DEGREE PROJECT IN COMPUTER ENGINEERING,
FIRST CYCLE, 15 CREDITS
STOCKHOLM, SWEDEN 2021

A study of phishing emails and their ability to mislead recipients depending on age and education level

KTH Bachelor Thesis Report

HÉCTOR GONZÁLEZ CAMPOS

KTH ROYAL INSTITUTE OF TECHNOLOGY
SCHOOL OF COMPUTER SCIENCE AND COMMUNCATION

A study of phishing emails and their ability to mislead recipients depending on age and education level

Authors

HECTOR GONZALEZ CAMPOS

Degree Project in Computer Science, DA150X

Examiner

Pawel Herman

Supervisor

Daniel Bosk

Abstract

Today, virtually every individual with access to an Internet connection also has a personal e-mail address. This has made it easier for companies, for example, to market their products to customers. Company employees also often have access to work e-mails, where information about upcoming meetings, new tasks, etc. is posted. Up to 45% of today's email traffic is made up of fraudulent emails that try to trick the recipient into providing personal data or clicking on a web link that then installs malicious software on the computer or mobile phone. This thesis examines how the recipient's age group and level of education affect their ability to identify fraudulent emails. The results show that this ability decreases significantly with increasing age. In contrast, level of education was not a significant factor affecting this ability.

Referat

En studie om bluffmejl och dess förmåga att vilseleda mottagare beroende på ålder och utbildningsnivå

Idag har i princip varje individ med tillgång till en internetuppkoppling även en personlig email adress. Detta har förenklat för exempelvis företag att marknadsföra sina produkter till kunder. Arbetare på företag har ofta även tillgång till jobbmejl där information om kommande möten, nya arbetsuppgifter et cetera. Hela 45% av mejltrafiken idag utgörs av bluffmejl som försöker vilseleda mottagaren till att ge ut personliga uppgifter eller klicka på en webblänk som sedan installerar skadlig mjukvara på dator eller mobiltelefon. Denna avhandling undersöker om en mottagares åldersgrupp och utbildningsnivå påverkar förmågan att identifiera bluffmejl. Resultaten visar att denna förmåga minskar avsevärt när åldern ökar. Utbildningsnivå var däremot inte en särskilt stor faktor i hur förmågan påverkades.

Acknowledgements

I would like to acknowledge everyone who played a role in my academic accomplishments and in my success. In particular, Daniel, my supervisor; Robert, my examiner; Carlos my supervisor in Spain and Pawel, the course responsible. I also appreciate the feedback given by Elias Rinzén and Jacob Wirgård. Thank you all for your unwavering support.

Contents

1 Introduction 1
- 1.1 Problem statement . . . 2
- 1.2 Approach . . . 2
- 1.3 Thesis Outline . . . 3

2 Background 4
- 2.1 Spam Mail . . . 4
- 2.2 Phishing Mail . . . 5
- 2.3 Definition: Scam . . . 6
- 2.4 Definition: Trust . . . 6
- 2.5 Related Work . . . 6

3 Method 9
- 3.1 Data collection . . . 9
- 3.2 Construction of survey . . . 11
- 3.3 Data compilation and analysis . . . 12
- 3.4 Ethical consideration . . . 13

4 Results and analysis 14
- 4.1 Real emails . . . 15
  - 4.1.1 Education level . . . 15
  - 4.1.2 Age Group . . . 17
- 4.2 Scam emails . . . 18
  - 4.2.1 Education level . . . 18
  - 4.2.2 Age group . . . 20

5 Discussion 22
- 5.1 Result discussion . . . 22
- 5.2 Possible improvements and future work . . . 23

6 Conclusion 24

References . . . 25

Chapter 1

Introduction

The ”never-ending” price decrease for technology has given more people an opportunity to own personal devices such as computers and smartphones [1] . The increase in personal devices has lead to a growth in the number of email accounts used and emails sent each year [2] . E-mails are an effective way of communicating considering how accessible it now is to reach a large number of people and companies can now easily market their products and services for practically no fee. However, this ease has unfortunately also opened the door for spam mail. Studies have shown that spam mail accounts for approximately $40 - 45\%$ of the total email traffic worldwide [3] . Spam mail is undesired and unsolicited email which may contain a link to a fake website, intending to capture the user's login credentials. Other spam mail may contain malicious software called malware that can be used to capture user information et cetera [4]. This type of spam mail is more specifically called “Phishing”, as the attacker tries to “phish” for a user's personal information. The word “Phishing” initially emerged in the 1990s, where hackers often replaced “f” with “ph” to generate new words used in the hacker community, since they most often hacked by phones [5].

We wish to investigate, depending on the recipients' age and education level, how the effectiveness of phishing email varies. In this context, an “effective” phishing email is if it managed to trick the recipient into thinking that the mail's content is “real” which would result in links being clicked or attached malware being installed.

This experiment will clarify if age and education level contribute to varied effectiveness of phishing. Furthermore, counteractive measures can be taken,

such as informing oneself, if the individual reading this report associates themselves with a particular group that appears more susceptible to phishing [6].

1.1 Problem statement

Scam emails have become a noticeable problem in society today. The purpose of this paper is to research the effect age and education level has on a person's ability to identify if an email is credible or not. Research has indicated that scammers often target the elderly [7]. Exploring why this is the case and if factors such as education also have an effect on judgement regarding emails is therefore worth investigating. This study aims to investigate the following:

- How does the effectiveness of phishing mail vary in recipients of different ages and educational levels?

1.2 Approach

In order to investigate our questions, we will first need to determine what types of phishing emails that are present and most commonly used for scamming. By doing this, a relevant representation of emails will be shown in the survey. The respondents will first answer questions about their age and education level. Afterwards, the respondents will be asked the question "Do you trust this email?" while being shown example emails, is the leading question but it also brings with it limitations as the answers will be "Yes/No/Not sure" . A consent form will naturally be used to ensure the security of the collected data. Using statistics, we will illustrate the respondents' choices using bar charts. The most relevant expected data will be whether the education level and/or age of the respondents has any correlation to the effectiveness of scam mail i.e., if the email managed to deceive them. If the results present a clear correlation between education level and the amount of "wrong" answers, we can confirm that education can affect how effective scam mail can be. The same correlation will be seen with age. We will distribute the survey through social media (different Facebook groups etc.). In this way, most respondents will be random people and this way of distribution will hopefully give us a good mix of ages and education levels. We are expecting at least 150 respondents but hope for more.

1.3 Thesis Outline

The following chapter of this paper present relevant background information, other research related to the study as well as definitions. The third describes the methodology used in the study. The subsequent forth chapter presents the results obtained from the survey. The fifth chapter discusses the results and their relevancy. The final chapter compiles the discussion and concludes the report research.

Chapter 2

Background

This chapter begins with a description of spam Mail, a unwanted mail, and follows a more concrete explanation of Phishing, a form of spam mail. What follows are definitions for recurrent terminology used in the paper, such as scam and trust. Lastly, the chapter presents a section on related research related to the subject.

2.1 Spam Mail

Digital spam can be defined as “The attempt to abuse of, or manipulate, a techno-social system by producing and injecting unsolicited, and/or undesired content aimed at steering the behavior of humans or the system itself, at the direct or indirect, immediate or long-term advantage of the spammer(s)” [8]. And spam mail can be defined as an inclusive definition from digital spam, but it also can be defined specifically as an unsolicited- unwanted or junk email- from the recipient or any email that the user does not want to contain in the inbox.

