ORGINAL SCIENTIFIC PAPER

UDK: 004.491
DOI: 10.59245/ps.32.3.1
Received for publication: January 2023

HRVOJE ČEMELJIĆ*, MARINA BAGIĆ BABAC**

Preventing Security Incidents on Social Networks: An Analysis of Harmful Content Dissemination Through Applications

Abstract

This study describes the dissemination of harmful content through malware applications deployed on the Facebook social network. A description of the cybersecurity incident is given, with a focus on the motive and purpose of the cyberattacks, as well as the description of the attacker and the victims. The attacker's tools are described in detail, as well as the techniques used by the attacker to reach many potential victims, infect them with malware, monetise the victims and hide the traces of the attack. Data analysis on the dataset containing information on more than two million victims is performed. The focus of the analysis is to model the dissemination of the malware and to determine the ratio of victims based on gender and country of origin. The study shows a significant statistical difference in the victims of the attacks based on their gender.

Keywords: *Facebook, cybersecurity, cyberattacks, malware applications, data analysis.*

1. INTRODUCTION

One of the key research topics of this study was to investigate the relationship between the person's gender and country of residence and the probability of becoming a victim of a cyberattack. The study aimed to determine if defining oneself as female, male, or neither could play a role in the probability of becoming a victim of a cyberattack, i.e., does defining oneself as female or male make it more probable for the user to become a victim of a cyberattack. Furthermore, the study aimed to determine if the user's country of residence

* University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia.

** University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia.

plays a role in the probability of becoming a victim of a cyberattack, i.e., does coming from a certain country or coming from a certain group of countries make it more probable for the user to become a victim of the cyberattack.

According to the research done by Redmiles et al. (2018), women are more likely to click on spam. This is of particular interest to our study since we have identified a path of actions (processes) required to be done by the user for the user to become a victim of the cyberattack. The first step in this process is a click on a spam post which immediately puts the user at increased risk of becoming a cyberattack victim, i.e., if the user does not click on the spam post, the user remains safe.

Related work on the impact of socio-economic and gender factors on cybersecurity behaviours in students was conducted by Fatokun et al. (2019). The research shows that male students had a higher score in their cyber security self-efficacy than female students. However, it remains unclear if male students self-assess their skills higher due to overconfidence. The research also shows differences in computer skills, where the male students scored higher than the female students, where around 77% of female students indicated in one of the survey questions that they were not comfortable dealing with technical issues such as: installing/upgrading computer software on their computers, as compared to the male students, where just 30% of them indicated such difficulties. This is particularly interesting to the study since it provides insight into the relationship between computer use proficiency and the probability of becoming a cyberattack victim.

Research done by Liu and Zhong (2017) showed that the top attack vector to mobile devices originates from harmful links on Facebook. The same study also describes the mechanisms of spreading threats over computer networks. This research interests us since all the victims described in this study originated from orchestrated attacks via Facebook.

In the paper by Whitty (2019), the research aimed to develop a theoretical framework to predict susceptibility to cyber-fraud victimhood. The research has found that cyber-fraud victims were more likely to be older, score high on impulsivity measures of urgency and sensation seeking, score high on addictive measures, and engage in more frequent routine activities that place them at great risk of becoming scammed. Data used in our study lacks information on the victim's age and psychophysical condition, but the related work references key insights in predicting the susceptibility of cyber-fraud victimhood.

In an analysis of what separates victims of scams from non-victims of scams by DeLiema et al. (2019), researchers found that those users who successfully evaded the scam attempt scored significantly higher in a test of financial literacy than those users who did not. When asked five financial literacy questions, those who scored higher had a lower probability of becoming victims of a scam. This related work provides insight into explaining why such a large pool of victims covered by our study was financially damaged by the attacker.

In our study, the original spam post that lured Facebook users into becoming victims of the cyberattack was titled "See why this woman killed herself after the wedding". Since the study shows that users who defined themselves as female were more likely to engage in the spam post, an examination of genders and their attitudes towards marriage was researched

to explain why the spam post lured more users who defined themselves as female. Servaty and Weber (2011) found that female students have, on average, stronger opinions and higher interest in matters regarding marriage than male students. This could explain why those users who defined themselves as male more frequently appear in the dataset (Norris et al., 2019).

2. DESCRIPTION OF THE SECURITY INCIDENT

2.1. Motive

The motive of the attacker was solely related to financial gain (Citigroup, 2016). By exploiting Facebook's security features described in this study, the attacker could achieve a significant financial gain in a relatively short period of time. This was achieved by infecting millions of Facebook users with malware and redirecting them to landing pages that served embedded video content of the victim's interest. To view the video content, users were required to send an SMS message to a subscription service. For each sent SMS, the attacker took a fee.

2.2. Purpose

The purpose of any cyber-attack is to gather the information that will be used against the victim (Han and Dongre, 2014). In this case, given that the attack was successful, personal information about victims was obtained by the attacker and stored in a database. Each record contained a victim's full name, country of residence, gender, Facebook's unique ID number, and the victim's email address. With such well-structured data, the attacker could reuse the database to launch an email spam campaign or could sell the database on the black market. With the first and last names known to the attacker, it was also possible to attempt to hijack the victim's email account by trying to log in with variations of the victim's name and surname as a password (WP Engine, 2020).

The attacker used the Facebook social media platform to gain access to such a large pool of potential victims. Using the techniques described in this paper, the attacker violated Facebook's terms of use (Facebook, 2020).

2.3. Attacker

In the case of security incidents described in this study, the attackers can be divided into two main categories;

Software development groups specialised in the production and distribution of software required by the attacker to launch an attack. These groups developed and sold software whose purpose was to automate the process of adding many Facebook connections (that is, Facebook *friends*) to the attacker's pool of potential victims. These groups also supplied the attacker with lists of fake Facebook accounts that would later be used to create and launch malware applications on the Facebook network. The attacker bought the automation software and the fake accounts from these groups. The main services that these groups offered were:

Automation software – used to automate the process of adding Facebook friends by simulating human behaviour on each of the imported fake accounts. The automation

software was sending connection requests, automatically approving received connection requests, *liking* random status updates, pictures, and videos posted by real users, and posting generic content to the news feed. With too many Facebook users, these fake accounts were indistinguishable from genuine accounts. The automation software also allowed the attacker to manage many fake accounts from a centralised location, allowing him to simultaneously launch an attack across his entire pool of potential victims (Bagić Babac and Jevtić, 2014).

Fake Facebook *Phone-Verified Accounts* (PVA) - the developer had to have a phone-verified account to publish a Facebook application. The attacker used these phone-verified accounts to develop and publish many malware applications that were used against the victims once the attack was launched. The same accounts were also imported into the automation software and used to mimic human behaviour to acquire a large pool of genuine contacts (*friends*) who would later be used as the first line of potential victims.

Proxy servers are used to simulate different physical locations of fake accounts, thus preventing the Facebook anti-malware algorithm from recognising that all the fake accounts are connected from the same location.

The attacker himself used the fake Facebook accounts and the automation software obtained from the development groups to gain access to a large pool of potential victims (Cvitanović and Bagić Babac, 2022). The remainder of the study will focus on the attacker rather than the software development groups.

