How information sharing on Online Social Networks may allow for personalized cyberattacks

R. Waterval
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
r.waterval@student.utwente.nl

ABSTRACT

Since many interactions take place online nowadays, a significant amount of personal information is shared in a semi-public environment such as an online social network (OSN). Even if a profile has been made private and access to information on the profile is limited to specific users, the profile is not completely safe. Once one of those users' profile is compromised, the information that was access restricted is now public. Additionally, cybercrime appears to occur more often and lead to more significant damage. Personalizing these attacks may increase their effectivity. This research aims to determine what information is shared and how this may increase the effectiveness of personalized cyberattacks. Participating students at the University of Twente will be surveyed on their information sharing habits on OSN. They will also be asked how they distinguish cyberattacks from legitimate communication to determine how likely they will fall victim to a cyberattack.

Keywords

Information Sharing, Online Social Network, Personalized Cyber Attacks, Security & Privacy

1. INTRODUCTION

Ever since the beginning of the modern internet, the number of online interactions between users have increased. Many of these interactions take place at an online network, where users may interact with each other. Examples of these (OSN) are Facebook, Twitter, LinkedIn, et cetera. Facebook, (2021) [1] reported roughly 2.91 billion active users monthly in their last financial quarter. Waters et al. (2011) report that people use OSN to share information with each other. They also report that 81% of millennials realize that they give up a part of their privacy when they communicate through an OSN. This means that a significant number of people are aware of their privacy but will still share some privacy-sensitive information. Cyber criminals that obtain access to this information can use it against their respective owners. They disguise a cyberattack as legitimate looking communication and can gain access to the owner's device or files.

As the world becomes more digitized, the attack surface for cyberattacks increases. Especially with the current pandemic, the number of cybercrime committed are significantly higher than previous years as reported by Lallie et al., (2021). Jain et al., (2021) and Thakur et al., (2019) have already reported on conventional attacks such as (spear) phishing, malware attacks or social engineering attacks. Many businesses have already been on the receiving end of these attacks. Although the total costs to defend against these attacks and repair any damage that is done are hard to estimate, Sophos Ltd, (2021) [9] reports the average cost of a ransomware attack for small and medium-sized businesses is US$1.85 million.

In Section 2 other research on this topic are shortly outlined. In section 3 the research questions that this research will focus on are presented. The methodology used to answer the research questions is discussed in section 4, which is followed by the results and a discussion in section 5. Finally, section 6 will summarize the conclusions of this paper and give an outlook on follow-up work.

2. RELATED WORK

The related work was found through searches using the databases of Google Scholar and IEEE using terms such as "Social Media", "Online Social Networks", "Information Sharing", "Voluntary Disclosure", "Personalized", "Cyberattacks".

OSN are the subject of an active field of research, as can be seen in the paper written by Kapoor et al., (2018). They found an overview of 13,177 records and discussed the findings of 132 papers on social media and social networking. Regarding the security of OSN, Gupta et al., (2015) created a model for security and privacy concerns. This model was empirically tested with users on the Facebook platform. They concluded that users believe their privacy is more secure, the more they trust the platform. This does introduce a false sense of security, as the platform may not be as secure as users think it is. That in turn may lead to sharing more information than is wise. Potgieter (2019) conducted a case study on the Security Awareness of students at the Central University of Technology, who suggested that students lack awareness for cybersecurity. Since students generally belong to the age group that use OSN the most, they are at high risk of a cyberattack. Finally, Hadlington (2017) notes that the prevention of cyberattacks is shifted more to human factors, which will require more security awareness and a better awareness of what information could be used against you.

Thakur et al., (2019) has presented a number of attack types that can be used against OSN. They also conclude with some general advice on how to stay safe from these attacks. Similarly, Jain et al., (2021) also presents a num-

36th Twente Student Conference on IT Febr. 4th, 2021, Enschede, The Netherlands.

ber of possible attacks and solutions, but states that the solutions are hard to implement in a real-world environment.

Although research has been done on available attacks and security on social networks, the literature search did not show any papers evaluating the effectiveness of the personalized cyberattacks.