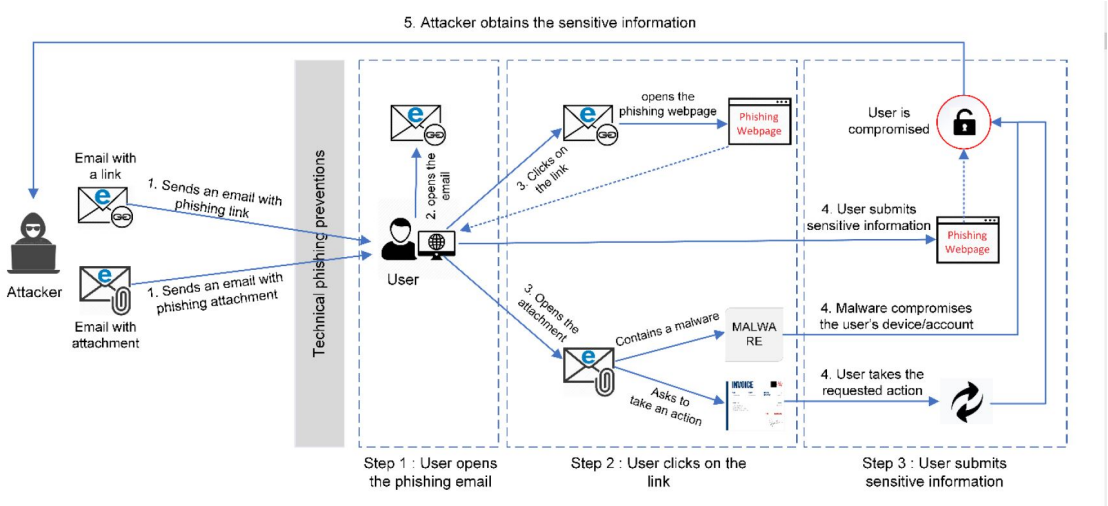
Although the research community has made tremendous efforts to alleviate the spam mail problem in the past two decades, the sense of urgency has not changed. In addition, when spam is intended to deceive or influence on a large scale, it may change the structure of society and our behavior [8]. On consequence, there is more spam mail these days and the amount is increasing day by day.

2.2 Phishing Mail

Phishing is a form of social engineering where cyber criminals make use of the Internet to fraudulently obtain sensitive information either from companies or individual users, commonly impersonating legitimate websites (Twitter, PayPal…) [9].

Nearly 90% of companies experienced targeted phishing attacks in 2019, which is a big percent to contemplate as a global problem nowadays. The volume of reported email increased 67% year over year, augmenting concern about the issue. At the same time, the main factor of these emails is to prevent user of being scammed since it has been seen that the vast majority lack some cybersecurity knowledge that could prevent these scams. For example, 45% of working adults admit to password reuse, more than 50% do not password-protect home networks, and 90% said they use employer-issued devices for personal activities [10].

Figure 1: Example of how a phishing attack is performed



For instance, a scammer may send a forged email that appears to come from a legitimate source and use the victim's weaknesses to build trust (figure 1). For example, they can ask the recipient to click a web link to win a prize. Some people have a high risk attitude or desire to gamble, which means they will click on a link to open a phishing website. Then, some of them may decide to enter sensitive information on the page [11].

2.3 Definition: Scam

Scam as a noun is defined as “a fraudulent or deceptive act or operation”, and as a verb “to obtain (something, such as money) by a scam” [12].

2.4 Definition: Trust

Trust as a noun is defined as “assured reliance on the character, ability, strength, or truth of someone or something”, and as a verb is defined as “to place confidence in” [13].

2.5 Related Work

Ludl et al. (2007) [14] investigated how effective phishing is, and looked for solutions to remedy it, mainly two popular anti-phishing solutions. For 3 weeks they tested the anti-phishing solutions integrated into the Firefox 2 (i.e., Google blacklists) and Microsoft’s Internet Explorer 7 by automatically testing them against a blacklist of 10,000 fake URLs maintained by Google and Microsoft, in order to determine their effectiveness. In addition, by analysing a large number of phishing pages, they explored the existence of page attributes that can be used to identify phishing pages. And how **these attributes (links, suspicious urls, forms, input fields) can be crucial for the users to be scammed**.

Oest et al. (2020)[15] isolated and identified detection gaps by measuring the end-to-end lifecycle of large-scale phishing attacks. They developed a unique framework that allowed them to passively measure victim tracking to phishing pages while proactively protecting tens of thousands of accounts in the process. Over one year, their network monitor recorded 4.8 million victims who visited phishing pages, excluding tracker tracking. They used these events and related data sources to dissect phishing campaigns: from the moment they connect to the network, through email distribution, visitor tracking, ecosystem detection, and finally, to account engagement. They found that the average campaign, from inception to the last victim takes only 21 hours. **At least** 7.42% **of visitors provide their credentials and eventually experience a compromise and subsequent fraudulent transaction**. In addition, a small collection of highly successful campaigns is responsible for 89.13% of victims. Based on their findings,

they highlight potential opportunities to respond to these sophisticated attacks.

Siadati et al. (2017)[16] conducted a systematic analysis of data from a large real-world embedded phishing campaign involving 19,180 participants from a single organization and used 115,080 test phishing emails. The first part of their research focused on developing methods to correct some sources of bias in order to make a more reasonable assessment of the effectiveness of embedded phishing campaigns and training. Then, they use these methods to analyze the effectiveness of embedded phishing campaigns, and through the analysis to determine how to improve the design of these campaigns. Using their method, they found that improvements in training seemed to be limited to more persuasive phishing emails, and there was no improvement to less persuasive phishing emails. **Based on their findings, they can suggest improvements to the design of the embedded phishing campaign, which may increase its efficiency and effectiveness**.

Alghamdi (2017) [17] investigated the effectiveness of phishing education and training in helping users to identify different forms of phishing threats. Users were tested for their ability to recognise fake emails, SMS phishing (SMshing), fraudulent phone calls (Vishing) and phishing via social networks. The aim of the study was to measure users' ability to recognise phishing threats and to evaluate the effectiveness of online anti-phishing educational materials. To achieve this goal, a phishing questionnaire was designed to conduct a pre- and post-test experiment to test whether there was a significant difference in the average pre- and post-scores of participants following phishing education and training materials. **The research results revealed that the scores of 43 subjects after phishing education was provided to participants, no significant changes were observed in the test scores**. The research looked at factors that may affect the results, for example, difficulty in understanding phishing education materials. However, further research is needed to address these issues and several avenues for further research are being considered.

Walrave et al. (2018) [18] used an integrative lifestyle exposure model to study the effects of routine risky activities that make a target more likely to encounter a motivated offender. To achieve these objectives, data collected in 2016 from a representative sample (n= 723) was used. And different variables that play an important role in victim hood were analysed. **A relationship was found**

between online shopping behaviour and digital copying behaviour, and phishing. In addition, a relationship was found between all online activities (except online shopping behaviour) and impulsivity. The present study suggests that especially online shoppers and users who often share and use files copied online should be trained to deal with phishing attacks appropriately.

Chapter 3

Method

The research conducted in this study is intended to be quantitative using survey methodology. A quantitative approach should be chosen since the purpose is to investigate whether distinct factors are related to a particular phenomenon [19]. Different focus groups of different ages and education levels were used to gather the relevant information needed to investigate the problem statement. The focus groups are non-academics and academics of different levels, along with their respective age groups, spanning from ages 18 to 65 and above. This chapter begins with a section describing the process of data collection, which refers to the selection of email examples relevant to the study and the reason for their relevance. Later is a section describing how the survey and questions were constructed in order to extract the relevant data needed to investigate and answer the problem statement.

3.1 Data collection

A collection of example emails had to be retrieved in order for the experiment to be conducted. To make the experiment as fair and unbiased as possible, example emails were constructed and/or found using newly conducted research regarding the content of phishing email and their specific frequency. For example, in $Q4$ of 2020, 25% of general email subjects used as phishing were email asking for "Password Check Required Immediately" [20] .