2.4. Victim

Victims of these security incidents were Facebook users who mostly came from European countries. These users accessed malicious content by clicking on posts visible in their news feeds, after which they were prompted to install an application developed by the attacker. The sequence of events required for a user to become a victim of the attack is described below:

- Facebook user sees spam post in their news feed advertising allegedly premium video content. If the user ignores the post, the user is safe.
- If the user clicks the post, the dialogue box opens and asks for the user's permission to install the application. If the user ignores the prompt, the user is safe.
- If the user allows the application to install, the user becomes a victim. The attacker gains the victim's Facebook information, containing the victim's full name, gender, email address, country of residence, and Facebook's unique ID number. Two events now happen simultaneously:
  - The victim is redirected to a landing page created by the attacker. Once on the landing page, the victim is prompted to send an SMS message to a subscription service to gain access to the embedded video content. If the victim sends the SMS message, the attacker achieves instant financial gain.
  - A copy of the post that initially lured the victim is re-published by the victim. This broadens the circle of potential victims to all the victim's friends.

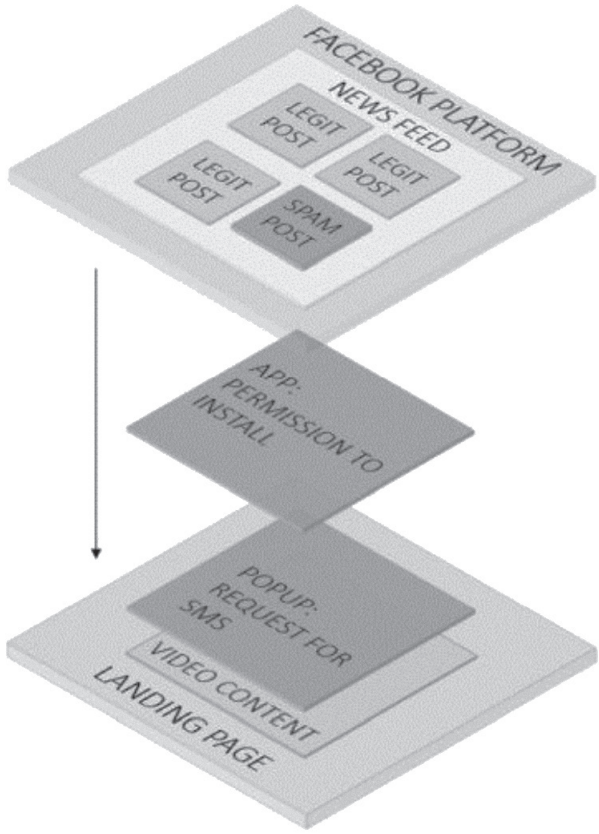

Figure 1 The path of a victim from Facebook's news feed to the attacker's landing page

Figure 1 shows the path of a user from Facebook's news feed to the attacker's landing page. Points in the path where the user's decision takes place are marked in red. These are:

- *Spam post* – the user decides whether to click it or not.
- *App* – the user decides whether to install the app or not. If the user does install the app, the user becomes the victim.
- *Popup* - victim decides whether to send the SMS or not.

The attackers' success ultimately depended on their ability to manipulate the user into clicking the spam post, installing the app, and sending the SMS.

3. ATTACKER'S TOOLS

3.1. Fake Facebook Accounts

Fake Facebook phone-verified accounts were bought in bulk from the software development groups and were used for the following purposes:

- Fake Facebook accounts were imported in bulk to the automation software. The software

then mimicked human behaviour to acquire many Facebook friends connected to fake accounts. These Facebook friends were later used as the first line of the attacked users.

- In order to develop and publish the Facebook application, a phone-verified account was required. Fake accounts were used to publish malware applications. In case Facebook's anti-malware algorithm detected and banned the application, a new one could quickly be deployed on another fake account.

Most of the fake accounts were impersonating young women. Men were mainly used as a target group of these fake accounts. This denotes elements of social engineering (Albladi and Weir, 2020) well understood by the attacker.

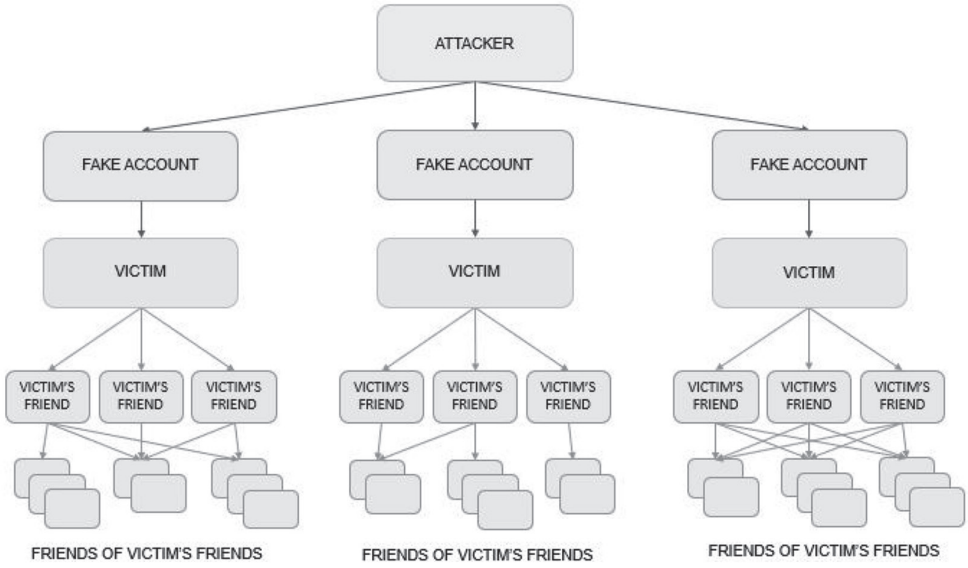

Figure 2 Graphic representation of the attacker's fake account with real people and their friends.

Figure 2 displays the network topology between the attacker and the target groups. The attacker gains access to a large pool of potential victims by orchestrating the attack via automation software.

Fake accounts were bought from the groups responsible for the development and distribution of tools required by the attackers to launch an attack. The relative price of the fake accounts was dependent upon the following:

- Degree of completion of profile information - address, education, languages, marital status, etc. Although all the information is fabricated, a higher degree of profile completion provides a greater impression of legitimacy.
- Quantity and quality of the personal photos - photos give the impression of the legitimacy of the profile, enable interaction with users, and stimulate interest and curiosity among legitimate users (Pennington, 2010).

- The number of friends already added - more friends allow more potential friend requests (Knowak, 2017), greater and faster distribution of attacks, and greater reach to the second line of potential victims (Catanese et al., 2012).
- Account creation age - older accounts have greater legitimacy than newly opened accounts, especially when adding new friends via automation software.

Occasionally, cheaper, non-PVA accounts were used to boost the probability of a successful viral attack by extending the number of potential victims in the first line of the attack. Note that non-PVA accounts could not become administrators of the malware apps.