3. RESEARCH QUESTIONS

The available research has not shown how “personal information shared on an OSN” influences the effectiveness of the various cyberattacks. Therefore, this research will attempt to fill in this knowledge gap by asking and answering the following questions:

- RQ1: What personal information do active users of OSN share online?
- RQ2: Which types of personalized cyberattacks are frequently used, and how does their effectiveness depend on personal information of the target?
- RQ3: Does the effectiveness of the frequently used personalized cyberattacks significantly increase, with personalized information shared by active users on OSN?

4. METHODOLOGY

In this section, the used methods will be outlined to obtain answers to the research questions. In subsection 4.1 the survey will be discussed to determine what information is being shared. Subsection 4.2 will discuss the method to determine which personalized cyberattacks are common. Finally, subsection 4.3 shows how the obtained information will be combined to determine the effectivity of personalized cyberattacks.

4.1 Personal information sharing

To determine what personal information is shared by active users of OSN, a survey was held under students of the University of Twente. These students are from various studies, and were only filtered based on whether they interacted with OSN in the past thirty days. They were then asked to reflect on the OSN they used in general and how often they use them. For their three most used networks, they were asked to check which privacy and security features were activated for their account, which personal data is shown on their profile, and what user-generated content they share on that network. The exact list of questions can be found in appendix A.

4.2 Personalized cyberattacks

Frequently used cyberattacks were determined through existing literature. However, the exact frequency was not mentioned in any of the literature found during a literature search. The precise number would also be very hard to determine for several reasons. The first reason being that millions of cyberattacks are performed per day. A large group of participants would have to be tracked over a prolonged period of time to obtain an overview of all types of attacks, the frequency they occur and how effective they would be. The second reason is that many of the attacks are not reported for several reasons. The attacks can be so frequent that they are ignored by most people, such as spam emails. Some attacks happen by or to children, e.g., cyber-bullying or cyber-grooming, which is hard to track for authorities. Other attacks cause that the targets cannot or do not want to talk about it, e.g., blackmail. A third reason is that many cyber incidents that are reported will not be solved. Since it is hard to track down the offender, many people could see it as a waste of time.

The research originally planned to simulate some personalized cyberattacks against the participants of the survey. Participants would be asked to distinguish attacks from legitimate messages by showing them an environment where a target of these attacks would be logged into their OSN. This was not executed because not enough information was found on the structure and procedures of certain attacks. If the setup would still be executed, the effectivity of the attacks would therefore not have been tested, but on how well the researcher could fabricate what they thought were the attacks. Instead, the survey was changed to include questions what participants looked at to determine whether a given message was an attack or legitimate. This has given insight whether personal information of the target would influence the effectiveness of an attack, but is not conclusive.

Since the attack types are already known, this paper will summarize the attacks that can be categorized under personalized cyberattacks. A further distinction will be made to filter out attacks that are not relevant to the target group of this paper.

4.3 Effectivity of personalized cyberattacks

Since the third research question builds upon the answer of the second research question, no conclusive, numeric answer can be given to the question. Instead, a prediction and suggestion will be given whether attacks' effectiveness changes based on which personal information is shared and what participants check in a message to determine legitimacy. This information can be derived from the results of the survey.

5. RESULTS AND DISCUSSION

In this section, the results of the survey will be presented in subsections 5.1 on general OSN usage, 5.2 on OSN specific usage and 8 on how participants try to recognize cyberattacks. The overview of personalized cyberattacks is presented in subsection 5.4.

5.1 General OSN usage

The survey was held under a total of seventeen participants, all of whom fulfilled the requirements to participate and consented to participate after knowing what the information would be used for.