KnowBe4, an award winning service that collects phishing statistics using user reports was used to collect the most up-to-date data. KnowBe4's quarterly reports

on phishing were added together, creating an average email subject frequency through the years 2017 – 2019. A decision was made on not using data from 2020 nor 2021, since the COVID pandemic brought several new subjects that usually are not used in phishing, such as "Vaccine Registration Open". Therefore, any data from 2020 – 2021 was left out. The decision on only using data from 2017 – 2019 is further based on the assumption that phishing subjects will continue on the same trend as the years 2017 – 2019, following the end of the pandemic.

Below are two pie-charts showing the top general and top social media email subjects from the years 2017-2019. The survey examples are based on these main topics.

Figure 2

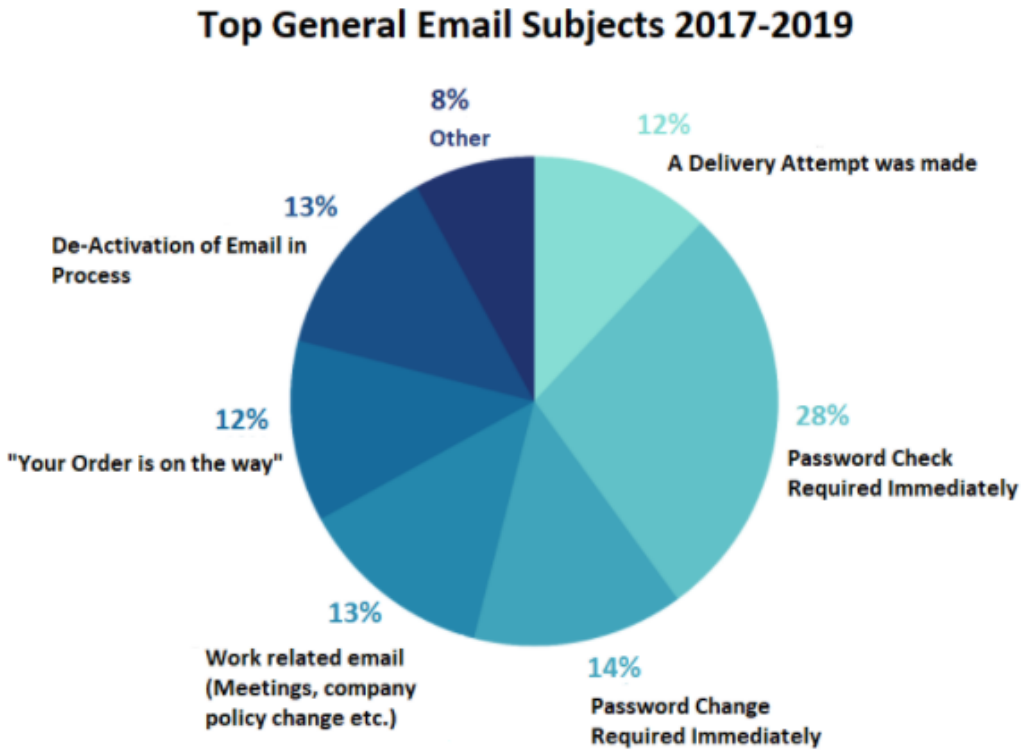
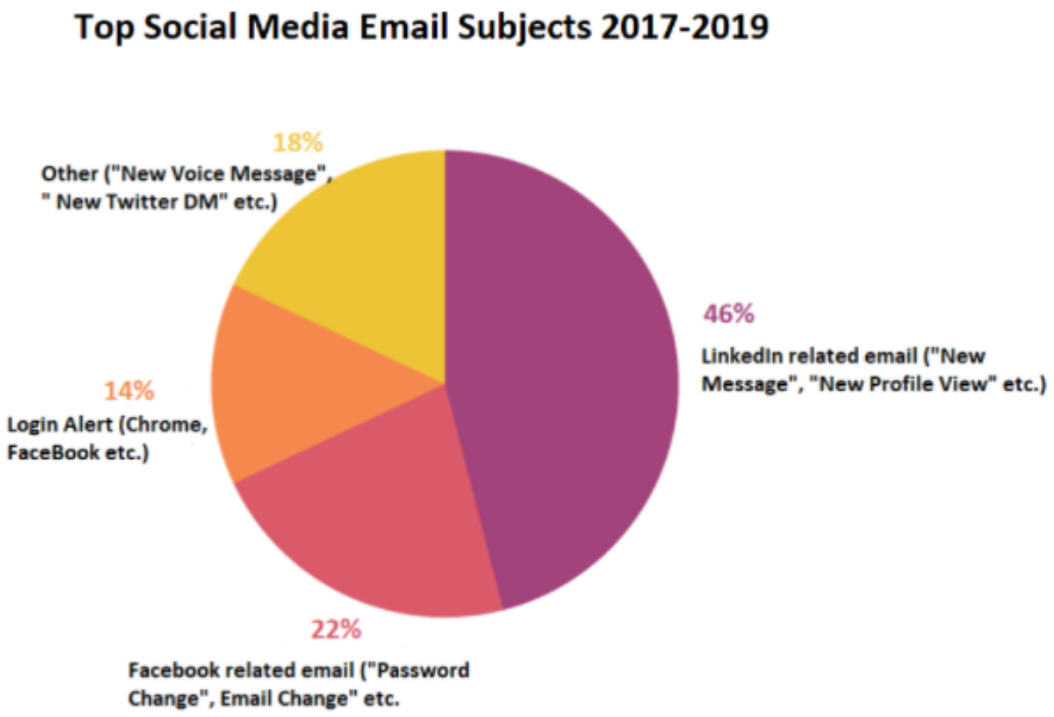


Figure 3



3.2 Construction of survey

Data was gathered with the use of a survey. The use of a survey gave us an opportunity to reach out to a greater number of respondents. We as researchers could therefore not affect the participants and since no physical contact was made between researcher and participant, a greater anonymity was established [19].

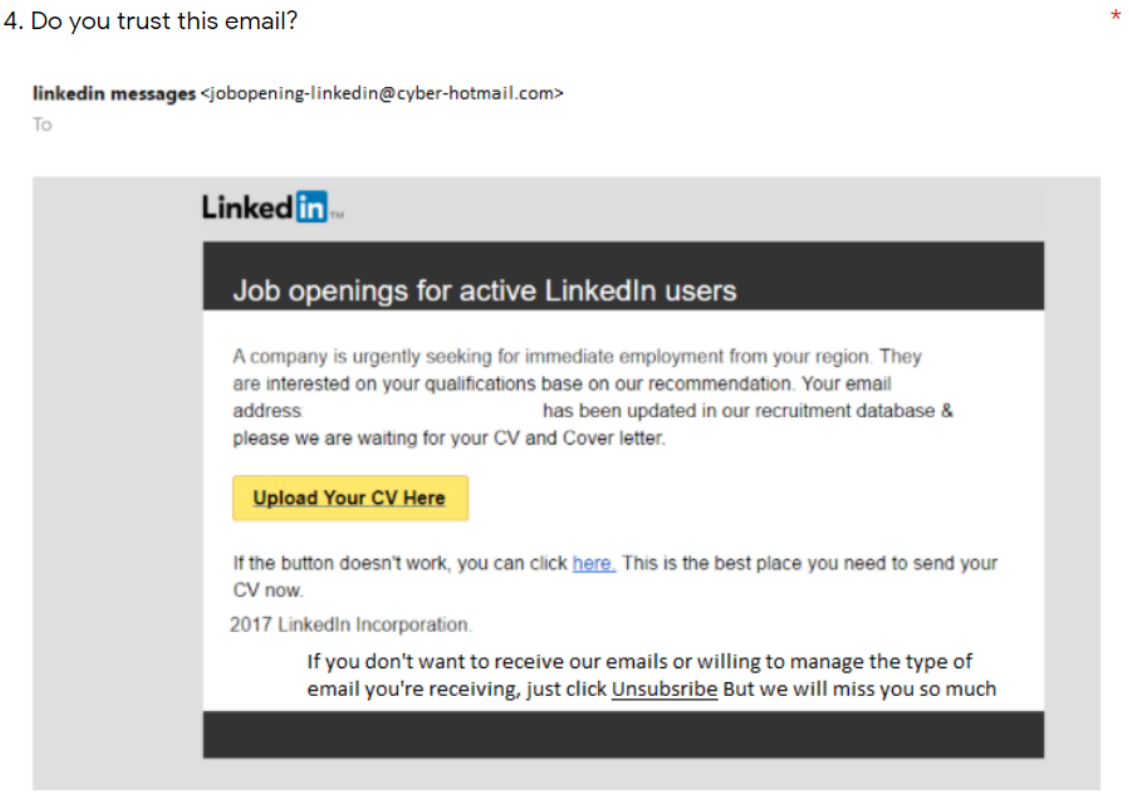
A pilot study was initially conducted in order to evaluate the used data collecting instruments, such as the pictures shown and their corresponding questions. The pilot study showed that all participants considered that the survey was well designed and understandable. Therefore, apart from the removal of concluding comments used for the pilot study, the survey was left unchanged.

