Module:	LD7028 – Research and Project Management
Module tutor:	Hamid Jahankhani / Shelagh Keogh
Assignment title:	MSc Research Proposal
Student name:	
Student university identifier:	
Course of study:	MSc Cyber Security
Supervisor:	
Second marker:	<TBC>
Review date:	<TBC>
Word count:	2753 Excludes cover sheet, contents, tables and references
Word count:	4464 Includes cover sheet, contents, tables and references

Table of Contents

Table of Figures

Table of Tables

Title

An approach to utilising Machine Learning to personalise spear phishing simulations to change individual behaviours.

Aim

Produce personalised spear phishing awareness from information within the end users email environment. To produce an algorithm or method to analyse mailbox data to this effect. Enhance the Human Firewall by changing default behaviours towards email risks.

Problem Statement

When an email is received the recipient needs to be instinctively checking the validity of the message. Recipients should be capable of finding and understanding key cues for legitimate and malicious messages.

However, recipients do not regularly check for authenticity of emails (Google, 2017), only doing so when the message is blatantly out of the ordinary. As most messages can be trusted there is a feel-good factor to responding quickly to received emails. Recipients are not vigilant to the risks of responding or clicking malicious messages or are overconfident in the ability of security controls to mitigate the threat posed. It only takes one successful malicious email to potentially expose personal or corporate information or systems. It is believed that the majority of cyber incidents are aided by unintentional insider assistance (Mangelsdorf, 2017). Traditional phishing simulations are run against a population rather than individuals resulting in a generalised, less personal campaign.

A successful malicious email can result in multimillion pound losses in remediation, regulatory action, civil lawsuits (Osborne, 2017) or direct fraud (Justice.gov, 2017). Market and customer confidence can be impacted (Rosati et al., 2017) resulting in diminished investment and growth opportunities. For the affected recipient, personal information may be leaked resulting in reputation loss (Kleinman, 2017), or financial loss (IT Governance Ltd, 2016), or identity theft.

Background

In February 2017 the researcher witnessed three cyclists in London running red traffic lights during a journey of a little over a mile. These risks were taken even though three cyclists were killed in the city the previous week (BBC News, 2017).

Having worked as a system administrator and security analyst for many years the researcher has noticed that human behaviour has not fundamentally changed with respect to risky decisions with information technology either, such as clicking on links in email. Conversely, the threat, skill and cunning of potential attackers have advanced and now prey directly on ingrained end user risky behaviours (Mansfield-Devine, 2017).

The researcher has had an interest in email security and its limitations for many years, especially around spear phishing. According to Lastdrager (2014) “phishing is a scalable act of deception whereby impersonation is used to obtain information from a target”. Spear phishing targets more selective groups of users often with defined responsibilities such as finance, technological, healthcare, or human resources access (Garretson, 2007a). Whaling, which dates to at least 2007 (Garretson, 2007b), is an even finer defined attack that targets high profile individuals; c-level executives (Payne, 2017).

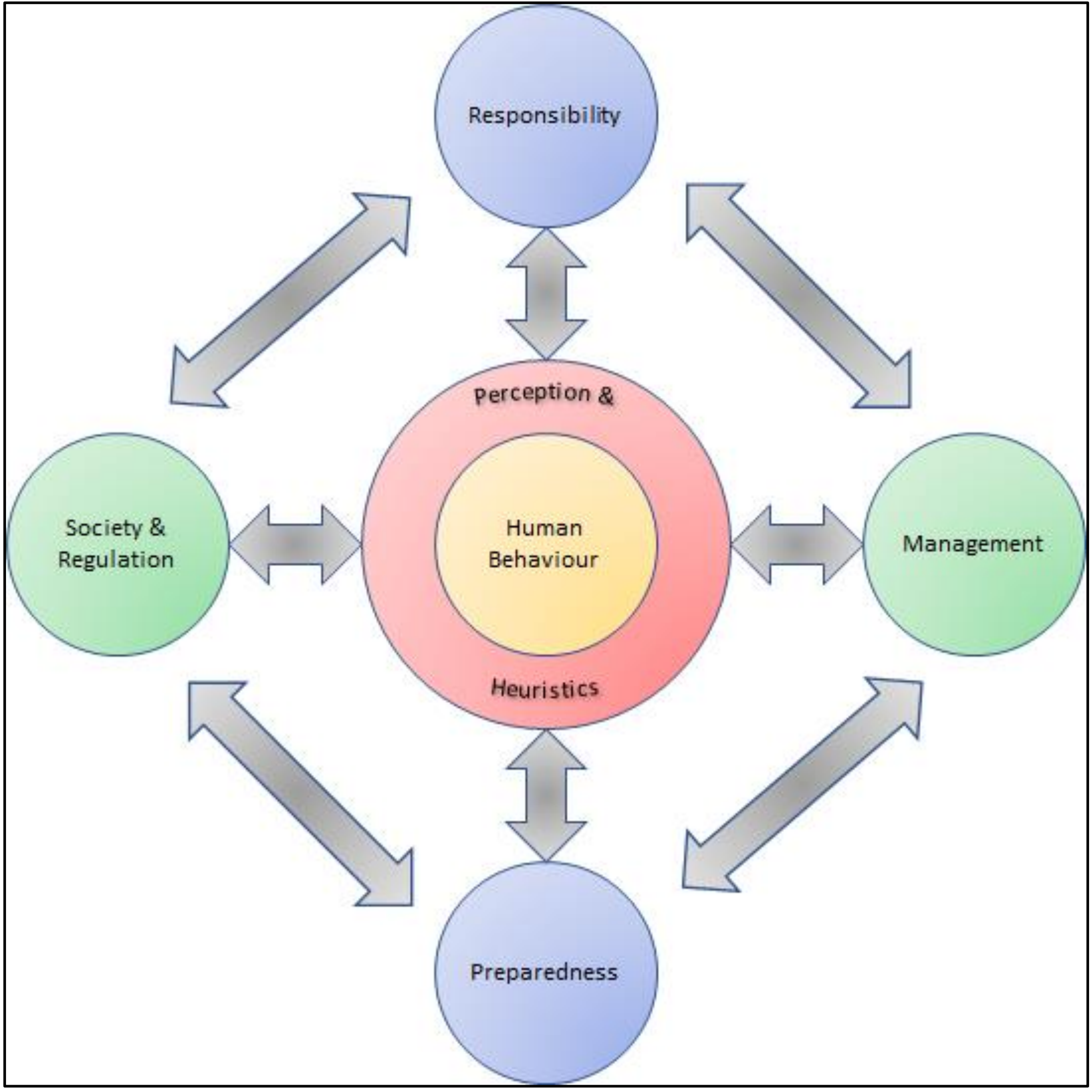
Because a malicious spear phishing campaign has been personalised to the individual victim it can be much harder to distinguish from legitimate messages which the individual opens without thought every day (Newman, 2017).