It is important to note that in case of detection of a fake Facebook account, the Facebook algorithm deletes such an account, and together with the account, the application set up by that account is irreversibly destroyed (Schiffer and Statt, 2019).

The opposite is also true: in the case of detecting an illegitimate application, the Facebook algorithm deletes such an application, and together with the application, the user account that set it up is irreversibly destroyed.

This demonstrates that the relationship between a fake Facebook account and the application set up by that fake account is inseparable. To lose an account is to lose an application and vice versa.

3.2. Malware Applications

The goals of orchestrating the attack via malware applications were:

- Turn the Facebook user into a victim – the Facebook user becomes a victim only after the application is installed. For the application to install, the user has to explicitly click the "Allow" button in the installation prompt.
- Collect information about users – by allowing the app to install, the user gives the attacker his name, email address, gender, Facebook ID number, and country of residence.
- Redirection of users to a landing page – by allowing the app to install, the user is redirected to a landing page where a popup requests the user to send an SMS to a subscription service in order to view the video content advertised by spam posts and malware applications. If the user complies, the attacker achieves instant financial gain.
- Redistribution of the attack – by allowing the app to install, the victim re-distributes the same spam post that initially lured the victim, thus expanding the pool of potential victims to all his connections.

Each fake account owns the malware application and has full administrative rights to manage the app.

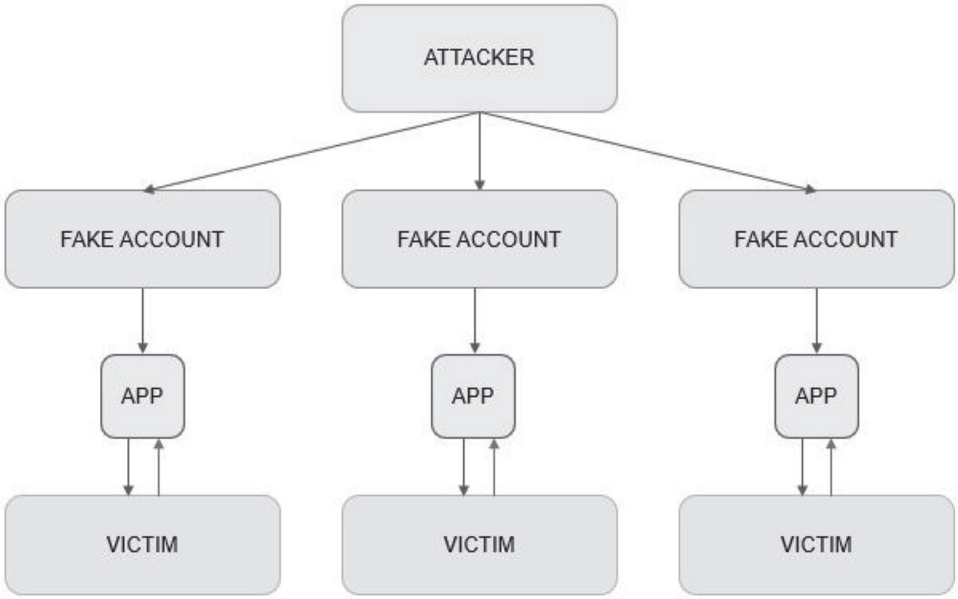

Figure 3 Network topology explaining the relationship between the attacker and his victims

Figure 3 demonstrates the network topology between the attacker and his victims:

- The attacker controls fake accounts.
- Fake accounts control the malware app.
- The malware app posts spam content to Facebook's news feed for the users to see. Once the user installs the malware app, the user becomes a victim.

While the app itself does not have to do anything harmful, it is a medium that the attacker will later use to exploit the legitimate user accounts of those users who have installed the app.

To further amplify the strength of the attack, the malware application can be customised to use other channels to propagate the attack. The basic attack technique uses a Facebook news feed in order to publish spam content that will be viewed by potential victims. Two other channels could be used:

- Facebook chat – malware takes control of the victim's messaging and sends messages to all the victim's friends. The message contains a link to the application together with a message that advertises the content. If the friend of a victim installs the app, the friend becomes a victim, and the process is repeated.
- Notifications – the app tags the victim's friends under spam posts or invites them to a Facebook event (Koolwal, 2016) that hosts the same malware application. If the friend of a victim installs the app, the process repeats, and the line of attack is broadened.

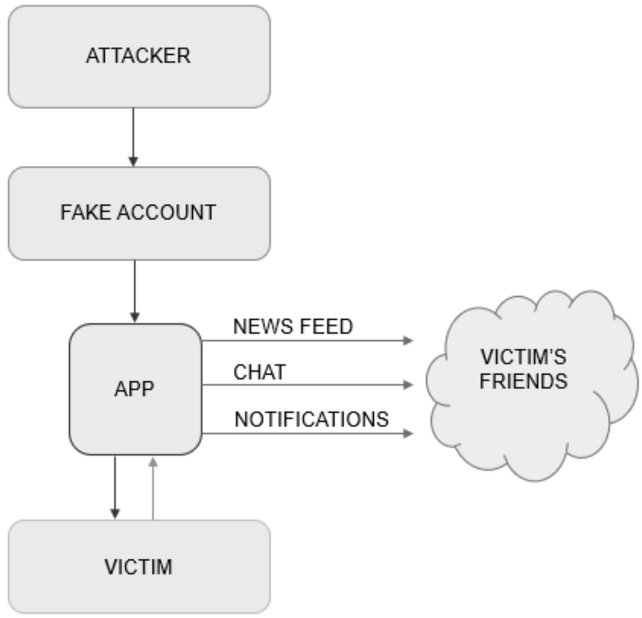

Figure 4 Graphic representation of the application spread to friends of the victim by three different channels

Figure 4 demonstrates the network topology between the attacker, the victim, and the victim's friends. By utilising all the available channels, the strength of the attack is amplified.

3.3. Attacker's Landing Pages

After each successful installation of the malware app, the victim was redirected to a landing page. Each landing page contained an embedded video covered with a popup requesting an SMS to a subscription service. In order to view the content, the victim had to send an SMS message. Once the SMS was sent, the popup would fade away, and the victim could see the video content.

In the case of this study, a video of a wedding was shown. The video shows the wedding participants visibly drunk, dancing, drinking, and occasionally rolling on the ground. At one point in the video, the groom's hair is washed in red wine in front of the bride.

Figure 5 Screenshot of the moment in the video when the groom's hair is washed in red wine (Nakita, 2011)

This screenshot was used in the application's spam post, together with the title "*See why this woman killed herself after the wedding*". The spam post lured the users into installing the app, eventually turning them into victims.

3.4. Software for Mimicking Human Behaviour and Automatically Adding Facebook Connections

In order to grow the pool of potential victims, the attacker bought the automation software used to import fake Facebook accounts in bulk. Once the fake accounts were imported, the software automatically mimicked human behaviour on each of the imported accounts to grow the number of genuine Facebook friends on that account. These genuine accounts were later used as the first line of potential victims once the attack was launched.