Google Forms was used to create the survey since it is free and has survey tools such as automatically creating and filling a Google Sheet with data. The survey consisted of two sections. The first section was made up of a consent form and later questions regarding the participant's demographic profile comprising age and level of education. The second section was made up by example emails as well

as questions corresponding to each email. There were 10 emails in total, where 5 are considered phishing emails and 5 are real emails from legitimate senders. The subjects of the emails were based on the above-mentioned statistics from KnowBe4.

Each email had three answers, "Yes", "No" and "Not sure" with the accompanying question "Do you trust this email?". A deliberate decision was made on not including the words "scam" or "phishing" in the question, in order to minimize any influence on the respondents' decision-making and answers.

Figure 4: Example email from survey



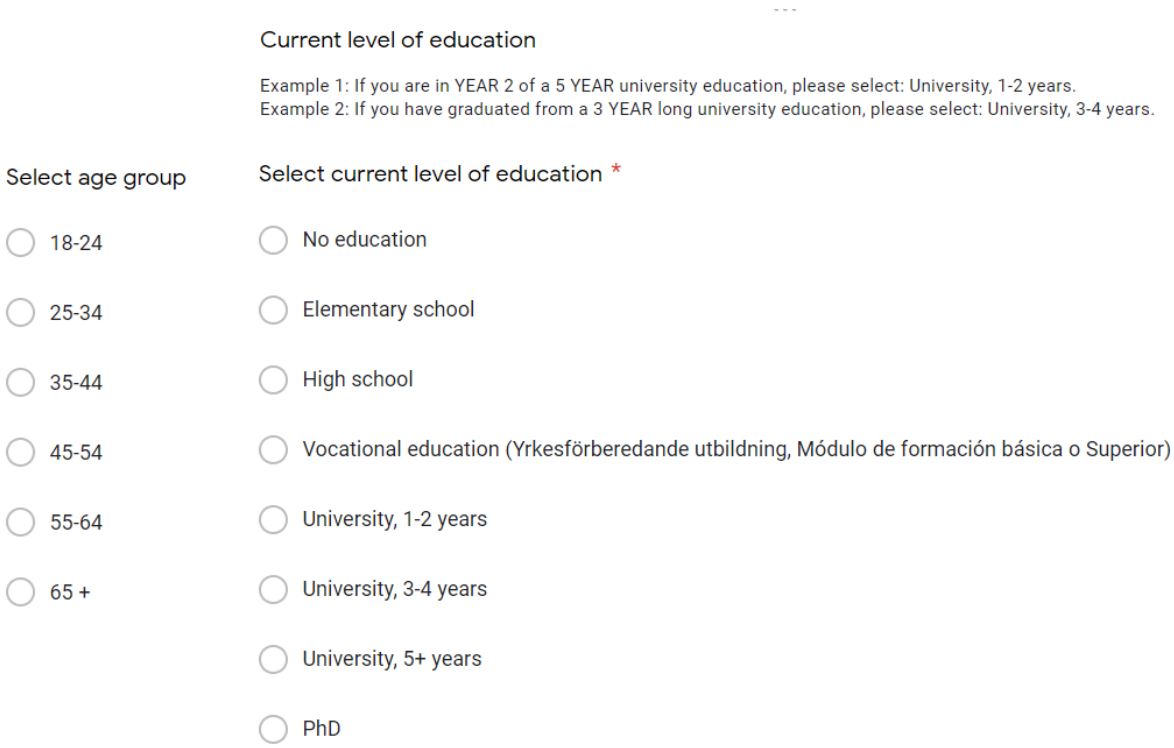
3.3 Data compilation and analysis

The survey was distributed through social media via Facebook groups and WhatsApp. The survey can be found **here**.

3.4 Ethical consideration

Participation in this study is voluntary. An option to abort an individual's participation is deemed unnecessary considering the anonymity factor of an online survey where only age and education level is given. The probability of identifying a particular individual only based on age group and level of education is deemed very low.

Figure 5: Questions about age group and education level along with examples



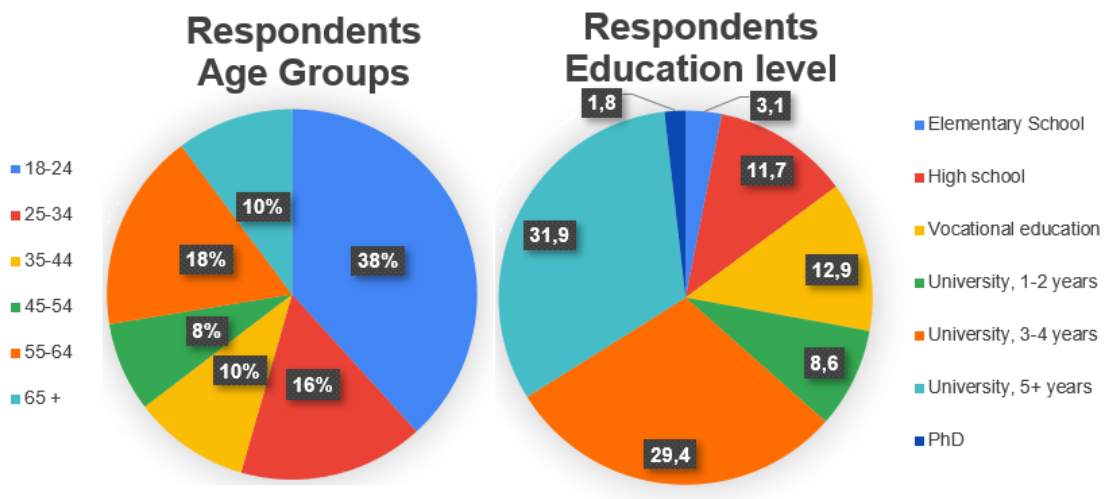
Chapter 4

Results and analysis

This chapter describes the results found in the survey. A total of 164 responses were collected over a course of 10 days. After dividing the responses gathered from real emails and scam emails respectively, the responses were further divided into two groups, one for education level and one for age group. The amount of answers from each group was calculated as well as the proportion between each "Yes", "No" and "Not sure" answer from both groups.

