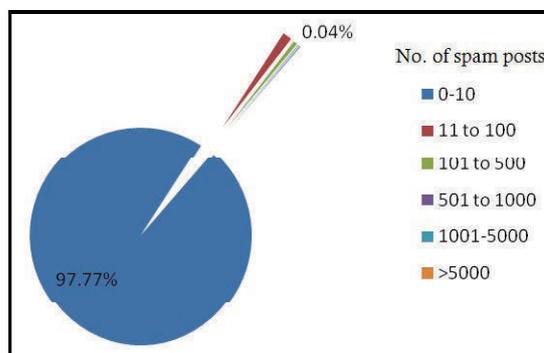


Figure 5. Percentage of spam posts sent by spammers divided into six categories.

Throughout the dataset, it is observed that spammers do not spam at the same volume. We categorized the data into six categories based on the number of spam posts sent by spammers as shown in Figure 5. These includes spammers that send less than 10 spam posts during these 6 months duration, 11 to 100, 101 to 500, 501 to 1000, and more than 5000 posts. It is apparent that most spammers send less than 10 spam post as that category accounted the largest percentage with 97.77%. Spammers that send spam more than 5000 spam posts only accounted for 0.04% from overall data. Other categories also accounted less than 1%. We further investigate the largest percentage categories here in Figure 6.

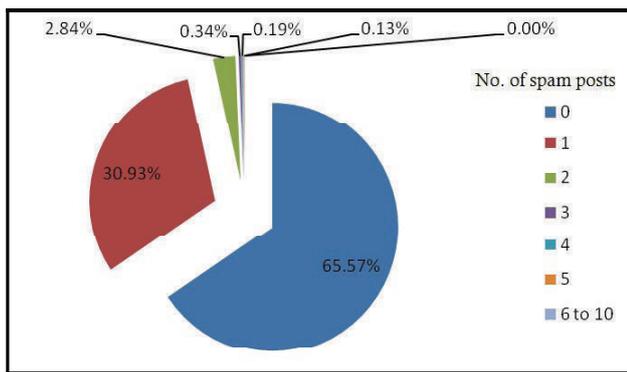


Figure 6. Percentage of spam posts sent by spammers for the first category (0-10).

The percentage of a more detailed fraction for the first category as we have described earlier are now presented in Figure 6. There are no spammers who have spammed with 6 to 10 posts; hence the percentage for this category is 0%. The largest fraction accounted for 65.57% for 0 posts shows that there are quite a huge number of spammers who just registered in the forum but have not posted anything. It is unclear of what their motives are, but they could still spam with a huge number of characters while registration.

Second large fraction owned by spammer who posted 1 post accounted for 30.93%. It is wise to take note that even though some of the spammers send only 1 spam, the size of each spam is different. Hence, there is a probability that they send only 1 spam post but with large number of characters.

Furthermore, this analysis is based on the fact that different profile owned by different spammer. There is a possibility that they are logged in as different user while they are spamming. It could be the same spammer logged in as different users and send spam at one time. Tracking their geo location from the IP address is not sufficient enough to track whether each profile originated from the same source as one spammer could be using different IP address for each account. Next sections will further reveal the detailed amount of storage consumed by each spam unit.

5.1.2 Spam Profile

This section will focus only on spam profile data obtained from the dataset. Table 3 which consists of spam profile data is generated based on Equation 1. Table 3 presents the number of spam profile created from June 2010 to June 2011. Total number of characters for spam profile are addition of three forum attributes which are website title, website url and signature. June 2010 was the month with the lowest number of profile created as we just started our forum and started advertising it. The number started increasing and continued to grow gradually in the following months.

In the first phase, September 2010 recorded the highest number of profile created with 7,961 spam profiles followed by August 2010 with 6,242 spam profiles created. Total number of characters seems to be proportional with the number of spam profile created in a month where spam profile is registered with around 100-200 characters. It is also interesting to see that there is a sudden drop in the number of profile created in November 2010. There is a possibility that there are only a few spammers attacking our websites contributing to a very low number of registrations per day. There is an average of only 6 new registrations per day and 7 days without any new registrations.

Table 3. Profile spam.

Month-Year	No of profile	No of characters	Size (MB)	Accumulated Size (MB)
Jun-10	39	6,341	0.0060	0.0060
Jul-10	1,615	272,715	0.2601	0.2661
Aug-10	6,242	874,888	0.8344	1.1005
Sep-10	7,961	1,338,842	1.2768	2.3773
Oct-10	1,626	307,644	0.2934	2.6707
Nov-10	190	19,815	0.0189	2.6896
Feb-11	8,910	1,094,089	1.0434	3.7330
Mar-11	32,040	4,799,605	4.5773	8.3103
Apr-11	387	20,281	0.0135	8.3238
May-11	374	19,627	0.0187	8.3425
Jun-11	3,414	416,908	0.3976	8.7401
Total	62,798	9,164,648	8.7401	

In the second phase, spammers created the highest number of profiles in March 2011 with 32,040 profiles and 4,799,605 number of characters. There were not many profiles created in February 2011 as the forum just restarted. We also implemented Captcha for our registration hence there is a sudden drop of spammer's registration in April and May 2011. The number of profile registered in June 2011 increased again in June 2011. Since we did not change any settings in the forum, we assume that spammers are using smarter bots in order to break our Captcha and bypass the first security layer, hence successfully registered as user. Storage used to keep all these spam profile is 2.6896MB in the first phase and this accumulates to 8.7401MB in June 2011.

5.1.3 Spam Personal Message

This section will now focus on the next spam unit which is the spam personal message. Table 4 is generated based on Equation 3 where the storage size of personal message is an addition of personal message subject and personal message content.

Table 4 shows the number of spam personal messages with total number of characters and storage size from June 2010 to November 2010. Total number of characters for personal messages is based on the summation of personal messages subject and its content. Based on this table, spam personal messages are only sent in August 2010 with 14,214 total number of characters and in April 2011 with 3000 characters consumed less than 1MB of storage. This does not happen often therefore it can be concluded that spammers does not prefer to use personal messages as their spamming unit.

Table 4. Personal message(pm) spam.

Month-Year	No of pm	No of characters	Size (MB)	Accumulated Size (MB)
Jun-10	0	0	0	0
Jul-10	0	0	0	0
Aug-10	111	14,214	0.0136	0.0136

Sep-10	0	0	0	0.0136
Oct-10	0	0	0	0.0136
Nov-10	0	0	0	0.0136
Feb-11	0	0	0	0.0136
Mar-11	0	0	0	0.0136
Apr-11	30	3,000	0.0028	0.0164
May-11	0	0	0	0.0164
Jun-11	0	0	0	0.0164
Total	141	17,214	0.0164	

5.1.4 Spam Post

Total number of spam post with total number of characters is shown in Table 5. A spam post size is calculated based on the total number of characters in post subject and post message as what we have defined in Equation 4.

Table 5. Spam post.