Any information security sphere can be weakened by poor decision-making by end users. Whilst security controls can be effective against many threats they can leave the end user unaware of the attacks that have been automatically mitigated, resulting in a low understanding of the underlying threat. Consequently, it has become easier to attack the human than bypass or breach controls such as firewalls, antivirus and antimalware (Hong, 2012). The favoured mechanism to fill the awareness gap is user education programmes.

Education programmes are structured around tick box training programmes, gamification (Arachchilage, Love and Beznosov, 2016), and scheduled simulations all of which are highly generalised and target groups of users rather than individuals. This fits into the preparedness domain proposed by AlHogail (2015) in the Human Factor Diamond framework. This model does not consider end user perception at the point of making a decision. Figure 1 shows how perception and unconscious decisions can override logical decisions based on education or responsibilities. One approach to improve on this framework would be to incorporate elements of end user psychology.

Pfleeger and Caputo (2012) surveyed multiple psychological aspects of human behaviour and how they may be applied with information security. One approach to improving an individual’s resistance to social engineering attacks is to improve the available metrics in an easily consumed manner, such the work done by Lötter and Futcher (2015). An appropriate, psychological, approach would appear to be to influence the availability heuristic. The availability heuristic combines how memorable an event is with the regularity of that event (Colman, 2009). By looking at how the user is responding to all email an awareness campaign can be run on an individual basis. By highlighting and rewarding secure, thoughtful conduct we can convince users to take fewer risks. Changing this behaviour requires the end users to start instinctively deducing when they must check the validating of messages (Best, 2009), overriding their heuristic inclinations, which may well be wrong.

Figure 1: Adjusted influences on human behaviour in information security (Pfleeger and Caputo, 2012).



A personalised phishing simulation campaign may create a resistance to personalised, highly targeted attacks such as spear phishing and whaling.

It is proposed that the techniques used in this research will provide a basis for a new style of anti-phishing education or awareness product that can be integrated into large, corporate mail systems, either intrinsically or as a plug-in. Such a system could create a human firewall making successful spear phishing far more unlikely.

The primary objective is to deliver a method or algorithm for analysing raw mailbox data in order to define or produce personalised spear phishing campaigns. Although several papers have been written about detecting spear phishing emails (Akinyelu and Adewumi, 2014; Daeef et al., 2016; Nizamani et al., 2014; Olivo, Santin and Oliveira, 2013) or phishing websites (Gowtham and Krishnamurthi, 2014; Kazemian and Ahmed, 2015; Lakshmi and Vijaya, 2012) utilising Machine Learning the novel approach taken here is to generate a new tailored simulation.

This is an enhancement to and can be run side by side with traditional education techniques such as the “Stop. Think. Connect.” campaign (2016) started cooperatively by the Anti-Phishing Working Group and National Cyber Security Alliance launched in 2010.

Scope

- Research will concentrate on one small but high visibility security sphere; spear phishing.
- The project will only be investigating English language spear phishing simulations.
- The primary data will be obtained from a single organisation.
- The primary data will be taken from a single mail service, Google Mail.
- Integration into existing phishing simulation tools or mail environments is excluded from the scope of this project.
- Automation of the generation of the simulations is excluded from the scope of this project.

Korpela (2015) discusses how awareness and education campaigns fail to identify individuals at highest risk to ensure effective application of resources. This research will not attempt to answer this question but, rather, provide an alternative education mechanism.

Objectives

- Deliver a framework for running a personalised phishing awareness campaign.
- Investigate and document secondary data sources on active or recent real world phishing campaigns.
- Review available Machine Learning systems and associated algorithms for interpreting email headers and bodies.
- Develop new algorithm or combination of trained algorithms to create targeted phishing campaigns.
- Survey participants on existing knowledge of email security and phishing awareness prior to participation.
- Survey participants on existing knowledge of email security and phishing awareness after participation.
- Analysis of improvements in ability to identify phishing emails.

Methodology

The work by Jingguo, Yuan and Rao (2016) discusses how checking the validity of a message is dependent on the recipient’s confidence in their ability to distinguish between valid and malicious messages. Jingguo et al further discuss the impact of perception and cognitive effort on overconfidence. This can lead to incorrect decision-making processes with respect to identifying phishing emails. By making the simulation more personal and less easily detected as a test we hope to affect the long-term heuristic or instinctive process that may be too trusting of emails without checking voracity. To do this the participant must start to make active decisions and not active instinctively.