Below are two pie-charts that represent the overall spread of demographic profiles, divided into education level and age group (figure 6). Education level "No education" was removed from the statistics since no respondent picked this option.

Figure 6: Illustrating the spread of demographic profiles, divided into age group and education level



4.1 Real emails

Education levels "Elementary school" and "PhD" were removed from the following graphs as the number of respondents with t he above-mentioned educational profiles were $3,1\%$ and $1,8\%$ respectively. As the emails shown are real emails and the respondents answered the question "Do you trust this email?", any "Yes" answers are therefore considered correct and "No" answers are considered incorrect.

4.1.1 Education level

Figure 7

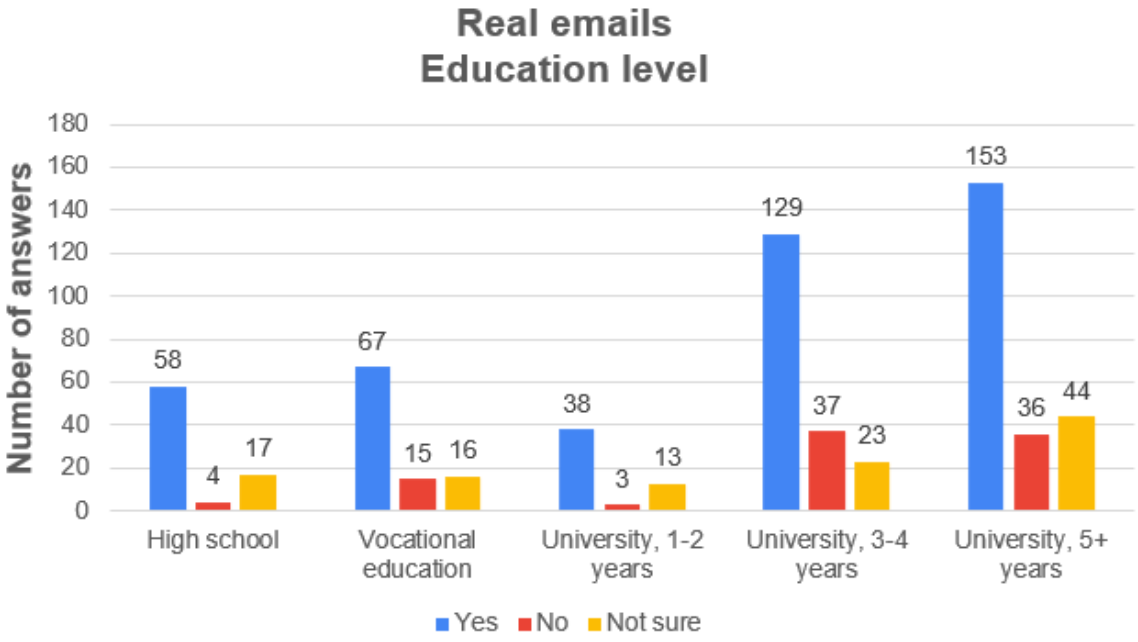
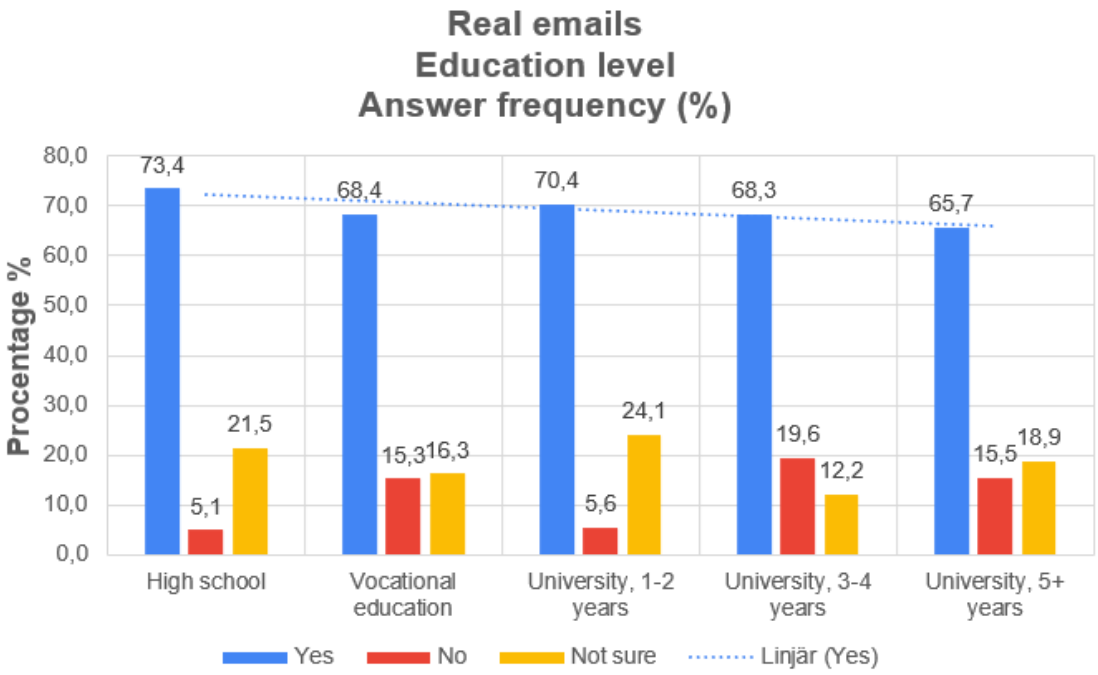


Figure 8



Education level as factor for real emails shows very similar distribution between all three answers when looking at real emails from legitimate senders. ”Yes” answers of each education level are essentially identical (figure 9).