Month-Year	No of posts	Size (MB)	Ratio of (Post Size/No of Posts)	Accumulated Size (MB)
Jun-10	39	0.2395	0.0061	0.2395
Jul-10	10,566	89.8141	0.0085	90.0535
Aug-10	25,527	213.6747	0.0084	303.7283
Sep-10	47,297	346.2396	0.0073	649.9679
Oct-10	72,724	638.9952	0.0088	1288.9631
Nov-10	59,268	625.6723	0.0106	1914.6354
Feb-11	21,203	54.1053	0.0026	1968.7407
Mar-11	81,241	253.9601	0.0031	2222.7007
Apr-11	61,901	213.4939	0.0034	2436.1946
May-11	40,650	173.4988	0.0043	2609.6934
Jun-11	30,356	138.5368	0.0046	2748.2302
Total	450,772	2748.2302	0.0061	

Our forum received the highest spam post with 72,724 in October 2010 and the lowest spam post with only 39 in June 2010. Interestingly, even though October 2010 stated as the month receiving the highest number of spam posts, the next month which is November 2010 recorded as having the highest ratio of size of each posts which means that spammers send spam with a larger text in this month. This further indicates that the number of spam posts does not signify the size of the post.

Despite of having a huge decrement of new profile registration in November 2010 for the first phase, the forum still received quite a huge amount of spam post messages posted in the forum. This situation points out that the spammers are just using the same profile created before to send spam post.

Table 5 shows that the number of spam posts received in the forum is decreasing in second phase. Since we did not amend any content in the forum, the storage consumed to store all spam post messages reaches nearly 3GB in June 2011.

6. ANALYSIS OF STORAGE COST

This section which refers to stage estimating the cost of storage will further provide a cost analysis from the data obtained from our experiments in an attempt to answer the following questions:

- What is the total cost for the discussion forum?
- What is the average cost for each spam post?

6.1.1 Total cost for discussion forum

Based on Equation 6, we now present the storage cost according to month and total accumulated cost based on self-hosted server, commercial web-hosting packages and cloud hosting packages in Table 6.

Total storage cost for our self-hosted server is AUD23.66; for commercial web-hosting is AUD133.90 which is highest price; and for cloud hosting is AUD11.53. Even though these costs are pretty low but it can be seen that it will increase rapidly once it exceeds the standard storage in a package. This will also happen accordingly if we took basic packages that contain certain amount of storage. The additional cost will increase dramatically once the storage space goes beyond the quote given. Hence, it is very important to choose a suitable package in the first place.

Table 6. Storage cost.

Month-Year	Accumulated Size (MB)	Cost (AUD)		
		Self-hosted server	Commercial web-hosting	Cloud hosting
Jun-10	0.2455	0.91	5.15	0.44
Jul-10	90.3197	0.91	5.15	0.44
Aug-10	304.8423	0.91	5.15	0.44
Sep-10	652.3588	0.91	5.15	0.44
Oct-10	1291.6474	1.82	10.30	0.89
Nov-10	1917.3386	1.82	10.30	0.89
Dec-10	1917.3386	1.82	10.30	0.89
Jan-11	1917.3386	1.82	10.30	0.89
Feb-11	1972.4872	1.82	10.30	0.89
Mar-11	2231.0245	2.73	15.45	1.33
Apr-11	2444.5348	2.73	15.45	1.33
May-11	2618.0523	2.73	15.45	1.33
Jun-11	2756.9867	2.73	15.45	1.33
Total		23.66	133.90	11.53

Perhaps, it is better to see these situations from a larger point of view. It is clear that the costs will expand continuously until it reaches a point where the forum administrator needs to pay a large amount of money and this will happen only if the forum administrator is not doing anything towards the content kept in the server. It is also the reason why people opt to choose filtering services. These services aim to provide plug in to remove spam comments and act as the first layer security so that the administrator does not have to personally check on each post. We now consider the cost of implementing commercial filtering services listed in Table 7.

Table 7. Commercial filtering services.

Filtering Services	Prices Permonth	Notes
Akismet	AUD47	Limited for 5 sites, unlimited posts per month
Mollom	AUD40	Limited for 1 site, Unlimited spam posts, 1000 legitimate posts per day

Using Equation 5, the total costs is absolutely higher when implementing these filtering services than the cost that we have previously calculated in Table 6. Considering that these filtering are effective then the possibility of exceeding the storage quota because of spam posts is extremely low. Furthermore, without implementing these filtering services, the forum administrator will have to eliminate spam by themselves, which would incur additional human resource cost. We will investigate this costs in our future research.

Still, if the forum administrator decides to implement these filtering services, these services will only filter the new spam coming into the forum. Hence, it is the forum administrator's task to eliminate the old spam residing in the forum and there is definitely a cost associated with this activity.

6.1.2 Average Storage cost

This section will now further focus on the analysis for the storage cost. The objective of this section is to work out the storage cost for each spam units. In order to do that, we need to obtain the storage cost per MB and the average size of each spam unit.

We first determine the storage cost per MB used for self-hosted server, commercial web hosting and cloud hosting sources. Total storage used for spam units that we have defined earlier in Table 1 is 2756.9876MB. Therefore, we now estimate the average cost for the research period based on three sources as in Table 8.

Table 8. Storage cost per MB.

Source	Storage Cost	Storage Cost per MB
Self-hosted server	AUD23.66	AUD0.0086
Commercial web hosting	AUD133.90	AUD0.0486
Cloud hosting	AUD11.53	AUD0.0042

Next, we estimate the average size for each spam units; which is profile spam, personal message spam and post spam. We took the average size for each spam units and work out the calculation as in Table 9.

Table 9. Average size of spam units.

Spam Unit	No of spam	Storage Size (MB)	Average Size (MB)	Average Size for 10,000 posts (MB)
Profile	62,798	8.7401	0.000139	1.39MB
Personal message	141	0.0164	0.000116	1.16MB

Post	450,772	2748.2302	0.006097	60.97MB
------	---------	-----------	----------	---------

We then estimate the storage cost for 10,000 spam for each spam units based on self-hosted server, commercial web hosting and cloud hosting sources. These costs are as listed in Table 10 below.

Table 10. Storage cost for 10,000 spam units.

Spam Unit	Self-hosted server	Commercial web-hosting	Cloud hosting
Profile	AUD0.012	AUD0.068	AUD0.006
Personal message	AUD0.010	AUD0.056	AUD0.005
Post	AUD0.524	AUD2.963	AUD0.256

It is apparent that the highest cost of each spam units will come from commercial web-hosting packages as the basic cost itself is higher than other sources and spam posts is the spam unit that costs the highest as it consumed more storage.

7. CONCLUSION AND FUTURE WORK

This paper is an early work to produce a cost model for Spam 2.0. We have presented ways of estimating the amount of storage used and produced three total costs covering storage and related cost. Our costs have shown that without any further actions taken, spam definitely waste network resources. Furthermore, these data has gone through the pre-processing stage. Perhaps, with the original posts and legitimate posts, storage consumed to store all the related attributes is bigger and the cost would be higher. Still, spam has its tangible and intangible effect. Apparently, the storage used for spam could increase the cost of storage. Spam also affects user's trust towards the services provided by the websites thus putting website's reputation at stake.