When asked which OSN participants used in general, they gave a total of 87 networks, of which twenty were unique. To help participants get started, a non-exhaustive list of OSN was given as an example, as well as what the definition of an OSN was. If participants felt like a specific network was an OSN but were not certain, they could decide based on the definition or contact the researcher. The definition was given as: “OSN are networks where users can share user-generated content with each other, such as freely written texts, photos, videos, locations, music, files etc.”. Participants did not seem to feel limited by the list in listing their OSN. Forty-eight networks that were listed by the participants was on the list of examples. From the unique networks, seven out of the twenty networks were on the list of examples. The entire list can be found in Table 1. Notable is the low number of participants reporting to use Email and YouTube, which is likely because it is not traditionally seen as a “social media” platform. This can be explained in part that most people do not use them as a primary source of sharing user-generated content with

friends or family. On average, a participant uses 5.1 OSN, of which 2.7 OSN are used daily. Participants reported to use at least two networks and at most ten networks. All calculated values can be found in Table 2.
If users interact more frequently with an OSN, they

5.2 OSN specific usage

The results in this subsection are based on the top three most used networks of each participant. The participants reflected on 48 networks in total, since not all seventeen participants were active on three or more networks. A couple of the networks, such as Discord, Snapchat, Tumblr, Weibo, and YouTube were only reflected upon by one or two people. Although the numbers can still be used for the overall usage of certain features and sharing certain data, it is hard to draw conclusions from those networks due to the small sample size. The adoption of some features can be explained through other means, such as the goal of the network or a setting being unavailable.

5.2.1 Privacy features

When asking the participants about which privacy features were activated, the majority had at least one privacy feature activated, whereas only six people used a network where they did not make use of any privacy features. Most of the people using a privacy feature limit access to their account. The networks where profiles includes more personal data also seem to use more privacy features. This may suggest that people are aware that their data may be used for malicious purposes. Detailed numbers can be found in Table 3.
Other privacy features that were mentioned include: Limiting certain access of certain fields to friends (Facebook), only connections can see that you are online (LinkedIn), scrutinize an unfamiliar phone number before interacting (WhatsApp).

5.2.2 Security features

When the participants were asked about their activated security features, there were three people that had none activated and two that did not know, leaving 43 networks with at least one security feature activated. The most frequently used features include setting up a recovery email address (which is usually the address used to create the account) or a recovery phone number. These are usually the least intrusive security features, in contrast to 2-factor authorization, which takes more effort but protects the account better as well. The networks that supported Single sign on (Instagram, LinkedIn, Reddit, Tumblr, TikTok, YouTube) had the feature enabled 8 out of 26 times. Single sign on makes the impact of a compromised account worse, but it reduces resistance of people's aversion to safety measures since it is only required for one login. Detailed numbers on the security features can be found in Table 4.
Setting up a recovery email address or phone number will not help to protect the account, but will help recover the account if it is compromised. Security questions work similarly. Furthermore, some networks will inform their users if their account may be accessed from an unexpected location. Two-factor authorization will help protect an account because an attacker will need to access the second account as well to fully compromise the account. If an attacker does not have that, the user will be able to change their passwords and no information is leaked.

Other Security features include: Having to use a PIN before getting access, storing Encrypted backups (both WhatsApp).

5.2.3 Personal data included in profile

The most common personal information included in a profile is the first (81%) and last name (54%), followed by age (38%). Many of the networks do not include personal data on a profile apart from a name, due to the fact they are structured more around communities instead of individuals. Combined with the fact that LinkedIn users do not limit access to their profile, a lot of their data is publicly available. Detailed numbers can be found in Table 5.
Other personal data included: On Facebook: birth town, former employment, and that marital status and contact info are private. On Instagram: nationality, gender. On LinkedIn: known languages.

5.2.4 Shared content

Most of the content that is shared on these networks are photos, written texts, comments on other people posts and videos. Many people explained that they are conscious of what they are sharing on OSN. However, photos and videos are prone to include hidden details that can tell much about the situation where they were taken. Personal details on shared content can be found in Table 6.
Other content include: On Facebook: Ads on the marketplace, Events. On Instagram: About work. On Tumblr: Reblogging other's posts.

5.2.5 Habits

Participants were asked whether they shared any habits on OSN. On 41 networks, they responded that this was not the case. On the remaining 7 networks, they shared the following habits:

- Discord: The user has the setting enabled that shows everyone in your servers whenever you play a game, as well as which game.
- Facebook: The user used to share what they were doing when they were younger.
- Instagram: The user regularly posts about work.
- Snapchat: The user shares whenever they are cooking or sometimes when using the toilet.
- WhatsApp: One user sometimes shares when they are using the toilet. Another says they share everything they do. The other shares whenever they are available to do something.