4.1.2 Age Group

Figure 9

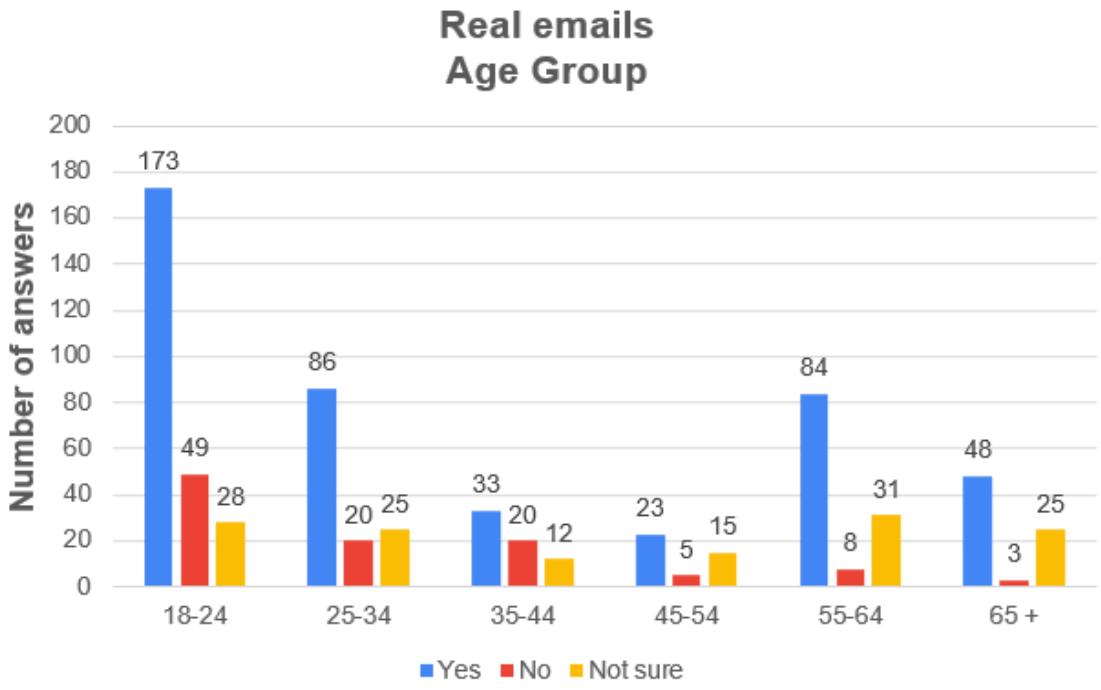
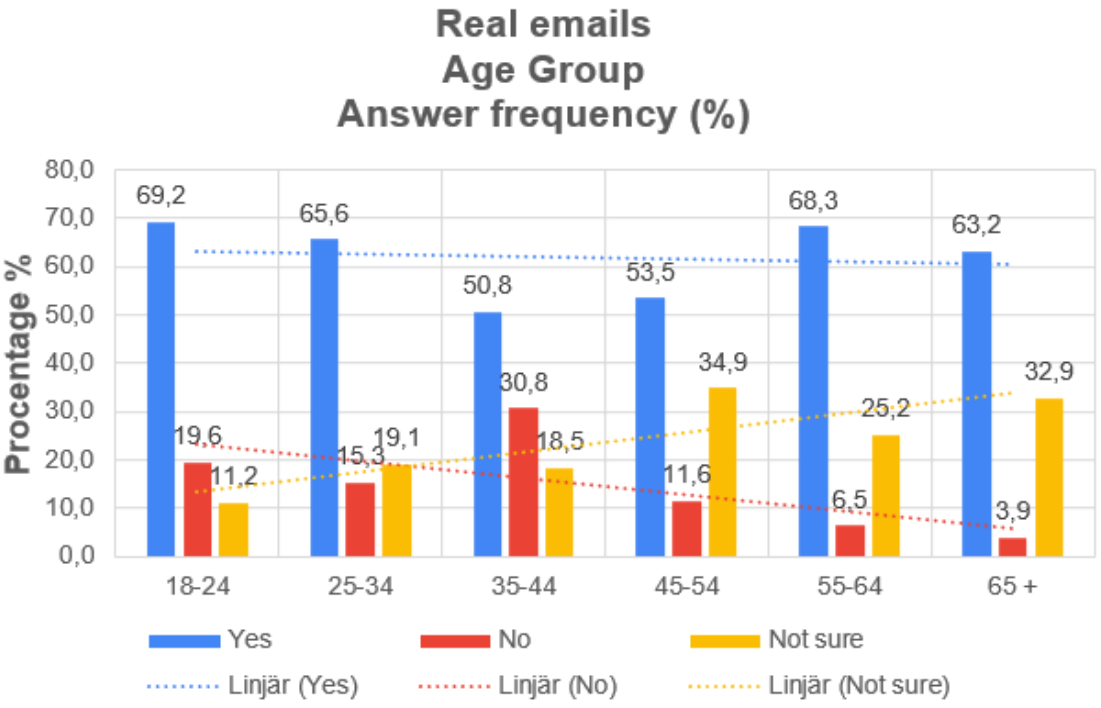


Figure 10



Age group as factor for real emails shows a similar distribution of answers for all age groups. The linear trend-line for “No” is leaning slightly downwards, going from $19,6\%$ to $3,9\%$, indicating that age group seems to have some negative effect on the ability to determine if an email is real or not. The trend-line for “Not sure” is leaning upwards, rising from $11,2\%$ to $32,9\%$, which indicates that uncertainty seems to rise in proportion to age group. ”Yes” answers drop for age groups $35-44$ and $45-54$. This might possibly be a result of the under representation of these particular age groups (figure 10).

4.2 Scam emails

The following graphs show the respondents answers while looking at scam emails. Therefore, as the respondents answered the question “Do you trust this email?”, any “Yes” answers are considered incorrect and “No” answers are considered correct.

4.2.1 Education level

Figure 11

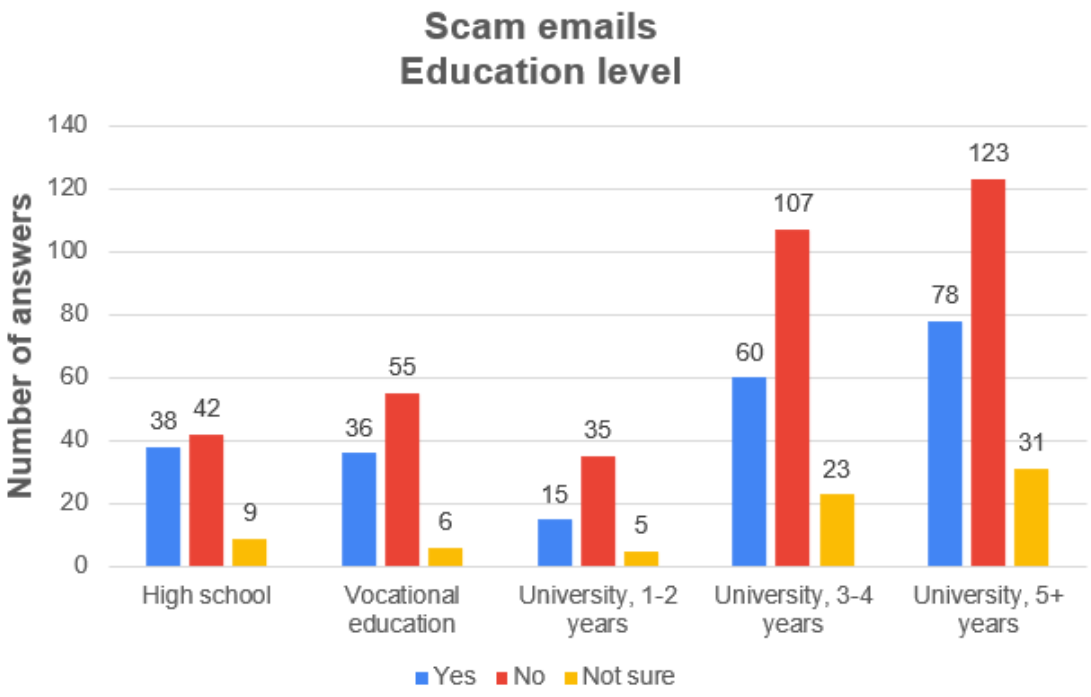
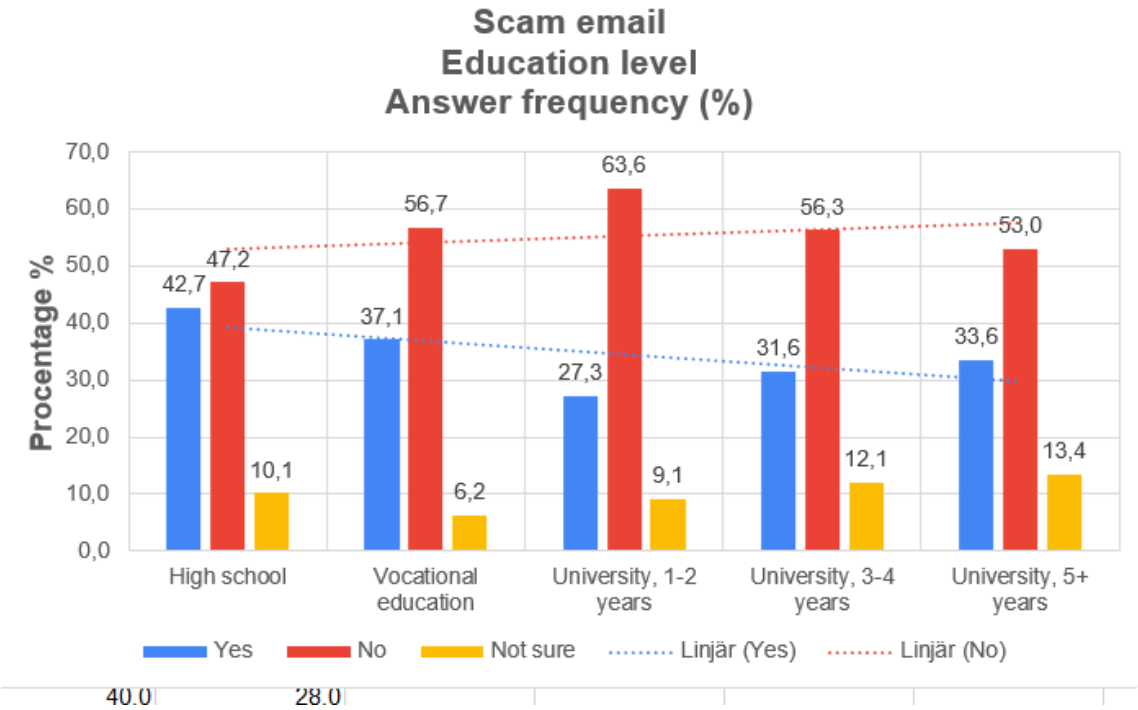