In addition, from the data collection, we could also use content based analysis and behavior based analysis to further learn about spammers and detect forum spamming. Future work could also involve with link spamming detection in the content. In order to further improve our cost model, we plan to further investigate the cost for a longer duration using continuous data. This experiment is only using storage price from three different sources. Future improvement on these prices could further improve our costing models. Unlike cloud hosting packages which clearly specified their price of storage per GB, commercial web hosting packages do not specify them. Hence, we only assume in this experiment that 25% of packages price is intended for storage and backup. The cost model could also be improved if we could further break down the cost elements to identify related cost specifically. We also consider other costs involves with spamming such as human resource costs, electricity cost and software costs which will be further explored.

8. REFERENCES

- [1] Yin, D., Davison, B.D., Xue, Z., Hong, L., Kontostathis, A., and Edwards, L. 2009. Detection of Harassment on Web 2.0. In *Proceedings of the Content Analysis In The Web 2.0 (CAW2.0) Workshop At WWW2009*. (Madrid, Spain, April 21, 2009). CAW 2.0 2009. Madrid, Spain.
- [2] Hayati, P., Potdar, V., Talevski, A., Firoozeh, N., Sarenche, S., and Yeganeh, E.A. 2010. Definition of spam 2.0: New

- spamming boom. In *4th IEEE International Conference on Digital Ecosystems and Technologies*. (Dubai, United Arab Emirates, April 12-15, 2010). IEEE DEST 2010. IEEE, Dubai, UAE, 580-584. DOI= 10.1109/DEST.2010.5610590
- [3] Shin, Y., Gupta, M., and Myers, S. 2011. The nuts and bolts of a forum spam automator. In *Proceedings of the 4th USENIX conference on Large-scale exploits and emergent threats*. (Berkeley, CA, USA, March 2011). LEET '11. USENIX Association, Berkeley, CA, USA, 3-3.
- [4] Shin, Y., Gupta, M., and Myers, S. 2011. Prevalence and mitigation of forum spamming. In *the 30th IEEE International Conference on Computer Communications*. (Shanghai, China, April 12-14, 2011) IEEE INFOCOM 2011. IEEE Computer Society, Shanghai, China.
- [5] Chai, K., Hayati, P., Potdar, V., Wu, C., Talevski, A. 2010. Assessing Post Usage for Measuring the Quality of Forum Posts. In *Proceedings of the 4th IEEE International Conference on Digital Ecosystems and Technologies (DEST)*.
- [6] Potdar, V., Ridzuan, F., Hayati, P., Talevski, A., Yeganeh, E.A., Firuzeh, N. and Sarencheh, S. 2010. Spam 2.0: The Problem Ahead. In *Computational Science and Its Applications – ICCSA 2010*, D. Taniar, O. Gervasi, B. Murgante, E. Pardede, and B.O. Aduhan, Eds., Springer Berlin/Heidelberg. 400-411.
- [7] Niu, Y., Wang, Y. M., Chen, H., Ma, M., and Hsu, F. 2007. A Quantitative Study of Forum Spamming Using Context-based Analysis. In *Proceedings Network and Distributed System Security (NDSS) Symposium*. February 2007.
- [8] Motoyama, M., K. Levchenko, C. Kanich, D. McCoy, G.M. Voelker, and S. Savage. 2010. Re: CAPTCHAs– Understanding CAPTCHA-solving services in an economic context. In *Proceedings of the 19th USENIX Conference on Security Symposium*. August 11-13, 2010, Washington DC.
- [9] Benjamin Markines, B., Cattuto, C. and Menczer, F. 2009. Social spam detection. In *Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web (AIRWeb '09)*. Dennis Fetterly and Zoltán Gyöngyi (Eds.). ACM, New York, NY, USA, 41-48. DOI= 10.1145/1531914.1531924
- [10] Ridzuan, F., Potdar, V., Talevski, A., and Smyth, W.F. 2010. Key Parameters in Identifying Cost of Spam 2.0. In *Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications*. Washington, DC, USA: IEEE Computer Society, 2010, 789–796. DOI = <http://doi.ieeeecomputersociety.org/10.1109/AINA.2010.163>
- [11] Hayati, P., Chai, K., Potdar, V. and Talevski, A. 2009. HoneySpam 2.0: Profiling Web Spambot Behaviour. In *Principles of Practice in Multi-Agent Systems*, J.-J. Yang, M. Yokoo, T. Ito, Z. Jin and P. Scerri Eds. Springer Berlin / Heidelberg, 335-344.
- [12] Hayati, P., Potdar, V., Chai, K., and Talevski, A. 2010. Web Spambot Detection Based on Web Navigation Behaviour. In *24th IEEE International Conference on Advanced Information Networking and Applications (AINA) 2010*, Apr. 2010, IEEE Computer Society. 797-803. DOI = <http://doi.ieeeecomputersociety.org/10.1109/AINA.2010.92>
- [13] Yaiche, H., Mazumdar, R.R., and C. Rosenberg. 2000. A game theoretic framework for bandwidth allocation and pricing in broadband networks, *IEEE/ACM Transactions On Networking (TON)*, v. 8, n.5, October 2000, 667-678. DOI = 10.1109/90.879352

Key Parameters in Identifying Cost of Spam 2.0

Farida Ridzuan, Vidyasagar Potdar, Alex Talevski, William F.Smyth
Anti Spam Research Lab,

Digital Ecosystems and Business Intelligence Institute, Curtin University of Technology, Perth, Western Australia
farida.mohdridzuan@postgrad.curtin.edu.au {v.potdar, a.talevski}@curtin.edu.au, smyth@mcmaster.ca

Abstract— This paper aims to provide an analytical view in estimating the cost of Spam 2.0. For this purpose, the authors define the web spam lifecycle and its associated impact. We also enlisted 5 stakeholders and focused on defining 5 cost calculations using a large collection of references. The cost of web spam then can be calculated with the definition of 13 parameters. Detail explanations of the web spam cost impacts are given with regards to the main four stakeholders: spammer, application provider, content provider and content consumer. Ongoing research in developing honey spam is also presented in this paper.

Keywords—component; web spam, cost of spam

I. INTRODUCTION

Spam in the context of email is defined as “unsolicited, anonymous and mass email messages” [1, 2]. Spam originated via email and one of the first spam emails dates back to the early eighties, when a lawyer sent out an advertising email on a newsgroup. Since then spam has evolved into what we know as spam today. A spammer is defined as “an entity that is involved in spamming”. Spammers use many different mediums to spam web users, drifting from the traditional email approach to new approaches that are termed Web Spam, Web 2.0 Spam or as we call it Spam 2.0 [3].

Spam 2.0 refers to the techniques employed by spammers to spread spam via websites in contrast to using emails. Spammers now use blogs, forums, wikis or even develop their own websites to post advertising material. Overall the motivation is still the same i.e. to generate revenue, increase page rank, promote product or services and steal user information [4].

Spammers use a number of techniques to drive traffic to their websites and one of those is to fine tune their websites to deceive search engines in increasing their ranking. It is quite common that when you search for a particular keyword, you are taken to a website which does not relate to what you are looking for, but instead it is an advertising page designed by spammers. Such websites are carefully crafted to make the search engines believe that it is providing genuine content by implementing keyword stuffing, incorporating fresh content and several other strategies [5].