Most of these habits are not directly harmful to the user. The habits on Snapchat and WhatsApp are shared in private networks. The presence of the information is not known until an account in the network is compromised. Even if an account is compromised, the information shared is unlikely to be useful for an attack. The Facebook used to share what they were doing but abandoned that habit. They will probably not be affected any more by any attackers, since the information is likely to be outdated by now.
On the other hand, the discord user may make it easier for attackers to create spear phishing attacks based on the games they are playing. The Instagram user's vulnerability depends on what their posts contain exactly. If it is clear from some posts that they work at a certain company, they may also be more likely to be targeted by cyber espionage or spear phishing attacks. The numbers for each network can be found in Table 7.

Network	Used by number of participants	Network	Used by number of participants
WhatsApp*	76%	Pinterest	12%
Instagram	65%	Signal	12%
Facebook*	59%	Twitch	12%
Discord*	59%	9GAG	12%
LinkedIn	47%	Email	6%
Reddit*	41%	Microsoft Teams	6%
YouTube	29%	Slack*	6%
TikTok*	24%	Tumblr	6%
Snapchat	18%	Unspecified chatroom	6%
Twitter*	18%	Weibo	6%

Table 1. The number of networks participants reported to use. Networks marked with a * were included in the example list.

	Networks	Daily users	Weekly users	Weekly to monthly users	Monthly to yearly users	Other interval
Sum	87	46	20	16	4	1
Average	5.1	2.7	1.2	0.9	0.2	0.1
Max	10	5	4	2	2	1
Median	5	3	1	1	0	0
Min	2	1	0	0	0	0

Table 2. Network usage per frequency. Values are rounded to one decimal.

	Total	*Discord*	*Facebook*	*Instagram*	*LinkedIn*	*Reddit*	*Snapchat*	*Tumblr*	*TikTok*	*Weibo*	*WhatsApp*	*YouTube*
Total entries	48	3	6	11	5	4	2	1	4	1	10	1
Limited access to who can view your account	46%	33%	67%	45%	0%	25%	100%	0%	25%	0%	70%	100%
Others cannot see when you are online	38%	67%	17%	55%	20%	75	0%	100%	25%	0%	20%	100%
Limited access to who can tag/mention your account	25%	33%	33%	27%	20%	75%	0%	100%	25%	0%	30%	0%
other	6%	0%	17%	0%	20%	0%	0%	0%	0%	0%	10%	0%

Table 3. Number of users that have enabled these privacy features per network.

	Total	*Discord*	*Facebook*	*Instagram*	*LinkedIn*	*Reddit*	*Snapchat*	*Tumblr*	*TikTok*	*Weibo*	*WhatsApp*	*YouTube*
Total entries	48	3	6	11	5	4	2	1	4	1	10	1
Recovery email address	73%	100%	83%	83%	60%	75%	100%	100%	50%	0%	60%	100%
Recovery phone number	46%	33%	67%	55%	20%	25%	100%	0%	0%	100%	50%	100%
2-factor authorization	19%	33%	17%	9%	20%	25%	0%	0%	0%	0%	40%	0%
Single sign on	17%	0%	0%	27%	9%	100%	0%	100%	100%	0%	10%	0%
Security questions (and memorized answers)	13%	0%	50%	17%	0%	0%	0%	0%	0%	100%	10%	0%
other	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	20%	0%

Table 4. Number of users that have enabled these security features per network.