A questionnaire will be designed to understand the existing awareness and email usage patterns of the population. A questionnaire has been selected in order to ascertain the skill of the participants in identifying real world phishing emails. This questionnaire can be combined with authorisation/consent to examine and process email on behalf of the participant for the purposes of the study.

A subset of the participants will be used to train or generate Machine Learning algorithms. In the first editorial in the journal Machine Learning, Langley (1986) discusses a definition of such a system, as distinct from Artificial Intelligence or Pattern Recognition systems. In this project Machine Learning is being used as a tool to generate knowledge from a significant amount of empirical data.

Machine Learning will be utilised to analyse information from each participant's mailbox and suggest a personalised phishing simulation campaign.

The individualised campaigns will be run over a series of weeks. Susceptibility will be tracked and immediate feedback given both interactively and by email response. The end users email platform will be tracked to provide the most accurate, personalised phishing cues available within the education provided. Platform cues are defined as information available within a mail client that can assist a user in deciding if a message is legitimate or malign, safe or dangerous. A personalised report will be provided to each participant at the end of the campaign.

The original questionnaire, with additional questions to test awareness of platform cues, will be requested of the participants at the end of the campaign.

Subject matter experts may be requested to give feedback on the algorithm and process used.

The methodology comprises two analytical components. A quantitative measure may be drawn from the trends in susceptibility to phishing campaigns generated during the study. A qualitative measure can be drawn from both expert feedback and answers from the final survey.

Project Plan

Table 1 shows the tasks that are required to complete this project on the schedule given (Keogh and Jahankhani, 2017). These have been used to populate the Gantt chart shown in Figure 2, Figure 3, and Figure 4.

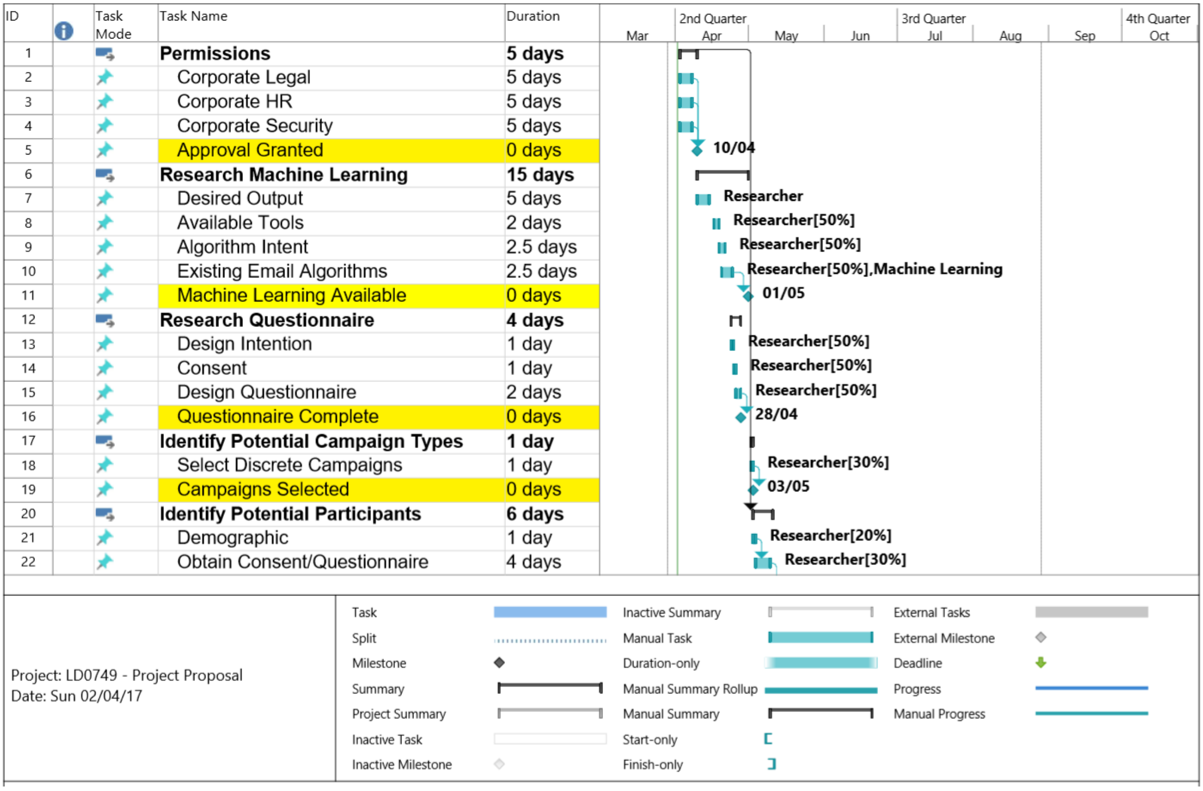
Table 1: Project tasks

Task No.	**Task**	**Effort**	**Start**	**End**	**Predecessors**	**Resources**
1	**Permissions**	5 days	03/04/17	10/04/17		
2	Corporate Legal	5 days	03/04/17	07/04/17		