Key features of the automation software:

- Automatic approval of all the friend requests received on each fake account;
- An automated request for new friends (sending friend requests) according to the attacker's criterion. The criteria could be anything from the user's location and membership in a specific Facebook group to the language of the end-user;
- Automated status updates, giving the impression that a human being is behind the fake account;
- Automated sending of messages to the connection's inboxes;
- Automated posting on the connection's timeline;
- Automated posting and updating of the fake account's photos, videos, activities, etc.

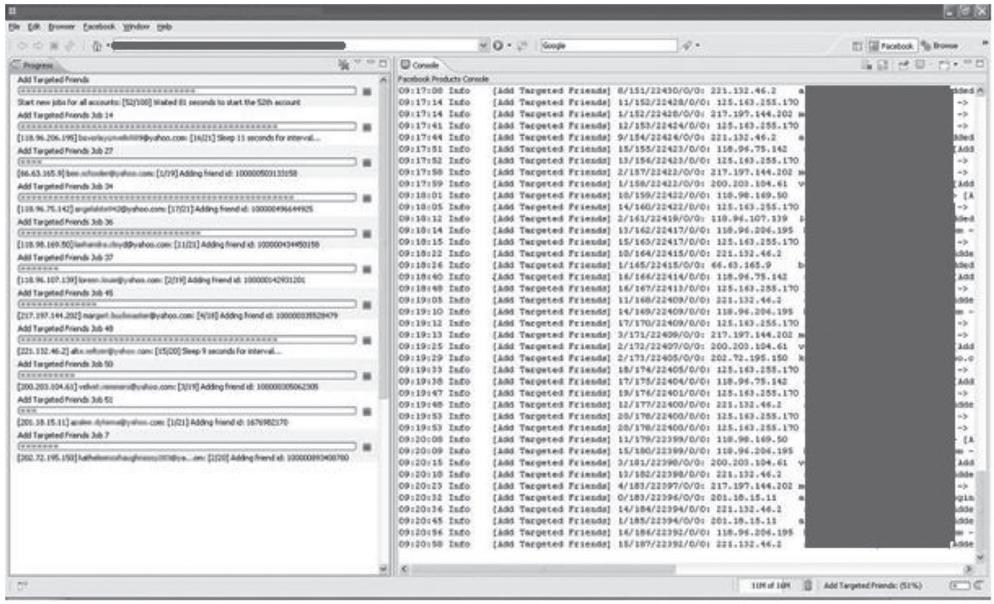

Figure 6 GUI of the software used for the automated adding of Facebook friends (Facebook FriendAdder Pro software[1])

The software provided options to import lists of proxy servers to simulate different physical locations of the connected fake accounts. This was done to evade Facebook's malware detection algorithms that would otherwise detect those hundreds of accounts are all connected from the same location.

Other features that were designed to help evade Facebook's malware detection algorithms were:

- Regularly signing out of the network and signing back in after a few hours;
- Randomising the number of sent friend requests in a given time window;
- Limiting the total number of send requests per time window;
- Taking approximately 8 hours off the network around the same time every day to simulate human sleep;
- Limiting the targets of sending friend requests to people living around the same area (most common was to target large cities);
- Randomise the messages sent to potential victims' inboxes by using synonyms. For example: "*nice*|*glad*|*happy to meet*|*see you*"; where "|" is the delimiter between randomly chosen synonyms.

[1] https://en.downloadastro.com/apps/facebook_friendadder_pro/

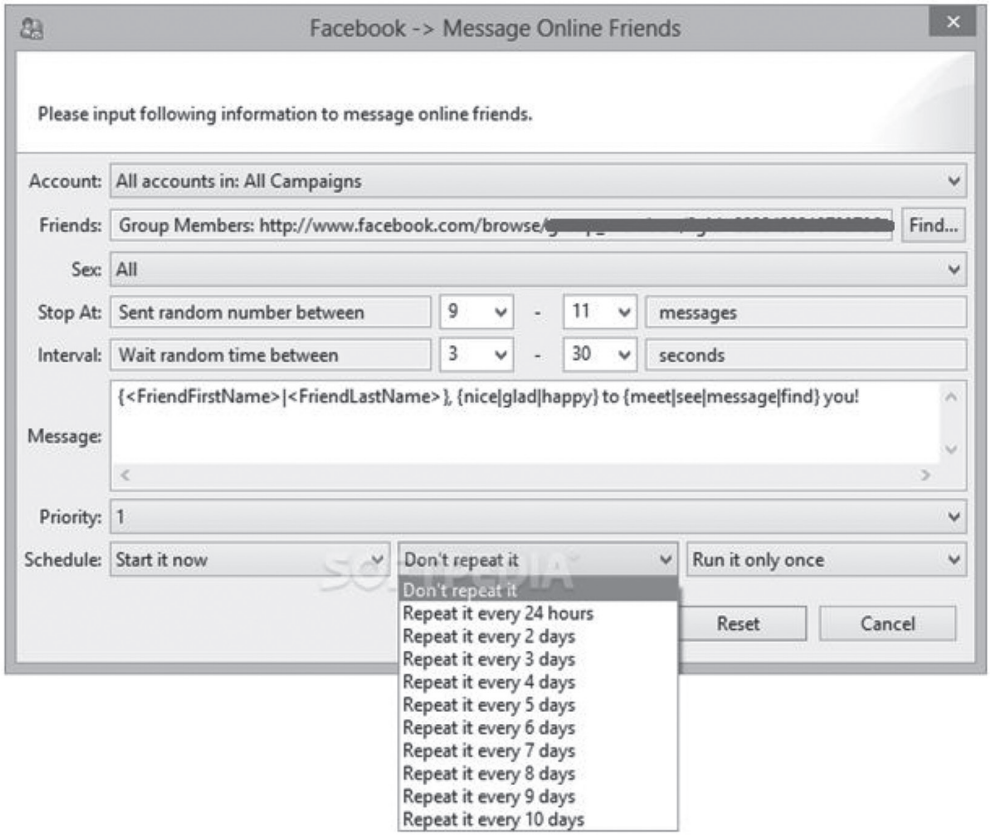

Figure 7 Fine-tuning the process of automated sending of messages to the potential victim's inbox (Facebook FriendAdder Pro software)

Figure 7 shows the fine-tuning of the parameters that is used when sending messages to friends. The parameters can be fine-tuned based on:

- Attacker's campaign;
- Membership status of a specific Facebook group;
- Sex;
- Total number of sent messages;
- The interval between sending messages;
- Randomising words in sent messages;
- Scheduled activity time.

Given the size of the Facebook network at around 2.8 billion users in the second quarter of 2021 (Tankovska, 2021), it was possible to gather tens of thousands of connections that will be used in the first line of the attack by using the described automation software. This initial pool of potential victims allowed the attacker to launch orchestrated attacks from a central location, quickly infecting thousands of victims at the early stages of the attack.

3.5. Other Requirements

To deploy an app, a unique domain name had to be registered. The importance of registering a unique domain name is explained as follows.