Spamming activities affects number of different parties involved in the Web 2.0 spam lifecycle, which includes the developer (i.e. those who tries to implement anti-spam techniques like introducing CAPTCHA [6] to discourage spammers but also introduce another level of annoyance for users), spammer itself, followed by the Internet Service Provider (ISP), Application Provider (i.e. those who host blogs,

forums, wikis etc.), Content Providers (i.e. users who add rich content to blogs, forums, wikis etc.) and finally the Content Consumer (i.e. the actual end users who uses this rich content). There is a cost associated for each an every stakeholder in the spam lifecycle i.e. the application provider has to spend time or money to ensure that their blogs or forums are free of spam, the content provider also spends time to filter out spam from their blog comments or forum posts, and finally the content consumer is adversely affected if spam content bypasses all the filters and is published on the web, since they cannot find the right information that they are looking for.

It is understood that there is a cost incurred by each stakeholder at each and every stage of the Web 2.0 spam lifecycle, however, to the best of our knowledge, there is no prior work that looks into detail at various cost parameters involved in the Web 2.0 spam lifecycle. No one has so far analysed the Web 2.0 spam lifecycle itself. Hence this paper aims to:

1. Understand Spam 2.0
2. List the key Spam 2.0 stakeholders
3. Identify different Spam 2.0 cost categories and cost parameters
4. Derive the cost associated with each stakeholder

This paper has been organized as follows. Section II will provide a detailed description of the Web 2.0 spam lifecycle. It will outline all the different stages in the web 2.0 spam lifecycle and associate different stakeholders to different stages. This section will also list different tools used by different parties for spamming or anti-spamming. Having understood the spam lifecycle, Section III then describes different costs categories for Spam 2.0 and its associated parameters used in deriving actual cost. Section IV then shows cost impact for each stakeholder. Section V then explains the prototype developed for capturing Spam 2.0, we call it HoneySpam. The prototype is being developed to estimate the costs for each stakeholder in the Spam 2.0 lifecycle. Section VI provides some thoughts on future research, ongoing work and concludes the paper.

II. SPAM 2.0 STAKEHOLDERS

In this section we list all the main stakeholders in the Spam 2.0 lifecycle. These include;

- Developer
- Internet Service Provider (ISP)
- Application Provider
- Content Provider

- Spammer
- Content Consumer

A. Developer

Developer plays the role in developing programs or software to either help spammers or anti spammers. Most of the services provided by them are not free. Developer for the spammers' side will try to create program that could break the latest anti spam techniques. On the other hand, developer on the anti spammers' side will try to create new techniques or method to avoid spammers from successfully sending spam to the applications, such as the CAPTCHA [6]. Even though such techniques have been proven to be ineffective [7], they do slow down spam attacks. Nevertheless, programs that they create usually have a few drawbacks on the users. Generally, developer on both sides aims for high profit.

B. Internet Service Providers (ISPs)

Internet Service Providers (ISPs), as the name suggests, provide web hosting servers and services that can be accessed by both spammers and non spammers. ISPs also provide several other services such as selling domain names, email hosting and others. Some of the popular companies that provide internet access in Australia are BigPond, OptusNet, AAPT LiveNet, Virgin and Vodafone.

Both spammers and other stakeholders need access to the internet to send and receive spam which makes ISPs the connector between spammers and spam receivers. Spam is transmitted through the service that the ISP provides. In order to maintain a high service standard, ISPs must implement strategies to avoid unnecessary bandwidth hogging load and protect from successful intelligent attacks and many failed brute-force spam attacks.

Nonetheless, it is still unclear of how ISPs manage web spam. Not only is it hard to detect web spammers, it is impossible for ISPs to stop providing services to spammers. Even so, useless/wasteful contents are transmitted through the networks makes the network becomes slower and this affects client's satisfactory towards the services.

C. Application Providers

Application providers play an important role in the lifecycle as they host web applications. Some of the most popular applications are Wordpress, phpBB, SMF, and Blogger [8]. They are generic freely available Web 2.0 tools. They have many templates and plugins that enhance their operation in order to provide a better user experience and reach a greater user base. These plugins may provide better interfaces, embedded applications and spam filters. Application providers would also want to ensure that their blogs or forums are free of spam so they spend significant amounts of time and money to develop and integrate spam filtering tools such as CAPTCHA [6].

D. Content Providers

Content providers have the ability to add, edit and delete web content. They usually need to register for an account from the application provider. They could be the web administrator hired by a company, a paying sponsor, they could simply be an application user, or for instance, an author of a blog.

If we assume a world without spam, the real job of a content provider would be just to add rich content to their website. Unfortunately, with the existence of spammers, the content provider's tasks widen requiring them to maintain their application / websites to be spam free. They have to regularly check for spam comments, spam posts, spam reports from users, and this include determining and detecting whether it is spam or not. Without proper management, users or viewers of the applications would leave.

E. Spammer

Basically, the lifecycle of Spam 2.0 starts with the spammer. Spammers may work in a team in order to make the spam campaign a success [9]. Importantly, spammers also pay for people to manually spam websites [10]. Spammers use various techniques to spam Web 2.0 applications in order to make profits. They will not only try to bypass filters ensuring that their spam content can get through to the content consumers, but also to ensure that content consumers read their spam content and visit the links provided. The interesting thing that must also be considered is that such evidently lucrative jobs may take away from the regular labour force and / or may drive up labour prices. Furthermore, spammers then require whole new matching job position that is dedicated to spam prevention. Further reading on [9] could give an in depth knowledge of what a spammer is.

F. Content Consumer

A content consumer is the final stakeholder in the lifecycle of Spam 2.0. Similarly to the content provider, all spam content sent by spammers is basically targeted to the content consumers. They could be someone who is spam aware or they likely could be someone who has limited knowledge of spam. Using the internet, it is common that a user may stumble upon spam content and fall for it. This could mean that they may;

- Make a misinformed conclusion or decision (this could range from something very small to very large)
- Unable to find genuine content
- Spend additional time on a website as they filtering and searching through genuine and spam content
- Attempt to inform staff of the problem (which of course may redundantly occur many times by many users)
- Simply give up and no longer visit the site / lose interest
- Become emotionally frustrated, angry etc
- Being redirected to another site which may be one that replicates the original, is an advertising page or may even be something illegal and/or offensive.
- Their computer becomes infected with Malware [11].

III. LIFECYCLE OF SPAM 2.0

In this section we enlist the six main stages in Spam 2.0 lifecycle. These are as follows:

- Getting a list of URLs
- Creating Spam Content
- Sending Spam Content
- Filtering Spam at Application Level

- Filtering Spam at Personal Level
- Spam that Bypasses all Filters

The six stages are shown in Fig. 1 along with five stakeholders. The first three stages involve spammers, the next stage involves the application provider, followed by content provider and finally the content consumer. Between the third stage and the fourth stage, spam traverses from Spammer to ISP and to content provider. At this stage it is not clear on what steps are taken by the ISP to filter out Spam 2.0, we have neglected this party in further discussions. We are also neglecting developer from this point forward as developer can be considered working on both sides. We now explain each stage in detail.