Task No.	Task	Effort	Start	End	Predecessors	Resources
3	Corporate HR	5 days	03/04/17	07/04/17		
4	Corporate Security	5 days	03/04/17	07/04/17		
5	Approval Granted		10/04/17	10/04/17	4,2,3	
6	**Research Machine Learning**	15 days	10/04/17	01/05/17		
7	Desired Output	5 days	10/04/17	14/04/17		
8	Available Tools	2 days	17/04/17	18/04/17		Researcher [50%]
9	Algorithm Intent	2.5 days	19/04/17	21/04/17		Researcher [50%]
10	Existing Email Algorithms	2.5 days	20/04/17	24/04/17		Researcher [50%], Machine Learning
11	Machine Learning Available		01/05/17	01/05/17	10	
12	**Research Questionnaire**	4 days	24/04/17	28/04/17		
13	Design Intention	1 day	24/04/17	24/04/17		Researcher [50%]
14	Consent	1 day	25/04/17	25/04/17		Researcher [50%]
15	Design Questionnaire	2 days	26/04/17	27/04/17		Researcher [50%]
16	Questionnaire Complete		28/04/17	28/04/17	15	
17	**Identify Potential Campaign Types**	1 day	02/05/17	03/05/17		
18	Select Discrete Campaigns	1 day	02/05/17	02/05/17		Researcher [30%]
19	Campaigns Selected		03/05/17	03/05/17	18	
20	**Identify Potential Participants**	6 days	03/05/17	11/05/17	1	
21	Demographic	1 day	03/05/17	03/05/17		Researcher [20%]
22	Obtain Consent / Questionnaire	4 days	04/05/17	09/05/17	21	Researcher [30%]
23	Possible Extensions	1 day	10/05/17	10/05/17	22	Researcher [50%]
24	Participants Selected and Contacted		11/05/17	11/05/17	23	
25	**Email Access Mechanism**	3 days	01/05/17	04/05/17		
26	Protective Measures for Data	1 day	01/05/17	01/05/17		Researcher [30%]
27	Search Criteria	2 days	01/05/17	02/05/17		Researcher [30%]
28	Possible Extensions	1 day	03/05/17	03/05/17		Researcher [10%]
29	Data Access Designed		04/05/17	04/05/17	28	
30	**Proof of Concept**	5 days	01/05/17	08/05/17		
31	PoC Analysis	5 days	01/05/17	05/05/17		Researcher [10%]
32	Review Results	5 days	01/05/17	05/05/17		Researcher [10%]
33	Adjust Algorithm	5 days	01/05/17	05/05/17		Researcher [20%]

Task No.	Task	Effort	Start	End	Predecessors	Resources
34	PoC Complete		08/05/17	08/05/17	33	
35	**Trial Campaign**	15 days	08/05/17	29/05/17		
36	Analyse Mailboxes	5 days	08/05/17	12/05/17		Researcher [20%], Machine Learning
37	Generate Campaign	5 days	08/05/17	12/05/17		Researcher [30%]
38	Run Campaign	5 days	15/05/17	19/05/17		Researcher, Phishing Tool, Participants
39	Analyse Results	5 days	22/05/17	26/05/17		Researcher
40	Trial Complete		29/05/17	29/05/17	39	
41	**Campaign**	25 days	29/05/17	03/07/17	35	
42	Analyse Mailboxes	5 days	29/05/17	02/06/17		Researcher [50%], Machine Learning
43	Generate Campaign	5 days	29/05/17	02/06/17		Researcher [50%]
44	Run Campaign	20 days	05/06/17	30/06/17	42,43	Researcher, Participants, Phishing Tool
45	Campaign Concluded		03/07/17	03/07/17	44	
46	**Termination Research Questionnaire**	5 days	03/07/17	07/07/17		
47	Analysis	10 days	10/07/17	24/07/17	46	
48	Individual Perspectives	5 days	10/07/17	14/07/17		Researcher [30%]
49	Corporate Perspectives	5 days	10/07/17	14/07/17		Researcher [30%]
50	MSc In-Depth Analysis	10 days	10/07/17	21/07/17		Researcher [30%]
51	Analysis Concluded		24/07/17	24/07/17	50	
52	**Reports**	20 days	31/07/17	28/08/17	47	
53	Review Analysis	5 days	31/07/17	04/08/17		Researcher [30%]
54	Conclusions	10 days	31/07/17	11/08/17		Researcher [30%]
55	MSc Dissertation	20 days	31/07/17	25/08/17		Researcher [30%]
56	MSc Dissertation Concluded		28/08/17	28/08/17	55	

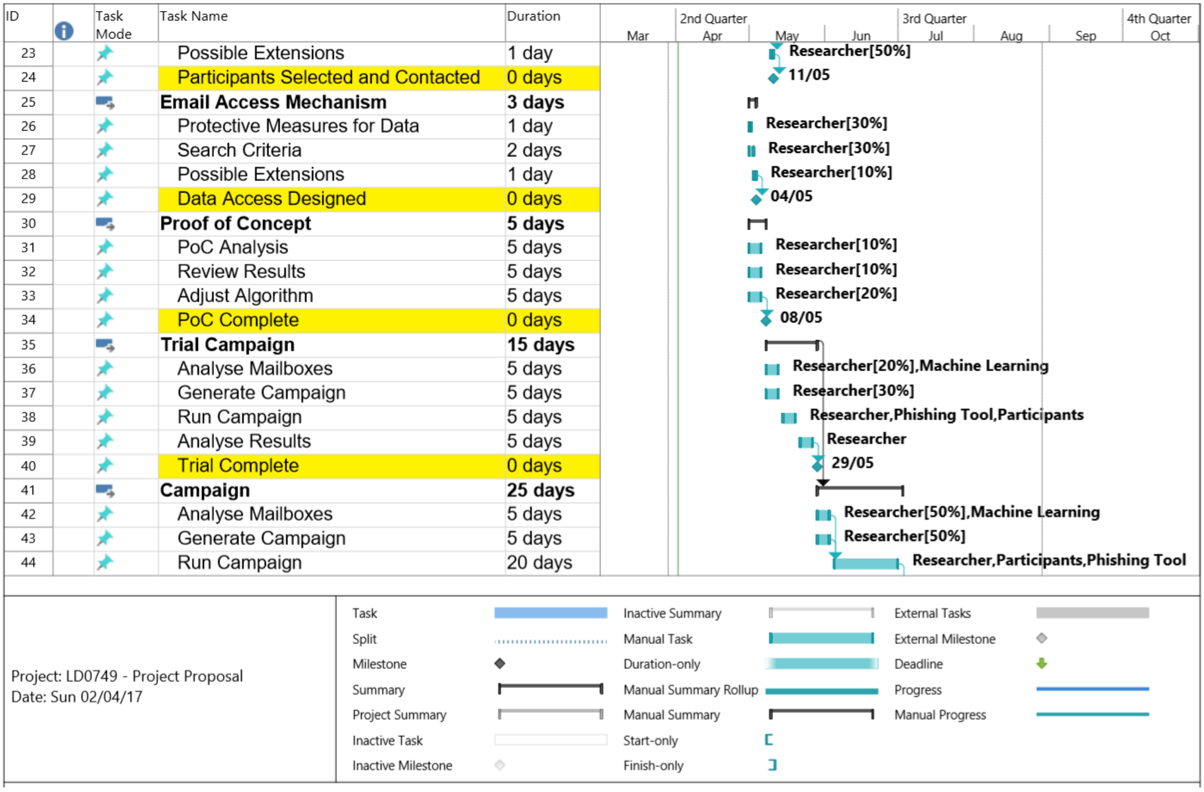
Figure 2: Project Gantt chart.