- Each application had to be associated with a domain name - it was not possible to publish a Facebook application without using a domain name associated with the host of that app. In order to evade Facebook's malware detection algorithm, each application was associated with a unique domain name, making it harder for Facebook to trace the source of the attack.
- Landing pages on which the advertised video content was embedded and covered with a popup requesting an SMS had to be hosted by the attacker. The same domain name used to publish an app was also used to point to a host that served the landing page. In case one landing page was identified by Facebook and the App hosted on that domain was banned, the rest continued to operate.

4. TECHNIQUES OF ATTACKS, MONETISATION, AND COVERUP

This part of the study describes techniques used by the attacker to orchestrate the attack, achieve financial gain, and conceal the traces of the attack.

4.1. Basic Attack Techniques

The attack is orchestrated in the following order:

1. The attacker accesses the administrator part of the Facebook application through his fake Facebook user accounts. Each fake account is already connected to hundreds (or thousands) of genuine friends using the automation software, thus providing the first line of the potential victims of the attack. In the administrator part of the application, the attacker enters the URL of his landing page that will be shared by all the fake accounts and the accounts that will later become victims, as well as the photo that will represent the post in the news feed (in this case the screenshot of the wedding described in the previous chapter).

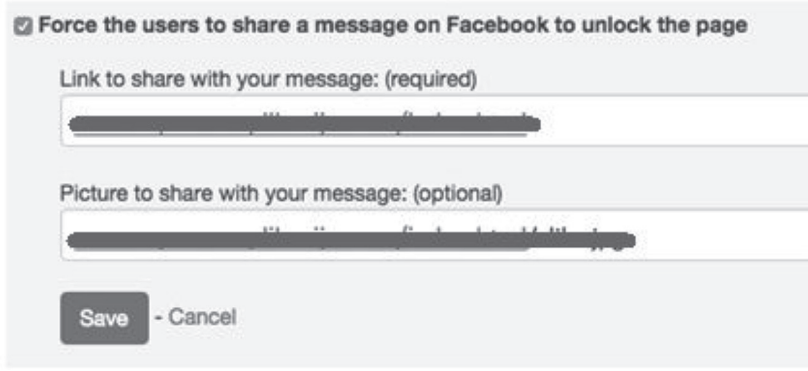

Figure 8 Launching the attack via the attacker's malware app (Facebook FriendAdder Pro software)

2. The attack is launched, and the following post appears in the news feed of all the genuine accounts connected to the attacker's fake accounts. Each user who clicks the post is prompted to install the application in order to view the video content.

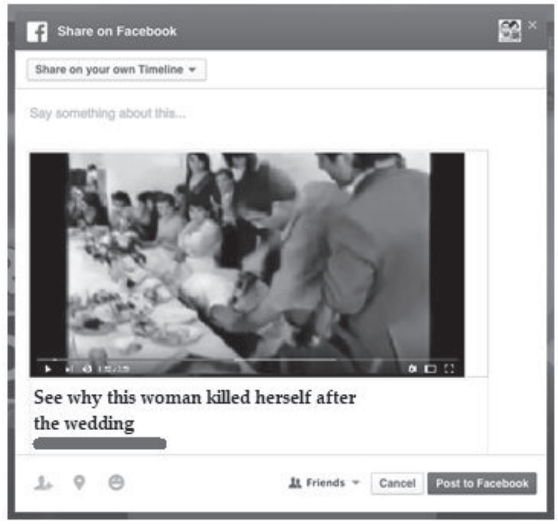

Figure 9 Facebook content-sharing form[2]

3. The user is prompted to install the malware app. At this point, the user's action is a decisive factor in determining if the user will become a victim. If the user clicks the "Allow" button, the user becomes the victim.

Figure 10 The form through which the user decides whether to install the application or not

[2] https://www.facebook.com

The functionalities of the application are provided by Facebook's app development platform, so the conditions of the app are clearly presented to the user. There is no way for the attacker to bypass this step or to get more permissions from the users than stated in the request[3]. In this case, the app requests the permissions for:

- Access to the user's basic information, including name, profile picture, gender, user ID, list of friends, etc.;
- Permission to post to the user's wall (*timeline*);
- Permission to access user's data when the user is not using the app;
- Permission to manage user's pages if the user has administrative rights to any Facebook pages.

Given the extremity of the control of the user that the attacker asks for, it is not entirely clear why such a large pool of users allowed the app to install. This could be due to numerous socio-economic factors, but no data can establish a hypothesis (Peltonen et al., 2018).

4. In the final step, the victim is redirected to the attacker's landing page, where a popup is blocking the video content. In order to remove the popup, an SMS has to be sent. If the victim sends the SMS, the attacker achieves instant financial gain. Again, it is the user's action that decides if the attacker will achieve financial gain or not.

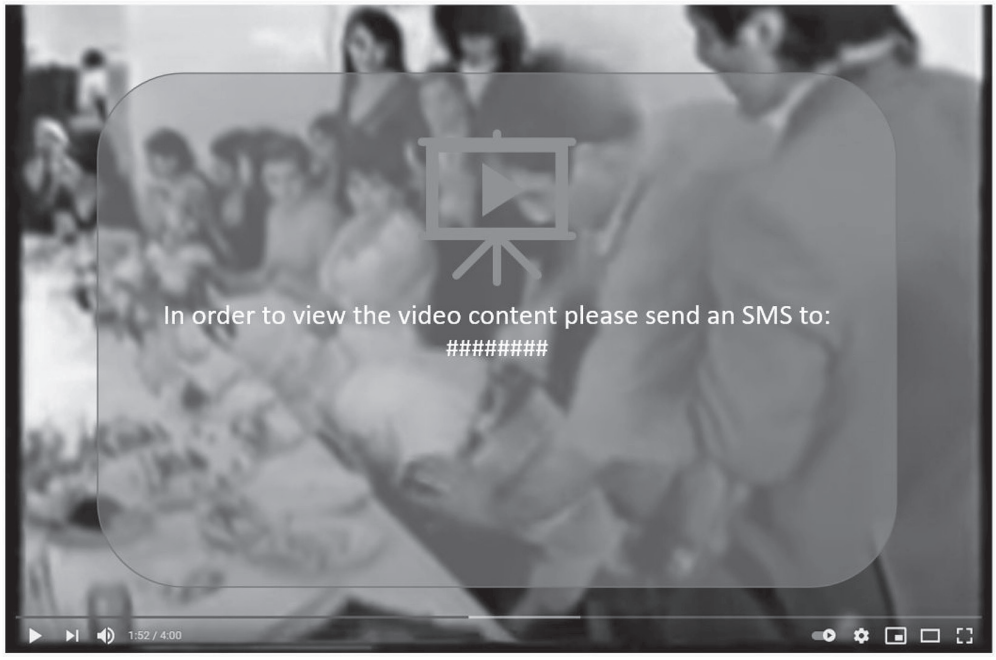

Figure 11 The video content is blocked by a popup requesting an SMS

[3] https://developers.facebook.com/docs/facebook-login/permissions/overview/

4.2. Further Enhancing the Productivity of Attacks Using Geolocations

Since the original spam post was titled "See why this woman killed herself after the wedding", it is reasonable to state that to perceive the post fully, the potential victim must first know the English language.