A. Getting a list of URLs

This is the first stage in the Spam 2.0 lifecycle. Here the spammers compile a list of URLs pointing to vulnerable web 2.0 applications like blogs, wikis, forums etc. Such application URLs can be used for adding comments or links on forum, wiki or blog threads. Vulnerable web applications can be found using shareware or commercial tools like Win Web Crawler, Web Data Extractor, Rafabot, Extract Link, Online Data Extractor, Visual Web Spider, Hrefer and Teleport [12-19]. Some of these tools are free for a limited time while others come with limited features in the free version. Alternatively, spammers may be opting for freeware such as Elite Web Crawler, WebReaper, URL Spider Pro, Heritix and WebSphinx [20-24].

It is not sure whether spammers are using any other sophisticated tools for crawling vulnerable sites or even detect dead links before actually spamming. Manual detection of dead links will be costly hence spammers may just spam all the collected links, given that the cost to spam 1 or a million websites would be marginal. From anti-spam perspective, web administrator could take some actions to prevent URL fetching by controlling which bots crawl their sites or index their pages.

B. Creating Spam Content

This is the second stage in the Spam 2.0 lifecycle. Creating spam content such that it can deliver the right advertising message while at the same time bypasses all anti-spam filters, is an extremely challenging task. Spammers are using intelligent methods to achieve this goal. It is observed that they create content using a combination of text messages, links and images [25]. It is understood that spammers have developed database of words, phrases and pictures for periodic use. It is also possible for the spammers to use SEO Text Generator or Keyword List Generator to create good spam content messages. In order to avoid being detected as spam content, spammers develop unique content so as to avoid being blacklisted. One of the spammers' tools that include this feature is X-Rumer Palladium [18].

The ultimate motivation for spammers is to provide a link or build a link farm that could generate revenue for them.

C. Sending Spam Content

This is the third stage in the Spam 2.0 lifecycle. Spammers can either manually or automatically send spam to Web 2.0 applications. If it is done manually, it can be done to a more specific target but as compared to automated approaches, this requires significant time. Hence, in order to send spam in bulk, spammers try to create or buy spambot that performs this task in an streamlined and automated fashion [26]. This works out well because most of Web 2.0 application uses generic templates which have the same format and data entry / validation requirements.

X-Rumer Palladium [18] is a tool that can be used for auto registration and posting spam on a forum, guestbook, wiki and other applications. This tool can be used to break most recent CAPTCHA and pass many antispam filters. With this tool, spammers can send spam automatically with a higher success rate.

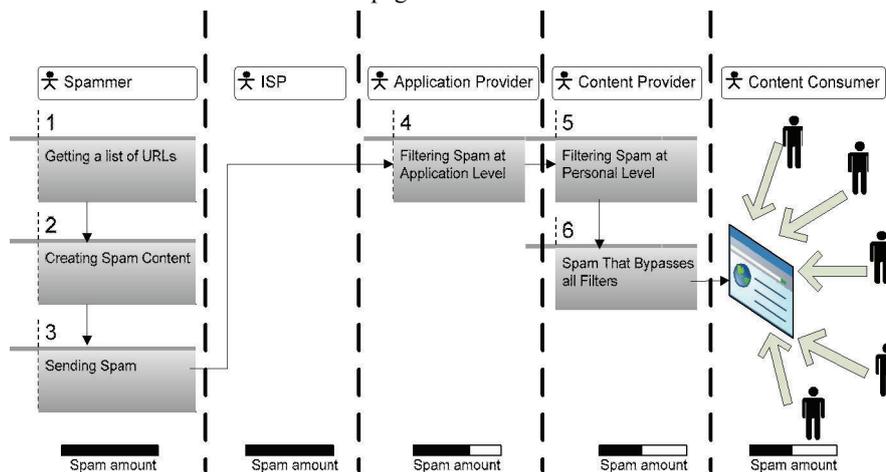


Figure 1. Lifecycle of web spam.

D. Filtering Spam at Application Level

This is the fourth stage in the Spam 2.0 lifecycle. Spam can be seen to be sent directly or indirectly. For instance, spam

messages can be sent directly to the users of a forum using private messaging. On the other hand, if it is a spam entry in a forum, then the forum's users can be considered to have

received spam indirectly. This works similarly for comments in blog and wiki. At the application level, web administrators hired by the application provider can install a number of antispam filters to quarantine possible spam content. Existing antispam filters apply blacklisting, whitelisting, keyword checking or other techniques to make an initial decision to quarantine.

To avoid spammers from easily sending spam to web applications, there are some tools that web administrators could use. For forums, TruBar and Anti-Spam Verification Questions for SMF are some of the antispam programs that can be applied [27, 28]. The latest version of phpBB has already embedded antispam filters including CAPTCHA into their package [29]. NoSpamNX, Typepad AntiSpam, AntispamBee, Trollguard Beta, WP Hashcash Plugin and WP-SpamFree are developed for Wordpress [30-35]. Other most applied antispam filter for web applications are Akismet and Defensio [36, 37]. Defensio supports various types of platforms other than Wordpress, such as AintaBlog, Drupal, Dotclear and Textcube [36]. Though these tools usually come with no cost for personal use, they usually require frequent updates and / or data for training.

E. Filtering Spam at Personal Level

While on this stage, the effort in eliminating spam is fully dependable on content provider's effort. The content provider has the appropriate permissions to add, edit and delete the spam content manually. Content providers can also report, delete, approve or even close the application. At this stage, there is almost nothing that spammers can do except to hope that the users would not delete spam content and somebody would fall for the trap by clicking on the link provided in the spam content. This is where the content provider plays an active and important role in managing their own application.

F. Spam that Bypasses all Filters

Web spam content usually consists of a spam message, followed by a link which will take a user to another site which generates revenue for the spammers. Spam that bypasses filters is likely to be seen by many users. If the link is clicked, then the search engine rank for that site linked with the spam will improve which is one the spammer's motivation for spamming.

At this stage, the targets of the web spammer are the content consumer. Content consumer have no access to edit or delete web spam, but they are able to view the content and possibly to report it to administrators. For forum content, consumers are the forum users. This applies similarly for blog and social networking applications. Meanwhile, for wikis, anyone could have the access to view and modify content.

IV. COST CATEGORIES OF SPAM 2.0

This section will show how the defined parameters are generally related to costs related to spam.

A. Defining Parameters

Based on previous research and several spam cost calculators that are currently available, there are several costs that can be calculated in order to estimate the price of email spam but no solutions currently exist that can calculate the cost of Spam 2.0. It is important to define the related costs that are measurable such as time and money. Hence, intangible cost

such as the users' level of annoyance in dealing spam is not included in the calculation.

Spam content for all type of applications have a basic unit. For email, the basic unit that is commonly used is the messages. We define spam content for all types of Web 2.0 applications as follows:

TABLE I. WEB 2.0 APPLICATION AND SPAM UNITS

Type of Web 2.0 Application	Forum	Wiki	Blog	Social networking
Spam Unit	post, poll, personal message, attachment	article, tag, reference	entry, comment, tag	post, comment, tag, personal message, user

Spam content will refer to the basic unit for each type of application. Further definitions of terms used in cost calculations and generic definitions for calculating each cost are defined in this section.