ID	ⓘ	Task Mode	Task Name	Duration
1		⇒	**Permissions**	**5 days**
2		📌	Corporate Legal	5 days
3		📌	Corporate HR	5 days
4		📌	Corporate Security	5 days
5		📌	Approval Granted	0 days
6		⇒	**Research Machine Learning**	**15 days**
7		📌	Desired Output	5 days
8		📌	Available Tools	2 days
9		📌	Algorithm Intent	2.5 days
10		📌	Existing Email Algorithms	2.5 days
11		📌	Machine Learning Available	0 days
12		⇒	**Research Questionnaire**	**4 days**
13		📌	Design Intention	1 day
14		📌	Consent	1 day
15		📌	Design Questionnaire	2 days
16		📌	Questionnaire Complete	0 days
17		⇒	**Identify Potential Campaign Types**	**1 day**
18		📌	Select Discrete Campaigns	1 day
19		📌	Campaigns Selected	0 days
20		⇒	**Identify Potential Participants**	**6 days**
21		📌	Demographic	1 day
22		📌	Obtain Consent/Questionnaire	4 days

Project: LD0749 - Project Proposal
Date: Sun 02/04/17

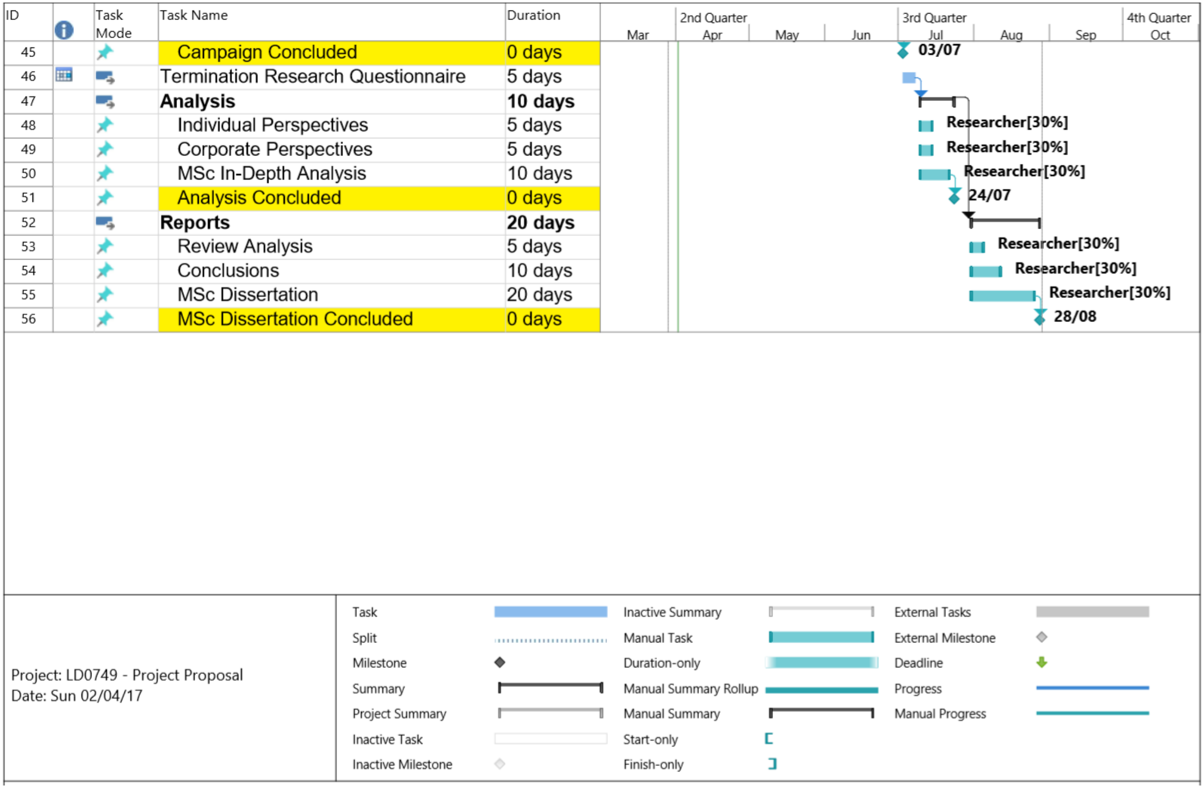
Figure 3: Project Gantt chart (continued).