The attacker translated the same spam post into numerous European languages. A script was developed that checked the geolocation of the victim and ensured that the re-publishing done by each new victim was done in the native language dominant in the victim's geolocation. For example, Croatian if the victim's geolocation is Croatia, etc.

This allowed the attacker to launch a smaller localised attack in one geolocation and wait until a bilingual victim was infected. For example, if an attack is launched in Croatia, and a bilingual French-Croatian victim is infected, given that the victim was in France during the time of the infection, the re-publishing was done in French. This allowed the attacker to "jump" from one community to another without necessarily having a pool of potential victims in the other community.

By orchestrating the attacks with automatic translations of the spam post, the attacks quickly spread across most of the European continent, infecting millions of Facebook users.

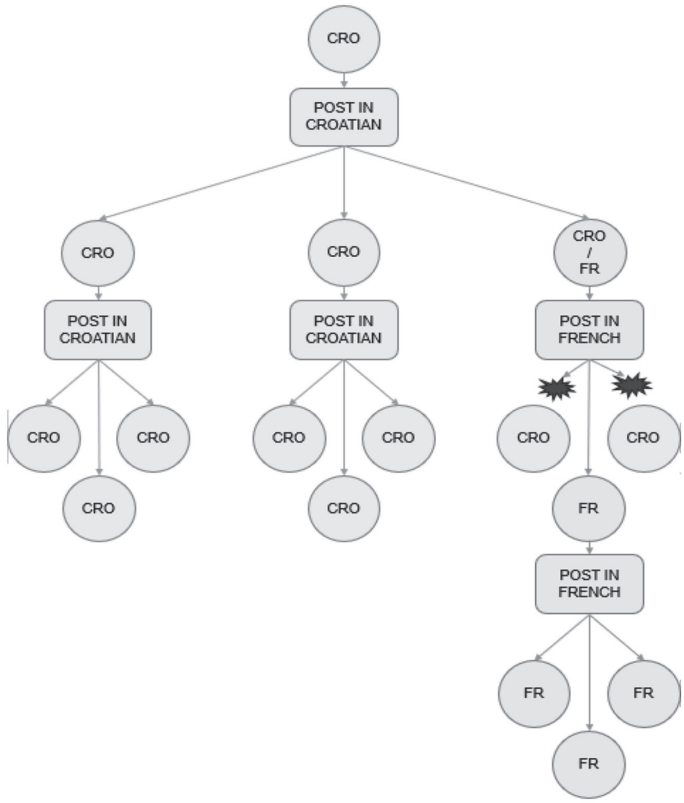

Figure 12 The effect of automatically translating the post from Croatian to French

4.3. Monetisation Techniques

In order to achieve financial gain, the attacker must manipulate the victim into sending an SMS message to a paid subscription service. The SMS service is provided by a 3rd party company. The details are not described in this study.

Note that the attacker is required to clearly state the SMS pricing and the subscription terms to the victim in order to comply with the 3rd party SMS payment provider. The pricing of the SMS varies according to the geolocation of the victim and the available subscriptions for that geolocation. On average, the SMS had a cost of between 0.4 USD and 6.00 USD. It is not entirely clear why such a large pool of users sent the SMS. In this study, we do not have the data required to establish the statistics that would show the percentages of users who willingly spent the money to see the video content against the users who did so without realising how much they are paying and what are the exact terms of their subscriptions.

4.4. Coverup Techniques

In order to reduce the probability of their applications being detected by Facebook's anti-malware algorithms, attackers designed random app rotators.

The purpose of the random app rotator is:

- Serve random apps to random users, making it more difficult for Facebook to identify sources and the network propagation of the malware apps.
- Reduce the rate of growth of users under each app. Instead of deploying the same app on every account, randomly selecting apps randomises the distribution of users per app.
- Easier administration for the attacker. In case one app gets banned, a new app can quickly be deployed to the random rotator.
- Randomised redirection to landing pages. Each random app randomly chose the redirection landing page, making the traffic evenly spread across all the attackers' sites.
- The attacker controls the random app rotator.
- Each fake account administers a random application.
- Each victim gets infected by a random application administered by the fake account.

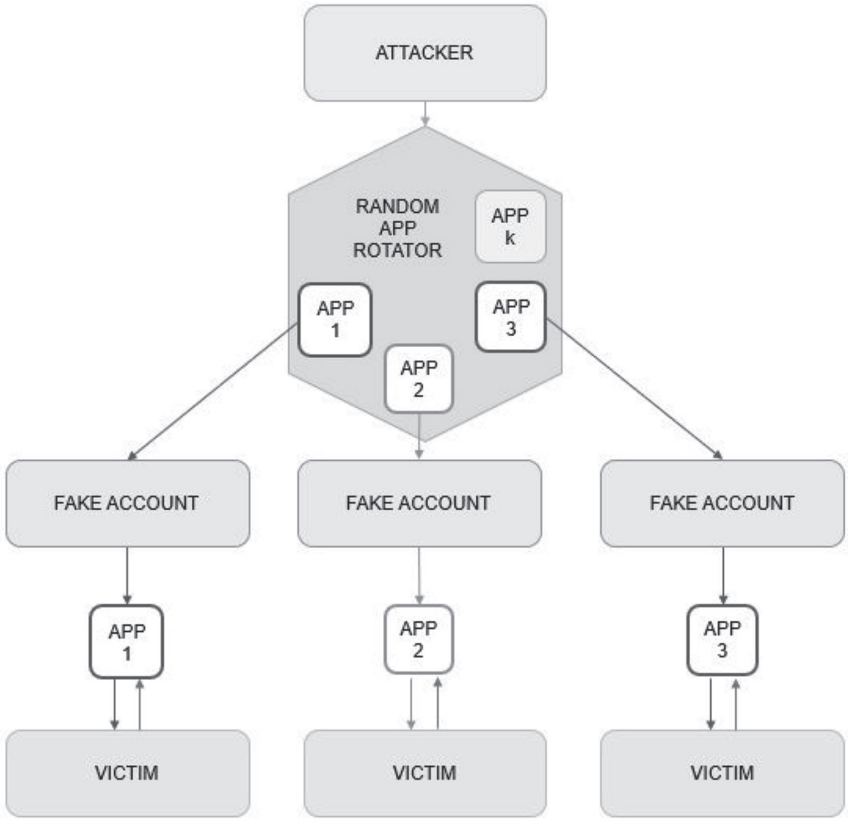

Figure 13 Use of the random app rotator for the purpose of concealing the traces of the attack

5. DATA ANALYSIS AND STATISTICS

Data cleaning, data analysis, and basic statistics (Lipovac and Bagić Babac, 2023) are given on the dataset containing the list of victims of the security incidents described in the study. Python programming language and the *pandas*[4] and *matplotlib*[5] libraries are used (Šandor and Bagić Babac, 2023).