1) *Storage Cost* : Storage cost as explained in [8, 36] is the cost for "the storage space used to keep message". In the case of web spam, storage cost is the cost spent for server storage used to store any information such as list of URL to spam for spammers and blacklisted IP addresses for company and ISP and most of the time, storage used to store actual spam content. Parameters for storage cost function are generally defined as follow:

$$C_s = f(a, b, c, d) \quad (1)$$

where a = monthly storage cost/GB
 b = total amount of spam content/day
 c = total spam content size
 d = duration of storage before elimination

2) *Bandwidth Cost* : Bandwidth cost as explained in [8] is "the bandwidth taken to download the message". In our case, we define bandwidth cost as the cost used for connectivity. In this case, all parties are going to bear the cost of connectivity with different amount. Bandwidth cost function needs parameters as defined below:

$$C_b = f(e, f, g) \quad (2)$$

where e = connectivity cost
 f = type of application percentage representing bandwidth
 g = spam percentage of all types of applications

3) *Human Resource Cost (Annual Support Cost for Spam Filter)* : Human resource cost or annual support cost for spam filter is the cost used by the associated party for spamming or spam filtering. This cost can be defined as follow:

$$C_{hr} = f(h) \quad (3)$$

where h = salary for human resource incharge of support spam queries.

4) *Annual Productivity Cost* : Annual Productivity cost in usual cases would consider the recipient time to delete spam messages. In this case, annual productivity cost is defined as the cost calculated in order to identify the cost of time that the recipient of spam spent to combat spam. Parameters for this cost function are as follow:

$$C_{ap} = f(i,j,k,l) \quad (4)$$

where i = time to clear out spam content/each check,
 j = time to look for false positive for marked spam content/each check,
 k = time used to determine that it's a spam/each check,
 l = how many times users check/day.

5) *Software Cost* : Spammers or users usually rely on software or program to spam or for spam filtering. There is a lot of free open source software but sometimes it requires settings, knowledge and skills to be able to use them effectively. Therefore, it is easier to opt for software that is easy to use, easy to setup and most of them come with a price. This cost can be defined as follow:

$$C_{sw} = f(m) \quad (5)$$

where m = software costs.

Listed in Table II below are the parameters used in calculation for spammer, application provider, content provider and content consumer. Even though we are trying to define a generic definition for each cost calculation, there might still be some cost calculation that is not going to be applicable to certain party thus showing that parameters used vary depends on the cost calculation.

TABLE II. PARAMETERS USED FOR SPAMMER, APPLICATION PROVIDER, CONTENT PROVIDER AND CONTENT CONSUMER.

Parameter	Abb.	Spammer	Application Provider	Content Provider	Content Consumer
Storage Cost					
Monthly storage cost/GB	a	X	X	X	X
Total amount of spam content/day	b		X	X	X
Total spam content size	c	X	X	X	X
Duration of storage before elimination	d	X	X	X	X
Bandwidth Cost					
Annual fee connectivity cost	e	X		X	X

Type of application percentage representing bandwidth	f			X	X
Spam percentage of all type of applications	g			X	X
Human Resource Cost					
Salary for human resource incharge of supporting spam queries	h		X	X	
Annual Productivity Cost					
Time to clear out spam content/each check	i			X	X
Time to look for false positive for marked spam content/each check	j			X	X
Time used to determine that it's a spam/each check	k			X	X
How many times users check/day	l			X	X
Software Cost					
Software costs	m	X		X	

This section has explicitly defined 13 parameters used in cost calculations. These cost calculations will have different effects on each stakeholder that receive the spam. This will further be explained in the next section.

V. STAKEHOLDER'S COST

Fig. 2 below shows the cost impact of web spam towards six parties: spammer, ISP, application provider, content provider, content consumer and country. Lifecycle of web spam starts from the spammers' side and continues to ISP followed by the application provider, content provider and content consumer. As we have mentioned earlier, we are not going to focus on ISP side as it is unclear of how much ISP played their role in filtering web spam. In this research, we are only going to focus the cost impact of spam towards four stakeholders that we have introduced in Section II. Based on the generic parameters set in previous section, each cost associated for spammer, application provider, content provider and content consumer are going to be identified.

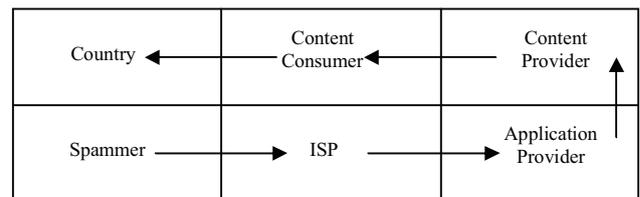


Figure 2. Web spam cost impacts.

A. Cost of Web Spam for Spammers

The fact that spammers also spend some amount of money [9] to spam questions the profit that they obtain from this activity. Hence, this section will further specify the related cost needed for spammers to spam in the web domain in each associated stage of lifecycle that we have introduced earlier.

The lifecycle of web spam starts with spammers as they gather a list of URLs. Spammers can use tools to collect a list of URLs such as Win Web Crawler, Web Data Extractor, Rafabot, Extract Link, Extract URL, Online Data Extractor,

Visual Web Spider and Teleport. These softwares range in price between AUD43 to AUD220. To keep the cost as low as possible, spammers could use freeware Elite Web Crawler, WebReaper, URL Spider Pro, Heritix and WebSphinx. Spammers need to have access to bandwidth and storage to keep a list of URLs. Storage and bandwidth cost is associated with stage 1 for spammers. Hence, parameters set for storage cost and bandwidth cost used by spammer are as follow:

$$C_s = f(a, c, d) \quad (6)$$

$$C_b = f(e) \quad (7)$$

Spammers could also find unsecured machines and use them to send spam. This could further reduce the cost for spammers. As can be seen, the cost that spammers have to bear are relatively small compared to other parties. In stage 2, spammers would have to generate spam content. Tools that can be used by spammers to generate spam content is SEO Text Generator which can be downloaded for free. However, there is an additional cost that spammers have to spend such as software costs. In order to send spam, X-Rumer 5.0 Palladium which costs AUD596. It has the capability to surpass different types of CAPCTHA. This software could also be used in all three stages of the lifecycle of web spam involved with spammers which are to find target for sending spam, create spam content followed by sending spam effectively.

B. Cost of Web Spam for Application Provider

In the effort of avoiding losing legitimate posts, it is easier to flag spam content so that it could be checked by the web administrator itself. Once the admin checks it, the admin would either read it and clear the content as non-spam or delete it if it is spam. This process requires additional storage and includes the cost of filtering because the efficiency of this method depends on the filter itself. Suppose if the email or content itself contains big attachment files or large images, this will increase the storage requirement and its cost. Storage cost is associated between stage 4 and 5 for users. Storage cost parameters can be defined as equation 1.

Application providers also play an active role in creating a better antispam filter. They create antispam plugins with better features in order to promote their services. We define the cost of creating a plugin as a human resource cost in equation 3. Nevertheless, there is a cost of deploying plugins that is used for commercialized purposes and this cost has to be paid by the content provider. This cost will further be explained in the next section.