	Total	Discord	Facebook	Instagram	LinkedIn	Reddit	Snapchat	Tumblr	TikTok	Weibo	WhatsApp	YouTube
Total entries	48	3	6	11	5	4	2	1	4	1	10	1
First Name	81%	33%	100%	91%	100%	25%	100%	100%	75%	0%	90%	100%
Last name	54%	33%	83%	55%	100%	0%	100%	0%	50%	0%	40%	100%
Age / Date of Birth	38%	33%	100%	36%	20%	0%	50%	100%	50%	100%	10%	0%
Current or past education	25%	0%	83%	27%	80%	0%	0%	0%	0%	0%	0%	0%
Hobbies	25%	0%	17%	45%	20%	50%	0%	100%	50%	0%	0%	0%
Nickname	17%	33%	0%	36%	0%	50%	0%	0%	0%	0%	10%	0%
Employer or current job	17%	0%	50%	9%	80%	0%	0%	0%	0%	0%	0%	0%
Contact information	17%	0%	17%	9%	40%	0%	0%	0%	0%	0%	40%	0%
Current Location / Home address	15%	0%	50%	18%	20%	0%	50%	0%	0%	0%	0%	0%
other	13%	0%	50%	18%	20%	0%	0%	0%	0%	0%	0%	0%
Marital status	8%	0%	50%	9%	0%	0%	0%	0%	0%	0%	0%	0%
Who your family members are	4%	0%	33%	0%	0%	0%	0%	0%	0%	0%	0%	0%
(One or more) middle name	2%	0%	0%	0%	20%	0%	0%	0%	0%	0%	0%	0%

Table 5. Number of users that include this personal information on their profile per network.

	Total	Discord	Facebook	Instagram	LinkedIn	Reddit	Snapchat	Tumblr	TikTok	Weibo	WhatsApp	YouTube
Total entries	48	3	6	11	5	4	2	1	4	1	10	1
Photos	54%	67%	67%	91%	20%	0%	100%	0%	0%	0%	60%	100%
Texts	44%	33%	33%	27%	40%	25%	50%	0%	0%	100%	100%	0%
Comments on other posts	31%	33%	50%	45%	0%	75%	0%	0%	50%	0%	0%	100%
Videos	21%	33%	17%	27%	0%	0%	50%	0%	25%	0%	30%	0%
Location	10%	0%	0%	27%	0%	0%	0%	0%	0%	0%	20%	0%
other	8%	0%	33%	9%	0%	0%	0%	100%	0%	0%	0%	0%
Files	8%	33%	0%	0%	0%	0%	0%	0%	0%	0%	30%	0%
Memes	8%	33%	0%	0%	0%	0%	0%	0%	0%	0%	30%	0%
GIF	6%	33%	0%	0%	0%	0%	0%	0%	0%	0%	20%	0%
Music / sound	6%	0%	17%	9%	0%	0%	0%	0%	0%	0%	10%	0%
(Live)Streams	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 6. Number of users that share this type of content per network.

	Total entries	habits
Total	48	15%
WhatsApp	10	30%
Discord	3	25%
Facebook	6	17%
Instagram	11	9%
Snapchat	2	50%
LinkedIn	5	0%
Reddit	4	0%
Tumblr	1	0%
TikTok	4	0%
Weibo	1	0%
YouTube	1	0%

Table 7. Number of users that share about their habits per network.

5.3 Recognizing a Cyberattack

The participants were asked to reflect which aspects of a message they check when they try to determine whether a received message is legitimate or a disguised attack. The data can be found in Table 8.

Sixty-five % of participants made sure they knew the sender and then checked whether the contents were unusual or unexpected. Fifty-nine % checked whether a message from a company they did not know looked professional. Fifty-three % checked the source or medium through which it was received and whether the message contained any links or buttons. Forty-one % checked whether correct grammar and spelling was used. Twenty-nine % reported they were suspicious of impersonal salutations in messages. Twenty-four % of people also specified they checked the mail whether all information was correct. Twelve % were suspicious of messages with a sense of urgency.

Important to recognize here is that most of these actions happen subconsciously. When people need to explain their actions in words, people often skip the subconscious steps because they often do not realize they perform them. This may mean that more people use some of these steps even though they did not report it. However, at least this number of people will perform these steps.

Twelve out of seventeen participants reported that they were more inclined to believe a message to be legitimate when it included personal information. Six participants further clarified that the information would need to be specific information that was not shared publicly. Two participants stated they were already more inclined to believe the message if it contained their first and a last name in the salutation.

5.4 Frequent personalized cyberattacks

During the literature search performed in preparation of this research, the following list of cyberattacks that could be personalized was obtained.