ID	ⓘ	Task Mode	Task Name	Duration
23		📌	Possible Extensions	1 day
24		📌	Participants Selected and Contacted	0 days
25		⇒	**Email Access Mechanism**	**3 days**
26		📌	Protective Measures for Data	1 day
27		📌	Search Criteria	2 days
28		📌	Possible Extensions	1 day
29		📌	Data Access Designed	0 days
30		⇒	**Proof of Concept**	**5 days**
31		📌	PoC Analysis	5 days
32		📌	Review Results	5 days
33		📌	Adjust Algorithm	5 days
34		📌	PoC Complete	0 days
35		⇒	**Trial Campaign**	**15 days**
36		📌	Analyse Mailboxes	5 days
37		📌	Generate Campaign	5 days
38		📌	Run Campaign	5 days
39		📌	Analyse Results	5 days
40		📌	Trial Complete	0 days
41		⇒	**Campaign**	**25 days**
42		📌	Analyse Mailboxes	5 days
43		📌	Generate Campaign	5 days
44		📌	Run Campaign	20 days

Project: LD0749 - Project Proposal
Date: Sun 02/04/17

Figure 4: Project Gantt chart (continued).



ID		Task Mode	Task Name	Duration
45			Campaign Concluded	0 days
46			Termination Research Questionnaire	5 days
47			**Analysis**	**10 days**
48			Individual Perspectives	5 days
49			Corporate Perspectives	5 days
50			MSc In-Depth Analysis	10 days
51			Analysis Concluded	0 days
52			**Reports**	**20 days**
53			Review Analysis	5 days
54			Conclusions	10 days
55			MSc Dissertation	20 days
56			MSc Dissertation Concluded	0 days

Project: LD0749 - Project Proposal
Date: Sun 02/04/17

Known Risks

Several risks to the project have been identified and outlined in Table 2.

Table 2: Identified project risks

Risk Type	Risk Event	Likelihood (1-10)	Impact (1-10)	Risk Value (1-100)	Risk Monitoring / Control Flag	Risk Management Strategy	Risk Review date
P	Availability of researcher	7	8	56	🟧	Schedule time and review regularly	29/Apr/2017
S	Permission to use work systems denied	3	9	27	🟦	Reach out early to HR and management	22/Apr/2017
T	Loss of data due to hardware failure	1	9	9	🟩	Multiple backups to be put in place	29/Apr/2017

Risk Type	Risk Event	Likelihood (1-10)	Impact (1-10)	Risk Value (1-100)	Risk Monitoring / Control Flag	Risk Management Strategy	Risk Review date
S	Permission to use personal data refused	5	8	40	🟦	Consent always sought Extend outreach to alternative sources	22/Apr/2017
P	Ability to design suitable questionnaire untested	8	7	56	🟨	Review previous questionnaires on similar topic for guidance	29/Apr/2017
T	Communication outage prevents working on project	2	7	14	🟩	Mitigated by multiple connectivity options	22/Apr/2017
S	Data breach affects participants	1	10	10	🟩	No personal information to be kept unencrypted	13/May/2017
T	Permission to use work resources	2	10	20	🟩	Reach out early to HR and management	22/Apr/2017
B	Closed population may result in biased outcome	7	8	56	🟨	Design questionnaire to capture bias Define and declare potential biases within analysis	27/May/2017
B	Population may be too small to be meaningful	3	8	24	🟩	Extend outreach to alternative sources	27/May/2017

Risk Type	Risk Event	Likelihood (1-10)	Impact (1-10)	Risk Value (1-100)	Risk Monitoring / Control Flag	Risk Management Strategy	Risk Review date
P	Participants may be reluctant to take part	2	8	16	🟩	Reassure participants of confidentiality	22/Apr/2017
F	Machine Learning may require financial investment	6	5	30	🟨	A small contingency has been made available	27/May/2017

Key:

- Risk types

F	Financial
T	Technology
P	People
S	Security
B	Bias

- Control Flags

≥ 75	***High risk***	- Action required urgently
≥ 50 < 75	***Medium risk***	- Action as soon as possible
≥ 25 < 50	***Acceptable risk***	- Analyse issues raised
< 25	***Low risk***	- No gain expected from mitigation

The risk will be re-evaluated on the dates indicated then new review dates will be ascertained. Impact and likelihood may change during the course of the project.

The availability of the researcher, which must be carefully monitored against the given milestones, is of significant concern and deliverables and scope have been defined to take this into account. To this end, simple Kanban style cards for project task-tracking and notification via the Kanbana (Version 1.9.4) iOS application. Kanbana will be populated from the task list and Gantt chart and more finely detailed tasks added as they become apparent.

In the event of complete loss of communication or technological services it is possible to travel to University of Northumbria's campus at Middlesex Street, London or target organisation head office locations.

Consideration is being given to utilising work resources. Preliminary permission has been granted by the Human Resources department (personal communication, 15 Mar. 2017), Group Legal Counsel (personal communication, 22 Mar. 2017) and the

Global Head of Information Security at the organisation (personal communication, 15 Mar. 2017).

As the participants are colleagues within a single organisation they may feel reluctant to participate on the grounds of personal or professional privacy. Efforts will be made to reassure participants that no personal or identity information will be published.