C. Cost of Web Spam for Content Provider

Suppose if a company would like to open a forum or a blog, this company plays the role as a content provider. In this case, storage cost as defined in equation 1 has to be paid by the content provider. Storage cost is wasted for spam content. Hence, there is a need for someone to manage this forum or blog. Therefore, the company then needs to hire a web administrator to handle this.

Taking into account that not everybody has knowledge of what spam is, the administrator is hired to handle any upcoming issues from spam. This could include help-desk support or a team specially hired for fighting spam. In reality, this administrator is not only being paid their salary, they will also need to attend training for antispam technologies that is deployed for the web applications. This support cost for spam filters is associated with stage 5 in the case of lifecycle of web spam and can be defined as in equation 3.

Cost of bandwidth is clearly an important issue as bandwidth is wasted when used to download unnecessary spam content. Spam that is transmitted across the network consumes the bandwidth. It consumes a larger bandwidth capacity whenever the spam content embeds large images. As a consequence, users in a company have slower access to internet and slower download rates to more important tasks. Indirectly, users will take a longer time to finish a given task thus gives an impact to loss of productivity. This cost is associated during the transmission of web spam from spammer's side to user's side which is between stage 3 and 4 and this cost can be defined as in equation 2.

Annual productivity cost is measured for the time that is spent on each spam messages or spam content. This cost may vary depends on the user's knowledge and how well-formed spam content is. This cost is associated with stage 5 and it is calculated as in equation 4.

In stage 4 of the lifecycle where a web application is deployed with the spam filter, there is a dependency on software usage. Most of the plug-ins used in this stage are free for personal use, but consume money if used commercially. For instance, Mollom which is an antispam filter for blog, social network and community website is free if used for personal but costs AUD5860/year for each site if used commercially. Akismet on the other hand is charging AUD55 for filtering spam on a company's blog.

D. Cost of Web Spam for Content Consumer

Spam which is transmitted at the same time with legitimate content causes increase usage of network bandwidth and storage capacity. In order to obtain information from web applications, content consumer also bear the costs of waste storage used to download spam content and their bandwidth is also exhausted for this purpose as defined in equation 1 and 2.

Content consumers could also play an important role in eliminating spam that bypasses the filters. Even though they have no access to delete or edit any content on certain type of web applications, they still have the ability to determine spam content and report it to web administrator. This reduces productivity. This cost can similarly define as in equation 4.

VI. PROTOTYPE DEVELOPMENT OF HONEYSPAM

Using 5 cost categories involving 13 parameters that we have defined earlier, we are now going to develop honey spam that could estimate the cost of web spam. This section will first explain honey spam followed by detailed discussion on how we plan to measure each cost categories.

Towards determining the cost of spam, it is essential to first decide the number of web that is infected by spam. Honey

spam that we are developing contains a crawler engine, content extractor and evaluation engine, such as in Fig. 3. The crawler engine is going to be used to crawl to discover Web 2.0 applications. The data collected is going to be used as a reference which will be used in estimating the total amount of web spam.

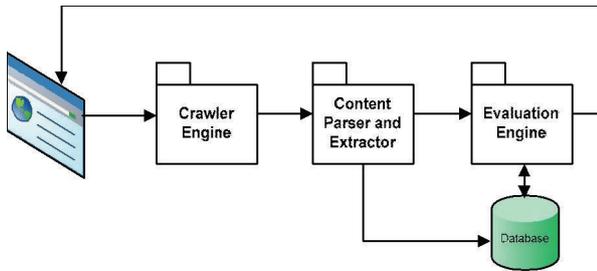


Figure 3. Prototype of HoneySpam Framework.

Meanwhile, the content extractor is going to be used to extract information that is available on a particular Web 2.0 application. The content extractor will parse HTML files that were downloaded earlier and extract valuable information.

This raw data will then be transformed to valuable information. This will then be passed to the evaluation engine which will decide whether the websites are spam-infected or not. The evaluation engine will contain a combination of several effective solutions to categorize a site into spam or not. This system will then report to the owner of the sites whether their site is spam infected or not. The evaluation engine will then be updated with new information.

Basically, based on the estimated amount of web spam for our honey spam above, we would be able to calculate the cost of spam. We will then be able to measure 13 parameters of five cost categories and this is going to be explained in the table below.

TABLE III. LIST OF PARAMETERS AND ITS MEASUREMENT.

Parameters	Abb	Measurement
Storage Cost		
Monthly storage cost/GB	<i>a</i>	Actual server cost paid for each GB
Total amount of spam content/day	<i>b</i>	Spam content received by all users in the certain duration of data collection recorded each day
Total spam content size	<i>c</i>	Actual size used to keep the content in storage
Duration of storage before elimination	<i>d</i>	Close observation towards the spam content requires the content of a specific type of application to be downloaded every day
Bandwidth Cost		
Annual fee connectivity cost	<i>e</i>	Current cost that users have to pay for the connectivity.
Type of application percentage representing bandwidth	<i>f</i>	Not decided yet.
Spam percentage of all type of	<i>g</i>	Storage that spam content uses compared to full downloaded data storage.

application		
Human Resource Cost		
Salary for human resource incharge of support queries	<i>h</i>	Current salary usually paid to the network administrator requires further survey on current situation in order to determine its precise value.
Annual Productivity Cost		
Time to clear out spam content/each check	<i>i</i>	Measurement for this cost depends on how fast a user can interact with system which also depends on how familiar users are with the application.
Time to look for false positives for marked spam content/each check	<i>j</i>	Measurement for this cost may vary depending on how knowledgeable users are. It is also possible to measure this based on author's experience.
Time used to determine that it's a spam/each check	<i>k</i>	Measurement for this cost has not been decided yet but it is also possible to measure this based on author's experience or several ongoing research.
How many times users check/day	<i>l</i>	It is possible to use a predetermined default value for this parameter.
Software Cost		
Software costs	<i>m</i>	Assuming that spammers would use software to spam, our calculation will consider the lowest cost software that could be used by spammers in each stage of lifecycle.

VII. ONGOING RESEARCH, DISCUSSION AND CONCLUSION

Sophos discovers one new infected webpage in every 3.6 seconds [38]. This statistic shows that even with all the technologies and methods that the anti spammers are using now; the spammers still could keep up with them [39]. It's still an "arm race" between these two parties. Unfortunately, while the race is going on between them, users are the ones who have to bear the consequences of this situation. In order to fight spammers, an individual has to spend their valuable time to check for spam content. As for the other parties, they have to prepare larger storage and spend extra on antispam technology.

The authors first identify the lifecycle of web spam and tools that can be used in completing each stage. We then listed the stakeholders involved in the webspam lifecycle. Afterwards, we identify five cost categories with their related parameters. Finally, we derive each stakeholder's cost based on the five cost categories. Considering that we are going to measure the cost of web spam accurately based on a large spam reference collection, there is a need to formulate all associated costs accordingly. It is important to take note that there are several key issues in calculating the cost of web spam.

A considerable amount of this report has been published on the cost of email spam. However, to the authors' knowledge there are no reports or studies on the cost of web spam. As we define Spam 2.0 cost, we noticed that some parameters can be easily defined and measured using our reference collection of downloaded data. Nevertheless, some parameters are highly dependable on current situation and need further survey to find the most acceptable value such as parameter *a* and *h*. In addition, there are some others that are not easily measured and are highly dependable on the user itself. For instance, parameters *i*, *j*, *k*, and *l* which could also be measured based on author's experience.