- Spam: Attackers will send unsolicited bulk electronic messages that lead to scam sites or sites that try to steal credentials.
- Malware: Software that will contaminate or access a computer system to take control of the system or steal data from the system. The malware is usually spread through USB-sticks, which are distributed by leaving them in public areas or through attachments to emails.
- (Spear) Phishing: A social engineering attack with the goal to obtain sensitive and confidential information from a person. Phishing is generally done in bulk, whereas spear phishing attacks are crafted to target a specific person.
- Identity theft: An attacker will try to impersonate the person whose identity they stole, to access accounts or information they normally would not have access to.
- Hijacking: Commonly referred to as hacking. The attacker takes control of an account and changes the password and recovery settings so that it is hard for the target to regain access. Commonly used to blackmail or extort the owner.
- Cyber espionage: Usually performed in a military or corporate setting. Common scenarios include sensitive data or intellectual property being stolen from another party or keeping an eye on technological advancements of the other party.
- Cyber-bullying: Bullying, except that it happens online. Happens most often to children or people in minority communities.
- Cyber-grooming: (Child) Grooming, except that it happens online.
- cyber-stalking: Stalking, except that it happens online.

From these attacks, there are some that are less likely to occur in the target group surveyed. For example, spear phishing attacks take more time to perform than regular phishing, since they need to account for personal data. Since students generally have less wealth, it will be harder for an attacker to make a profit from the attack. Similarly, cyber espionage requires the target to have access to sensitive data or intellectual property, which most students will not have. This may be the case however in private research of PhD students.

No data was found on how often these attacks occur.

6. CONCLUSIONS & FUTURE WORK

The research gives us a glimpse into the data that is shared on OSN. Although many users have activated privacy and security features, most of them only have one or two activated, leaving their account vulnerable to a breach. Furthermore, most users seem to be aware of the danger of sharing personal data online, as can be concluded by their explanations on the content that they share. The personal data that is shared, however, is also most useful to phishing attackers.

Since the number of participants for this study was rather low, future work would include repeating the survey with a large sample size. At the same time, it is a good idea to research what the various personalized cyberattacks precisely look like. Once the attacks structure are known, tests can be conducted to determine how effective they are by themselves, and how much better they can get once the attacker has access to personal data of the target. This will help us prevent victims from being targeted in the first place if we can educate ourselves on the subject.

Acknowledgments

This paper would not have been finished without the help of several people, whom I would like to thank for their efforts. I thank my supervisor Dr. Sarmah for her time and patience, even when I did not understand something or when we disagreed on a subject. I thank Dr. Bayhan for helping me find participants for the study. I thank Jelle Nijland for spending much time together while working on our papers and keeping me motivated. I thank my friends for helping me stay positive, when things did not go the way I wanted them to. And finally, I thank everyone who could participate on such short notice in this study and for providing the data to make the research possible.

7. REFERENCES

[1] Facebook. Facebook reports third quarter 2021 results, 2021. URL: https://investor.fb.com/investor-news/press-release-details/2021/Facebook-Reports-Third-Quarter-2021-Results/default.aspx. Last accessed on 2021-11-28.

Detail	looked at by number of participants
Knowing the sender	65%
The content is not unusual or unexpected from the sender	65%
Whether the message looks professional	59%
The used media or source through which it is sent	53%
Contains links or buttons to websites	53%
Correct use of grammar and spelling	41%
Contains an impersonal salutation	29%
Message contains correct information	24%
Message contains urgency	12%

Table 8. Aspects of messages participants looked at to determine whether a message was legitimate or a distinguished cyberattack.

If the message contains . . .	% people that are more likely to believe
personal information in general	71%
specific personal information that is not public	35%
only first + last name	12%

Table 9. Number of people that reported they are more likely to believe a message is legitimate if certain personal information was included.

[2] A. Gupta and A. Dhami. Measuring the impact of security, trust and privacy in information sharing: A study on social networking sites. *Journal of Direct, Data and Digital Marketing Practice*, 17(1):43–53, 2015.

[3] L. Hadlington. Human factors in cybersecurity; examining the link between internet addiction, impulsivity, attitudes towards cybersecurity, and risky cybersecurity behaviours. *Heliyon*, 3(7):e00346, 2017.

[4] A. K. Jain, S. R. Sahoo, and J. Kaubiyal. Online social networks security and privacy: comprehensive review and analysis. *Complex & Intelligent Systems*, pages 1–21, 2021.