Figure 12



Education level as factor for scam emails indicates that a higher level of education seems to have some positive effect on how well an individual can spot a scam email, as the trend-line for "Yes" leans slightly downwards, from $42,7\%$ to $33,6\%$ and "No" trend-line leans slightly upwards, going from $47,2\%$ to 53%. (figure 12).

4.2.2 Age group

Figure 13

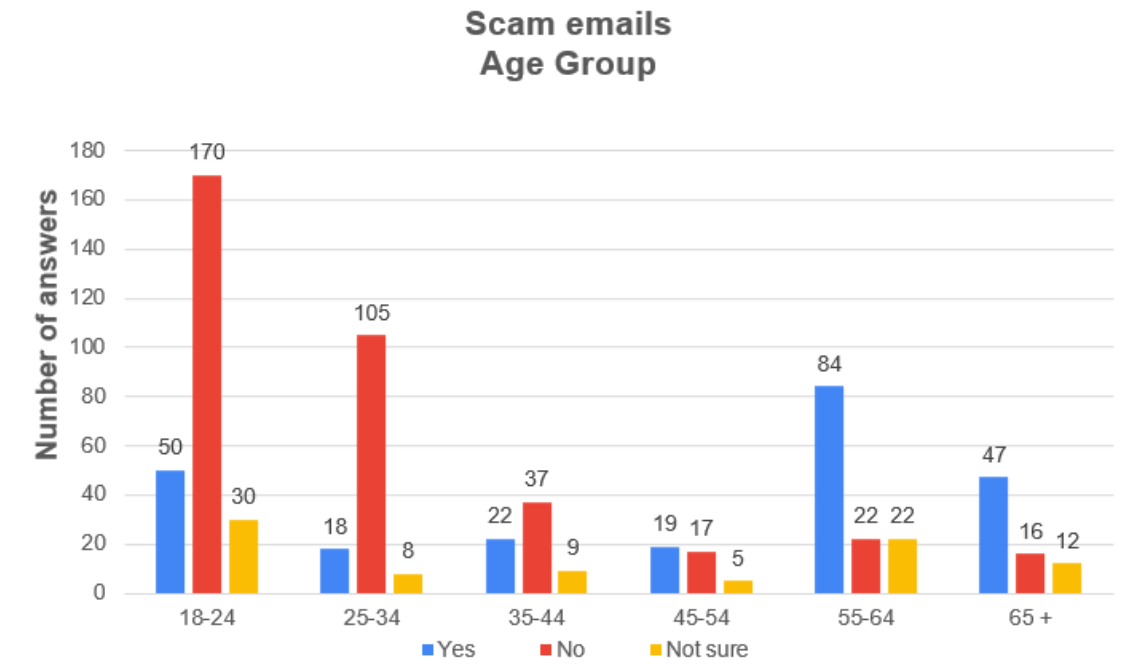
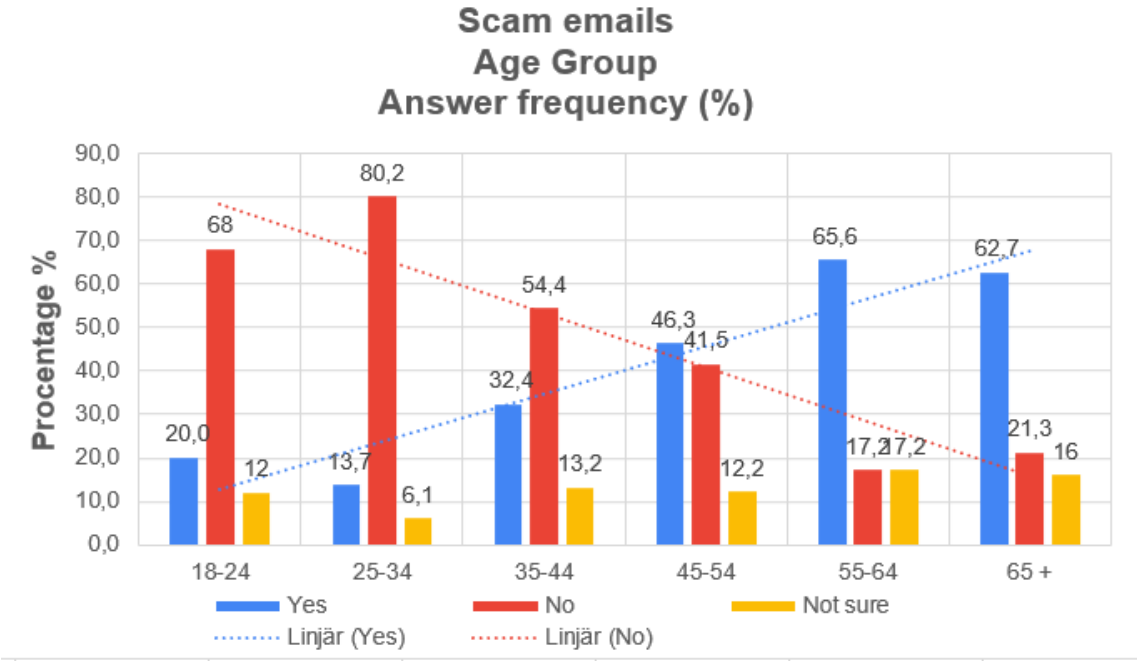


Figure 14



Age group as factor for scam emails has clear results, as the linear trend-line for "Yes" is steeply leaning upwards, from 20% to $62,7\%$, indicating that age group greatly affects how prone a person is to believing a scam email is real. The "No" trend-line further reinforces this observation, going from 68% to $21,3\%$ (figure 14).

Chapter 5

Discussion

This chapter discuss the results, possible improvements and possible future work.

5.1 Result discussion

As seen in the result for the real emails, neither age nor education level had a massive affect on how well the respondents could identify a real email. It should be noted that uncertainty seems to increase with age, as the ”Not sure” answer frequency rose from $11,2\%$ for the $18-24$ age group, to $32,9\%$ for age group $65+$. One particular email stands out from the rest, as a real ”NordVPN” email received 35.4% ”Not sure” answers. Closest to this number was a fake ”Apple” email, receiving only $24,3\%$ ”Not sure” answers. We assume that the reason for this is the unfamiliar logo, as VPN services are in a niche market. The above mentioned ”Apple” email had no clear logo. This seems to show that a well-known company logo within an email is an important characteristic to what recipients base their judgement. This assumption is further based on the fact that a fake ”Amazon” email, with a clear and well-known logo, received the most incorrect answers of all fake emails - $46,3\%$.