5.1. Dataset Structure

The dataset used in this study consists of two million records stored in a JSON file. Each record describes the victim. Data structure and data types are as follows:

[4] https://pandas.pydata.org

[5] https://matplotlib.org/stable/#matplotlib-visualization-with-python

{
“idx”: “int”,
“uid”: “int”,
“status”: “string”,
“name”: “string”,
“first_name”: “string”,
“last_name”: “string”,
“gender”: “string”,
“email”: “string”,
“lastUpdate”: “timestamp”
},

Where:
“idx”: Victim’s ID from the perspective of the attacker.
“uid”: Victim’s Facebook ID.
“status”: Victim’s country.
“name”: Victim’s full name as shown on Facebook.
“first_name”: Victim’s first name.
“last_name”: Victim’s last name.
“gender”: Victim’s gender
“email”: Victim’s email address.
“lastUpdate”: Timestamp at which the victim installed the malware.
Note: all the entries in the timestamp column are identical in this dataset, making any attempt to model the data with respect to timestamps extremely unlikely.

Example:
{
“idx”: “1”,
“uid”: “123456789”,
“status”: “HR”,
“name”: “First Last”,
“first_name”: “First”,
“last_name”: “Last”,
“gender”: “male”,
“email”: “first_last@example.com”,
“lastUpdate”: “2015-5-05 05:55:55”
},

5.2. Exploring the Dataset

Pandas and matplotlib libraries are imported. The entire dataset is loaded into a Pandas *DataFrame* structure. The Head of the *DataFrame* is printed:

```
DataFrame head:
    idx             uid status  ...  gender                      email          lastUpdate
0  1480  [redacted]      HR  ...    male  [redacted]  2010-11-13 09:18:45
1  1481  [redacted]      HR  ...  female  [redacted]  2010-11-13 09:22:01
2  1483  [redacted]      EN  ...    male  [redacted]  2010-11-13 09:33:37
3  1484  [redacted]      HR  ...    male  [redacted]  2010-11-13 09:35:02
4  1485  [redacted]      HR  ...  female  [redacted]  2010-11-13 09:35:04
```

The shape of the DataFrame is printed:

```
DataFrame shape:
(2031658, 9)
```

5.3. Data Cleaning

Unique values from the ‘gender’ column are printed:

```
Unique genders before data cleaning:
['male' 'female' '' 'moÅ¡ki' 'muÅ¡ko' 'mujer' 'mÃ¤nnlich' 'Å¾enski'
 'donna' 'hombre' 'Ð¾ÑƒÑˆÐºÐ¸' 'femme' 'Å¾ensko' 'feminino' 'homme'
 'Ð¶ÐµÐ¾Ð°' 'Ð¶ÐµÐ¾Ñ\x81ÐºÐ¸' 'vrouw' 'Ð¾Ð°ÑˆÐºÐ¾' 'man' 'uomo' 'kobieta'
 'kvinne' 'femÃ«r' 'masculin' 'masculino' 'weiblich' 'home' 'kvinna'
 'x-x›x¯' 'Ð¶ÐµÐ¾Ñ\x81ÐºÐ¾' 'fÃ©rfi' 'fÃ©minin' 'mashkull' 'erkek' 'mand'
 'Å¾ena' 'nÅ‘' 'mann' 'kvinde' 'Ð¾ÑƒÐ¶Ñ\x81ÐºÐ¾Ð¹' 'mate'
 'Î¬Î¼Ï„Ï\x81Î±Ï‚' 'Ð¾ÑŠÐ¶' 'Ø°ÙƒØ±' 'dona' 'vrouwelijk' 'muÅ¾'
 'Ð¶ÐµÐ¾Ñ\x81ÐºÐ¸Ð¹' 'laki-laki' 'à¸Šà¸²à¸¢' 'ç”·æ€§' 'ç”·'
 'Î³Ï…Î¼Î±Î¯Î¹Î±' 'ná»¯' 'mannelijk' 'lass' 'kadÄ±n' 'x\xa0x§x‘x”'
 'nainen' 'mies' 'moteris' 'kallkyn' 'feminin' 'emakumezkoa' 'baineann'
 'gizonezkoa' 'à¸«à¸\x8dà¸´à¸‡' 'vir'
 'áƒ›áƒ“áƒ”áƒ“áƒ\xa0áƒ\x9dáƒ‘áƒ˜áƒ–áƒ˜' 'mÄ™Å¼czyzna' 'É™lÉ\x90É¯É™â„²'
 'muller' 'f3m413' 'vyras' 'å¥³' 'lelaki' 'ì—¬ìž\x90' 'femenino'
 'à®†à®£à¯\x8d' 'Ø£ÙˆØ«Ù‰' 'Ð¶Ñ–Ð¾Ð¾Ñ‡Ð°']
```

Inconsistent unique values and corrupted namespaces indicate that normalisation is required. Data is cleaned in such a way that all the male genders and their associated synonyms from various human languages are defined as “male”. The same is true for “female” genders.

The list ‘undefined_genders’ contains all the strings that could not be translated due to corrupted characters and those fields that are intentionally left blank by the end-user. The set dates to the year 2011, when the only gender option during account creation was to either choose male or female. Thus, people who decided not to identify as either male or female could leave the entry field blank. Because of this, the blank fields are also added to the ‘undefined_genders’ list. In 2014, Facebook changed the account creation process and included non-binary options for gender entry (Griggs, 2014).

```
Genders after data cleaning for "male", "female" and "undefined":
['male' 'female' 'undefined']
```

Value count is performed on the ‘gender’ column:

```
Gender value count:
female       1018337
male          819078
undefined     194243
Name: gender, dtype: int64
```

Value count indicates that most victims defined themselves as female.

5.4. Data Visualisations

Data is visualised by matplotlib pyplot library for Python. The style defined for the matplotlib pyplot graphs is *ggplot*[6].

A pie chart displaying the proportion of victims by gender is presented:

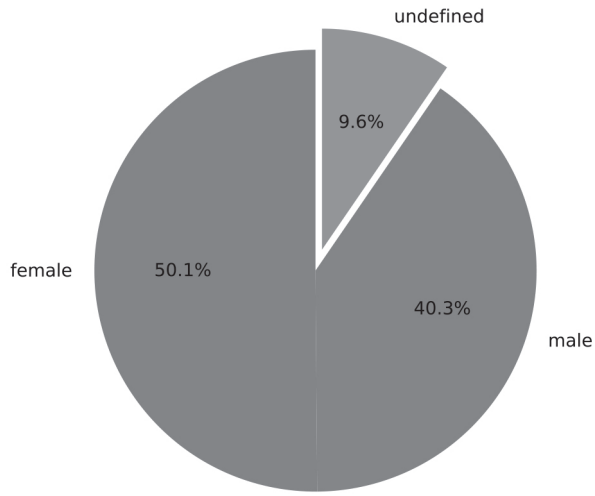

Figure 14 The proportion of victims by gender

The chart indicates that the number of victims who defined themselves as female is greater than the sum of the number of victims who defined themselves as male and the number of victims whose gender remains undefined. Thus, the dominant gender overall is female.

There is no clear conclusion as to why the proportion of females is dominant. This could be due to socio-economic factors or simply because the content could be more attractive to those users who defined themselves as female. Still, there is not enough data to determine causation.