[5] K. K. Kapoor, K. Tamilmani, N. P. Rana, P. Patil, Y. K. Dwivedi, and S. Nerur. Advances in social media research: Past, present and future. *Information Systems Frontiers*, 20(3):531–558, 2018.

[6] H. S. Lallie, L. A. Shepherd, J. R. Nurse, A. Erola, G. Epiphaniou, C. Maple, and X. Bellekens. Cyber security in the age of covid-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic. *Computers & Security*, 105:102248, 2021.

[7] A. Perrin. Social media usage. *Pew research center*, 125:52–68, 2015.

[8] P. Potgieter. The awareness behaviour of students on cyber security awareness by using social media platforms: A case study at central university of technology. *Kalpa Publications in Computing*, 12:272–280, 2019.

[9] SophosLtd. The state of ransomware 2021, 2021. URL: https://secure2.sophos.com/en-us/medialibrary/pdfs/whitepaper/sophos-state-of-ransomware-2021-wp.pdf. Last accessed on 2021-11-28.

[10] K. Thakur, T. Hayajneh, and J. Tseng. Cyber security in social media: challenges and the way forward. *IT Professional*, 21(2):41–49, 2019.

[11] S. Waters and J. Ackerman. Exploring privacy management on facebook: Motivations and perceived consequences of voluntary disclosure. *Journal of Computer-Mediated Communication*, 17(1):101–115, 2011.

APPENDIX

A. SURVEY QUESTIONS

Before participants in the survey could answer these questions, they were filtered based on whether they are a student at the University of Twente at the moment of filling it in and whether they interacted with an OSN in the past 30 days. They were then informed about the goal and methods in the study, that they could withdraw at any point before submitting the answers at the end, and about the fact information would not be shared with any other party. After they consented to the previous statements, they could continue with the survey. The questions of the survey itself, along with their answers in case the question was multiple choice, can be found in the list below. After these questions were answered, they were debriefed with some final information on what would happen next with their answers. They were also informed of some useful cybersecurity sources and were offered to leave feedback on the survey or research in general, as well as their email addresses if they wanted to be informed of the results after the paper is finished.

1. **Sharing personal information**

    (a) Which Online Social Networks do you use, and how often do you make use of it?

    (b) *For the up to 3 most used networks:*

      i. Which of the following privacy features are activated for your account?

        A. Limited access to who can view your account.
        B. Limited access to who can tag/mention your account.
        C. Others cannot see when you are online.
        D. Other(s): ... (filled in by participant).

      ii. Which of the following security features are activated for your account?

        A. Recovery email address
        B. Recovery Phone number
        C. 2 Factor Authorization
        D. Security Questions (And memorized answers)
        E. Single Sign On
        F. Other(s): ... (filled in by participant)

      iii. Which of the following personal data are included on your profile?

        A. First Name
        B. (One or more) middle name (s)
        C. Last name
        D. Nickname
        E. Age / Date of birth
        F. Who your family members are
        G. Current Location / Home address
        H. Marital status (Single/Married/Divorced etc.)
        I. Who your employer is or what job you have
        J. Current or past education
        K. Contact information (such as phone number/ email address)
        L. Hobbies
        M. Other(s): ... (filled in by participant)

      iv. What type of content do you mostly post on this network?

        A. Texts
        B. Comments on other posts
        C. Memes
        D. GIF
        E. Photos
        F. Videos
        G. Music / sound
        H. Location
        I. (Live)Streams
        J. Files
        K. Other(s): ... (filled in by participant)

      v. Do you share any habits on the platform? If yes, which?

      vi. Please write a description of what you posted, once for each of the types you selected above.

        - You can think of the following things in your answers: What is the topic of your post? Did you talk about specific information that could (help) identify someone's identity? Is your current location included in the post? In case of images/videos: Are other people included in the picture and are they (explicitly) aware and okay with that? Is there anything in the background such as a trophy / program open on a screen / something that could identify your location?

2. **Recognizing personalized cyberattacks**

    (a) What are things that you look at when deciding whether an email/ message is legitimate or fake?

    (b) Are you more or less inclined to believe a message when it contains personal information? Does it matter what personal information is included? Please explain your answer.