Source and Use of Knowledge

Several journals that have covered the problem of human behaviour in relation to phishing attacks are of interest to the researcher. “Computers in human Behavior”, editor M. Guitton, publishes research on computing focused on psychological approaches. Because the journal concentrates on theories of psychology as the starting point of research (Journals.elsevier.com, 2017) it is unlikely that a paper based on the proposed research would be accepted by this journal.
The “Information Security Journal: A Global Perspective”, edited by Jill Slay, has a computer security industry focus may be a more appropriate journal to consider (Isc2.org, 2016). The researcher is already a member of the International Information System Security Certification Consortium, commonly referred to as (ISC)^2, and the journal actively encourages the members that are engaging research at Masters level to submit papers. It would be appropriate to consider an article generated from this research as potentially reducing security risk and raising security awareness within organisations and the article may be viewed more favourably if directed in this manner. Taylor & Francis (Tandfonline.com, 2017) publish the journal on behalf of (ISC)^2. Articles must be original and be between three thousand and six thousand words. All tables and figures are considered to be equivalent to three hundred words.

Other journals that are also being considered as appropriate for this research are:

- International Journal of Cyber-Security and Digital Forensics
- Computers & Security
- Computer Fraud & Security
- Computer Journal
- International Journal of Human-Computer Studies

The author and researcher Steve Mansfield-Devine has published several papers on spear phishing and is also the editor of two journals of interest, Network Security and Computer Fraud & Security. It would be interesting both to watch Mansfield-Devin’s output in this area and discuss the research directly, though Mansfield-Devin does write and research other information security domains.

Ethical, Legal, Social, Security and Professional Concerns

It is important to clearly identify the scope of the data and what it is to be used for to all participants prior to the collection of such data. That data must only be used for the purposes of this projected and destroyed after a reasonable period has elapsed. To this end all participants must be resident within the United Kingdom and British nationals. Under both the British Data Protection Act (Legislation.gov.uk, 1998) and the European General Data Protection Regulation (Publications.europa.eu, 2016) the Information Commissioner's Office (ICO) will be considered the appropriate Supervisory Authority (Ico.org.uk, n.d.).

As research is directed at analysis of an individual mailing habits over a discrete period care must be taken to protect both that resource and the connectivity used. It may be appropriate to hold both the data and any system utilised within a protected network, such as that provided by the target corporation. It is suitable and appropriate to use encryption at rest and in transit for the raw data obtained from participants. At the conclusion of the project all primary or processed data will be destroyed. Any data exported out of this environment will be anonymised and summarised accordingly.

As the research may be directly employed by the target organisation there is a potential conflict of interest between the needs of the corporation, the university and keeping the research independent and objective. To this end the research will adhere to the ethical constraint appropriate to an MSc project. The same entities also have clashing claims on Intellectual Property that may arise from the project that needs to be amicably resolved. The project proposal is to deliver a method or algorithm for producing customised phishing campaigns and does not clash with the target organisation's standard operations or business sectors.

The ethics of constructing realistic or even reconstructing real-world phishing simulations needs to be examined. Trademark and copyright law must not be impinged which inhibits the realism of simulated attacks. Simulation providers such as PhishMe (n.d.) provide clauses stipulating that intellectual property rights must be respected. Efforts must be made, therefore, to generate non-infringing simulations.

Consideration must be made about how to incentivise the campaign to prevent it becoming, or participants feeling like it has become, a blaming, shaming or finger-pointing process. Anonymization partially assists but a change of culture will not occur without such moral questions being assuaged. It is suggested that perhaps gamification of phishing reporting may be a step in the right direction. Some organisations have "Star Awards" (personal communication, 2017) and this may be a step towards making participation a positive process.

References

Akinyelu, A. and Adewumi, A. (2014). Classification of Phishing Email Using Random Forest Machine Learning Technique. *Journal of Applied Mathematics*, 2014, pp.1-6.

AlHogail, A. (2015). Design and validation of information security culture framework. *Computers in Human Behavior*, 49, pp.567-575.

Arachchilage, N., Love, S. and Beznosov, K. (2016). Phishing threat avoidance behaviour: An empirical investigation. *Computers in Human Behavior*, 60, pp.185-197.

BBC News. (2017). *Third cyclist killed in four days on London's roads - BBC News*. [online] Available at: http://www.bbc.co.uk/news/uk-england-london-38923734 [Accessed 24 Mar. 2017].

Best, J. (2009). Need to Override Your Heuristic System? Better Bring Your Deductive Competence. *North American Journal of Psychology*, 11(3), pp.543-582.

Colman, A. (2009). A dictionary of psychology. 1st ed. Oxford: OUP.

Daeef, A., Ahmad, R., Yacob, Y., Yaakob, N. and Azir, K. (2016). Multi Stage Phishing Email Classification. *Journal of Theoretical and Applied Information Technology*, 83(2), pp.206-214.

Garretson, C. (2007a). *New E-Mail Scam Targets Executives*. [online] PCWorld. Available at: http://www.pcworld.com/article/139671/article.html [Accessed 12 Mar. 2017].

Garretson, C. (2007b). Whaling: Latest E-mail Scam Targets Executives; Social Engineering Fuels New Form of Phishing. *Network World*, (Nov 14 2007).

Google, (2017). *Safety Online | Phishing, Spam*. [podcast] The Apps Show. Available at: https://www.youtube.com/watch?v=5qGTy9S7ewc [Accessed 23 Mar. 2017].

Gowtham, R. and Krishnamurthi, I. (2014). A comprehensive and efficacious architecture for detecting phishing webpages. *Computers & Security*, 40, pp.23-37.

Hong, J. (2012). The state of phishing attacks. *Communications of the ACM*, 55(1), p.74.