The statistics for age groups 35-44 and 45-54 differ from the trend-lines, which we think might the result of under-representation of those particular age groups.

The results from the fake emails are what answered our research questions. The results show that education level has some affect on the ability to identify fake emails. The education level with the highest ability seems to be ”University, 1-2 years”. However, as this group only amounted for $8,6\%$ of the participants,

we think this could be a result of under-representation. In our case, the rate of incorrect answers drops around 2% each increasing education level along with a rise of ”Not sure” answers each level. This shows that people with a higher level of education are less likely to trust an email outright, and seem to be more cautious.

Age as factor shows great diversity between how the different age groups answered. The ability to identify fake emails looks to lessen with age, as age groups 55-64 and 65+ had the highest share of incorrect answers as well as the share of ”Not sure” answers. Correct answers drop from 65% to $21,3\%$, and subsequently, incorrect answers rise from 20% to $62,7\%$, going from the youngest to the oldest age group. The rate of incorrect answers rose between $8-10\%$ with each increasing age group, showing that age group is the major contributing factor to the inability to identify a fake email.

This result was rather expected as our personal experience has shown that this is usually the case. But it is true that the statistical significance is not very large as the population reached is only 164, so our sampling error may not be very representative.

5.2 Possible improvements and future work

More time would yield more survey responses and subsequently more data. This could negate any under-representation of particular demographic profiles. As the survey was distributed through social media and we as researchers belong to the younger age groups, it was hard to receive a sufficient amount of answers from older people. Notice board were also used to distribute the survey, and we assume that older people are less prone to type in a long web link on their phone or computer. Due to the pandemic, we were limited to the above-mentioned distribution methods. A larger amount of real and fake emails could also improve the results.

Future work could include more deep diving questions, asking respondents what they base their judgement on. By doing this, one could look at what the individuals with better judgement base their decision on, and later teach their particular strategies to people with poor judgement. Other example emails could then be shown, and armed with new knowledge, the amount of correct answers would hopefully increase across all profiles.

Chapter 6

Conclusion

The results show that overall, individuals with a higher education level appear slightly more sceptical and cautious when looking at emails. However, age appears to be the most significant factor to an individuals ability to correctly identify if an email is real or not. The results also show that an email consisting of a well-known company logo increases its chances to be interpreted as a real email.

Bibliography

[1] Saracco, Roberto. “A never ending decrease of technology cost”. In: *IEEE Future Directions* (Oct. 2017). URL: https://cmte.ieee.org/futuredirections/2017/10/18/a-never-ending-decrease-of-technology-cost/.

[2] Radicati, Sara. “Email statistics report, 2015-2019”. In: *The Radicati Group, INC., A Technology Market Research Firm, Palo Alto, CA, USA April* (2015).

[3] Cvetićanin, Nikolina. “What’s On the Other Side of Your Inbox – 20 SPAM Statistics for 2021”. In: (Feb. 2021). URL: https://dataprot.net/statistics/spam-statistics/.

[4] Mallikarjunappa, Basavaraju and Prabhakar, Dr. “A Novel Method of Spam Mail Detection using Text Based Clustering Approach”. In: *International Journal of Computer Applications* 5 (Aug. 2010), p. 15. DOI: 10.5120/906-1283.

[5] Chawla, Minal and Chouhan, Siddharth. “A Survey of Phishing Attack Techniques”. In: *International Journal of Computer Applications* 93 (May 2014), pp. 32–35. DOI: 10.5120/16197-5460.

[6] Chen, Rui, Gaia, Joana, and Rao, H. Raghav. “An examination of the effect of recent phishing encounters on phishing susceptibility”. In: *Decision Support Systems* 133 (2020), p. 113287. ISSN: 0167-9236. DOI: https://doi.org/10.1016/j.dss.2020.113287. URL: https://www.sciencedirect.com/science/article/pii/S0167923620300427.

[7] Karla Pak, Doug Shadel. “Consumer Fraud Victims - AARP Foundation Study”. In: *AARP* (May 2011). URL: https://www.aarp.org/money/scams-fraud/info-03-2011/fraud-victims-11.html.

[8] Ferrara, Emilio. “The History of Digital Spam”. In: *Commun. ACM* 62.8 (July 2019), pp. 82–91. ISSN: 0001-0782. DOI: 10.1145/3299768. URL: https://doi.org/10.1145/3299768.

[9] Ghazi-Tehrani, Adam Kavon and N.Pontell, Henry. “Phishing Evolves: Analyzing the Enduring Cybercrime”. In: *Victims & Offenders* 16.3 (2021), pp. 316–342. DOI: 10.1080/15564886.2020.1829224. eprint: https://doi.org/10.1080/15564886.2020.1829224. URL: https://doi.org/10.1080/15564886.2020.1829224.

[10] Egan, Gretel. “2020 ‘State of the Phish’: Security Awareness Training, Email Reporting More Critical as Targeted Attacks Spike”. In: *State of The Fish* (Jan. 2020). URL: https://www.proofpoint.com/us/security-awareness/post/2020-state-phish-security-awareness-training-email-reporting-more-critical.

[11] Abroshan, Hossein et al. “Phishing Happens Beyond Technology: The Effects of Human Behaviors and Demographics on Each Step of a Phishing Process”. In: *IEEE Access* 9 (2021), pp. 44928–44949. DOI: 10.1109/ACCESS.2021.3066383. URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9380285.

[12] *Scam Definition*. 2021. URL: https://www.merriam-webster.com/dictionary/scam.

[13] *Trust Definition*. 2021. URL: https://www.merriam-webster.com/dictionary/trust.

[14] Ludl, Christian et al. “On the Effectiveness of Techniques to Detect Phishing Sites”. In: (2007). Ed. by Bernhard M. Hämmerli and Robin Sommer, pp. 20–39.

[15] Oest, Adam et al. “Sunrise to Sunset: Analyzing the End-to-end Life Cycle and Effectiveness of Phishing Attacks at Scale”. In: *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, Aug. 2020, pp. 361–377. ISBN: 978-1-939133-17-5. URL: https://www.usenix.org/conference/usenixsecurity20/presentation/oest-sunrise.

[16] Siadati, Hossein et al. “Measuring the Effectiveness of Embedded Phishing Exercises”. In: *10th USENIX Workshop on Cyber Security Experimentation and Test (CSET 17)*. Vancouver, BC: USENIX Association, Aug. 2017. URL: https://www.usenix.org/conference/cset17/workshop-program/presentation/siadatii.

[17] Alghamdi, H. “Can Phishing Education Enable Users To Recognize Phishing Attacks?” In: *Masters dissertation, Technological University Dublin* (2017). DOI: 10.21427/D7DK8T. URL: https://arrow.tudublin.ie/scschcomdis/99/.

[18] De Kimpe, Lies et al. “You’ve got mail! Explaining individual differences in becoming a phishing target”. In: *Telematics and Informatics* 35.5 (2018), pp. 1277–1287. ISSN: 0736-5853. DOI: https://doi.org/10.1016/j.tele.2018.02.009. URL: https://www.sciencedirect.com/science/article/pii/S0736585317304677.

[19] Polit, Denise F and Beck, Cheryl Tatano. *Nursing research: Principles and methods*. Lippincott Williams & Wilkins, 2004, pp. 19, 350–251.

[20] Sjouwerman, Stu. “Q4 Work From Home Phishing Emails on the rise”. In: (2021). URL: blog.knowbe4.com/infographic-q4-2020-work-from-home-phishing-emails-on-the-rise.

TRITA-EECS-EX-2021:458