Moreover, some parameters depend on current technologies which may affect its price such as *a* and *m*. There is also a need

to collect data regularly/everyday hence a bigger collection of data would consume a bigger storage space such as parameter d which need a close observation towards determining its value. We are also aware of the resulting values for each attributes may vary depending on several factors such as popularity of the forum, thread, users/topic, number of posters and the administrator's effort.

It is obvious that ISPs are playing their role in filtering email spam. Therefore, it is interesting to find out how ISPs play their role in filtering web spam which requires further research in the future. By developing the HoneySpam and as the ongoing research is progressing, we would finally hope to develop a real time cost spam calculator based on the five cost categories and 13 parameters that we have defined earlier. It is believed that this cost calculator could provide a better overview of how serious the web spam problem is.

REFERENCES

- Courmane, A. and R. Hunt, An analysis of the tools used for the generation and prevention of spam. *Computers & Security*, 2004. 23(2): p. 154-166.
- Spamhaus. The Definition of Spam. [cited 2010 25 January]; Available from: <http://www.spamhaus.org/definition.html>.
- Hayati, P. and V. Potdar, Toward spam 2.0: an evaluation of web 2.0 anti-spam methods, in 7th IEEE International Conference on Industrial Informatics (INDIN 2009) 2009: Cardiff, Wales.
- Hayati, P. and V. Potdar, Evaluation of spam detection and prevention frameworks for email and image spam: a state of art, in *Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services*. 2008, ACM: Linz, Austria.
- Lin, Y.-R., et al., Detecting splogs via temporal dynamics using self-similarity analysis. *ACM Transactions on the Web (TWEB)*, 2008. 2(1).
- Ahn, L.v., M. Blum, and J. Langford, Telling humans and computers apart automatically. *Communications of the ACM*, 2004. 47(2 (February 2004)): p. 56-60.
- Yan, J. and A.S.E. Ahmad. A low-cost attack on a Microsoft captcha. in *Conference on Computer and Communications Security*. 2008. Alexandria, Virginia, USA: Proceedings of the 15th ACM conference on Computer and communications security.
- Su, C.-c. An Open Source Portal for Educators. *TESL-EJ* 9,1 2005 [cited 2009 23 November]; Available from: <http://tesl-ej.org/ej33/int.html>.
- Spammer-X, J. Posluns, and S. Sjouwerman, *Inside the SPAM Cartel: By Spammer-X*. 2004: Syngress.
- Gyongyi, Z. and H.G. Molina, Web spam taxonomy, in *1st International Workshop on Adversarial Information Retrieval on the Web*. 2005.
- Provos, N., M.A. Rajab, and P.i.M. ommatis, *Cybercrime 2.0 : when the cloud turns dark*. *Communication of the ACM*. Vol. 52. 2009. 42-47.
- Win Web Crawler - Powerful webcrawler, web spider, website extractor. [cited 2009 23 November]; Available from: <http://www.winwebcrawler.com/>.
- Web Data Extractor - Extract URL, Meta Tag, Email, Phone, Fax from Web. [cited 2009 23 November]; Available from: <http://www.webextractor.com/>.
- RafaBot - Download Websites! Bulk website downloading and spidering software [cited 2009 23 November]; Available from: <http://www.spadixbd.com/rafabot/>.
- Offline Email Extractor/Link Extractor [cited 2009 23 November]; Available from: <http://www.spadixbd.com/elink/>.
- Online Data Extractor - Extract Email, Phone, Fax, URL, Meta Tag from website, search engine. [cited 2009 23 November]; Available from: <http://www.onlinedataextractor.com/>.
- Visual Web Spider, Website Crawler, Web Robot, Website Ripper. [cited 2009 23 November]; Available from: http://www.newprosoft.com/web_spider.htm.
- Botmaster.net. Botmaster.net : autosubmitter's XRumer description. 2009 [cited 2009 13 November]; Available from: <http://www.botmasternet.com/more1/>.
- Tenmax.com. Teleport Pro - Offline Browsing Webspider. [cited 2009 13 November]; Available from: <http://www.tenmax.com/teleport/pro/home.htm>.
- Sourceforge. WebSphinx. 2009 [cited 2009 9 November]; Available from: <http://sourceforge.net/projects/websphinx/>.
- Heritrix. 16 October 2009 [cited 2009 9 November]; Available from: <http://crawler.archive.org/>.
- Brothersoft. URL Spider Pro 3.3.3. [cited 2009 9 November]; Available from: <http://www.brothersoft.com/url-spider-pro-5727.html>.
- Otway, M. WebReaper. [cited 2009 9 November 2009]; Available from: <http://www.webreaper.net/download.html>.
- Brothersoft. Elite Web Crawler. [cited 2009 9 November]; Available from: <http://www.brothersoft.com/elite-web-crawler-298613.html>.
- Nagamalai, D., B.C. Dhinakaran, and J.-K. Lee, An in-depth analysis of spam and spammers. 2009.
- Hayati, P., et al., HoneySpam 2.0: Profiling Web Spambot Behaviour. , in *PRIMA 2009*. 2009: Nagoya, Japan.
- PHP-Nuke. PHP-Nuke - TruBar 4.0(Silent) - anti-spam MOD. 2009 [cited 2009 13 November]; Available from: <http://phpnuke.org/modules.php?name=News&file=article&sid=8244>.
- Simple Machines. Anti-Spam Verification Questions for SMF 1.1.7 2009 [cited 2009 13 November]; Available from: <http://custom.simplemachines.org/mods/index.php?mod=1516>.
- phpBB. phpBB Features. 2009 [cited 2009 13 November]; Available from: <http://www.phpbb.com/about/features/?from=submenu>.
- Kubiak, S. Plugin Directory. 2009 [cited 2009 13 November]; Available from: <http://wordpress.org/extend/plugins/nospamnx/>.
- TypePad Antispam. 2009 [cited 2009 13 November]; Available from: <http://antispam.typepad.com/>.
- Müller, S. Plugin Directory. 2009 [cited 2009 13 November]; Available from: <http://wordpress.org/extend/plugins/antispam-bee/>.
- Trollguard Beta. 2009 [cited 2009 13 November]; Available from: <http://www.trollguard.com/>.
- WP Hashcash Plugin for Spam by Wordpress Plugins. 2009 [cited 2009 13 November]; Available from: <http://wordpress-plugins.feifei.us/hashcash/>.
- Allen, S. Plugin Directory. 2009 [cited 2009 13 November]; Available from: <http://wordpress.org/extend/plugins/wp-spamfree/>.
- Defensio - Wordpress Anti Spam Plugin. [cited 2009 9 November]; Available from: <http://defensio.com/downloads>.
- Akismet. [cited 2009 13 November]; Available from: <http://akismet.com/>.
- SOPHOS, Security threat report: July 2009 update -A look at the challenges ahead. July 2009.
- Ferris Research, Spam, Spammers and Spam Control. March 2009, Ferris Research: San Francisco, Calif, USA.