Ico.org.uk. (n.d.). *Accountability and governance*. [online] Available at: https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/accountability-and-governance/ [Accessed 1 Apr. 2017].

Isc2.org. (2016). *(ISC)2 Journal - Information Security Professionals Publication*. [online] Available at: https://www.isc2.org/isc2-journal.aspx [Accessed 26 Mar. 2017].

IT Governance Ltd, (2016). *Personalised phishing, Android Trojan, free pizza, and the Panama Papers*. [podcast] Weekly Podcast. Available at:

https://www.itgovernance.co.uk/blog/weekly-podcast-personalised-phishing-android-trojan-free-pizza-and-the-panama-papers/ [Accessed 23 Mar. 2017].

Jingguo, W., Yuan, L. and Rao, H. (2016). Overconfidence in Phishing Email Detection. *Journal Of The Association For Information Systems*, 17(11), pp.759-783.

Journals.elsevier.com. (2017). *Computers in Human Behavior*. [online] Available at: https://www.journals.elsevier.com/computers-in-human-behavior/ [Accessed 26 Mar. 2017].

Justice.gov. (2017). *Lithuanian Man Arrested For Theft Of Over $100 Million In Fraudulent Email Compromise Scheme Against Multinational Internet Companies*. [online] Available at: https://www.justice.gov/usao-sdny/pr/lithuanian-man-arrested-theft-over-100-million-fraudulent-email-compromise-scheme [Accessed 24 Mar. 2017].

Kanbana - your personal task manager. (2017). Denmark: Kanbana ApS.

Kazemian, H. and Ahmed, S. (2015). Comparisons of machine learning techniques for detecting malicious webpages. *Expert Systems with Applications*, 42(3), pp.1166-1177.

Keogh, S. and Jahankhani, H. (2017). Coursework Assessment Specification. *Research & Project Management, LD0749*.

Kleinman, Z. (2017). *Spear-phishing scammer demanded sex show - BBC News*. [online] BBC News. Available at: http://www.bbc.co.uk/news/technology-39338004 [Accessed 23 Mar. 2017].

Korpela, K. (2015). Improving Cyber Security Awareness and Training Programs with Data Analytics. *Information Security Journal: A Global Perspective*, 24(1-3), pp.72-77.

Lakshmi, V. and Vijaya, M. (2012). Efficient prediction of phishing websites using supervised learning algorithms. *Procedia Engineering*, 30, pp.798-805.

Langley, P. (1986). Editorial: On Machine Learning. *Machine Learning*, 1(1), pp.5-10.

Lastdrager, E. (2014). Achieving a consensual definition of phishing based on a systematic review of the literature. *Crime Science*, 3(1), pp.1-10.

Legislation.gov.uk. (1998). *Data Protection Act 1998*. [online] Available at: http://www.legislation.gov.uk/ukpga/1998/29/contents [Accessed 1 Apr. 2017].

Lötter, A. and Futcher, L. (2015). A framework to assist email users in the identification of phishing attacks. *Information and Computer Security*, 23(4), pp.370-381.

Mangelsdorf, M. (2017). What Executives Get Wrong About Cybersecurity. *MIT Sloan Management Review*, 58(2), pp.22-24.

Mansfield-Devine, S. (2017). Bad behaviour: exploiting human weaknesses. *Computer Fraud & Security*, 2017(1), pp.17-20.

Newman, L. (2017). *Phishing Scams Even Fool Tech Nerds—Here's How to Avoid Them*. [online] WIRED. Available at: https://www.wired.com/2017/03/phishing-scams-fool-even-tech-nerds-heres-avoid/ [Accessed 24 Mar. 2017].

Nizamani, S., Memon, N., Glasdam, M. and Nguyen, D. (2014). Detection of fraudulent emails by employing advanced feature abundance. *Egyptian Informatics Journal*, 15(3), pp.169-174.

Olivo, C., Santin, A. and Oliveira, L. (2013). Obtaining the threat model for e-mail phishing. *Applied Soft Computing*, 13(12), pp.4841-4848.

Osborne, C. (2017). *Seagate sued by angry staff following phishing data breach | ZDNet*. [online] ZDNet. Available at: http://www.zdnet.com/article/seagate-sued-by-angry-staff-following-phishing-data-breach/ [Accessed 24 Mar. 2017].

Payne, C. (2017). *There's Plenty of Phish in the Sea*. [online] Cm-alliance.com. Available at: https://www.cm-alliance.com/acs/theres-plenty-of-phish-in-the-sea [Accessed 24 Mar. 2017].

Pfleeger, S. and Caputo, D. (2012). Leveraging behavioral science to mitigate cyber security risk. *Computers & Security*, 31(4), pp.597-611.

PhishMe. (n.d.). *Terms of Use - PhishMe*. [online] Available at: https://phishme.com/terms-of-use/ [Accessed 1 Apr. 2017].

Publications.europa.eu. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council*. [online] Available at: https://publications.europa.eu/en/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1/language-en [Accessed 1 Apr. 2017].

Rosati, P., Cummins, M., Deeney, P., Gogolin, F., van der Werff, L. and Lynn, T. (2017). The effect of data breach announcements beyond the stock price: Empirical evidence on market activity. *International Review of Financial Analysis*, 49, pp.146-154.

STOP. THINK. CONNECT. (2016). *About Us*. [online] Available at: https://stopthinkconnect.cc/about-us/ [Accessed 24 Mar. 2017].

Tandfonline.com. (2017). *Information Security Journal: A Global Perspective :*. [online] Available at: http://www.tandfonline.com/action/authorSubmission?page=instructions&journalCode=uiss20& [Accessed 26 Mar. 2017].