Based on these results, the first hypothesis is formulated:

H1: *Facebook users who defined themselves as female became victims of the malware application to a greater extent than those users who defined themselves as male.*

[6] https://matplotlib.org/stable/gallery/style_sheets/ggplot.html

The first hypothesis will be tested with the Chi-squared[7] test in the remainder of the study.

Another data visualisation is given in the form of a bar chart. A bar chart displaying the total number of victims by country is presented:

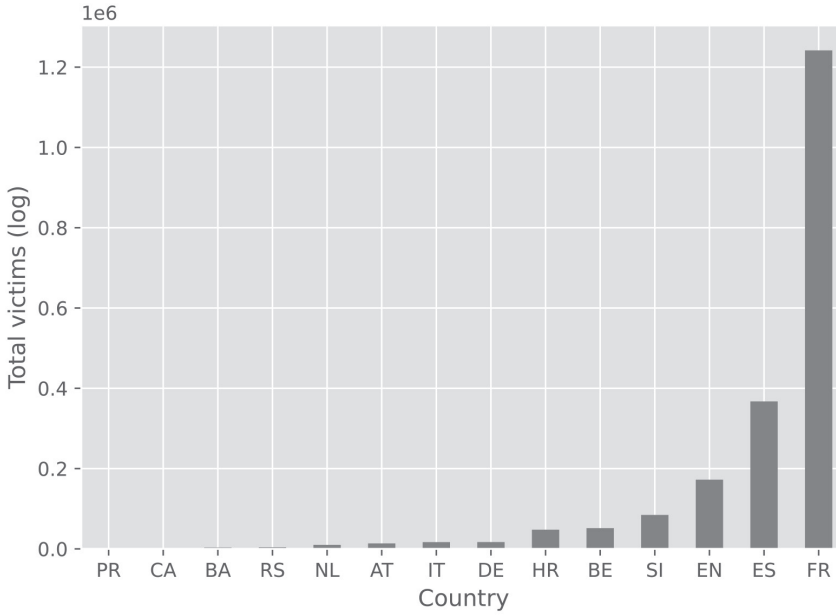

Figure 15 Total victims by country, linear scale

The chart indicates that the number of victims in France and Spain is dominant to such an extent that the number of victims in other countries is hardly visible. The graph is therefore adjusted to the logarithmic scale ordinate and drawn again.

It is unknown whether the 'EN' status indicates persons from the UK in general or England in particular. For the remainder of this study, the author refers to 'EN' status as England in particular.

[7] https://www.statisticshowto.com/probability-and-statistics/chi-square

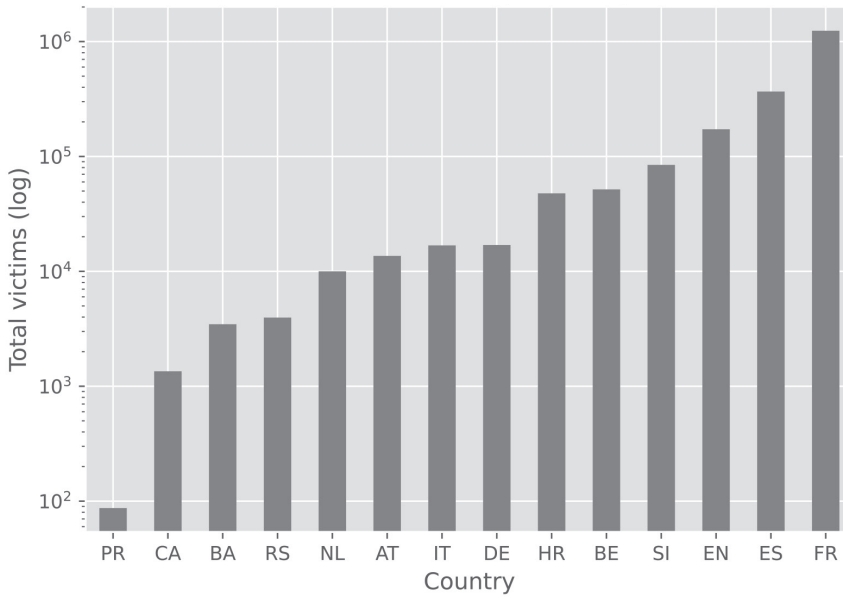

Figure 16 Total victims by country, log scale

Logarithmic scale ordinate provides better visibility of the number of victims in all countries.

In order to get more insight regarding the victims, a bar chart displaying the total number of victims by country, grouped by gender, is presented:

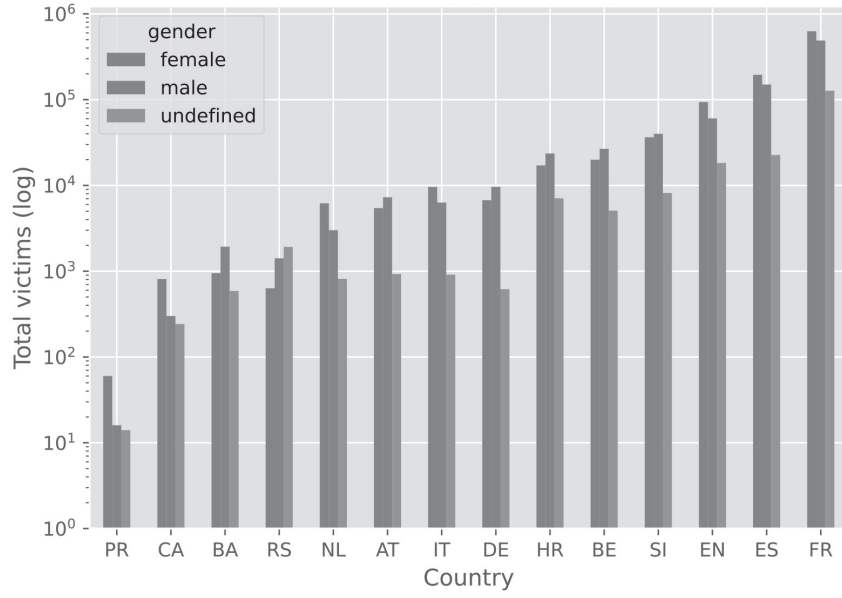

Figure 17 Total victims by country grouped by gender

The graph indicates that the dominant gender in the most developed European countries (western Europe) varies across different countries. The dominant gender in England, Spain, and France is female, while the dominant gender in Germany, Belgium, and Austria is male.

Furthermore, the graph indicates that the dominant gender in the least developed European countries (member countries of former Yugoslavia, such as Bosnia, Serbia, Croatia, and Slovenia) is male across each of these countries.

Altogether, the graph indicates that the dominant gender varies across different countries but that the dominant overall gender is female.

Based on these observations, a second hypothesis is formulated:

H2: *Facebook users who defined themselves as female became victims of the malware application to a greater extent than those users who defined themselves as male across both blocks of analysed countries (block of former Yugoslavia member countries and block of the remaining Western European countries)*.

The second hypothesis will be tested with the Chi-squared rest in the remainder of the study.

In order to create a line chart of the growth of the number of victims by country over time, a growth matrix is formulated in Python. In order to populate the growth matrix, a *for loop* is constructed. A spaghetti plot displaying the growth of the number of victims per country is presented:

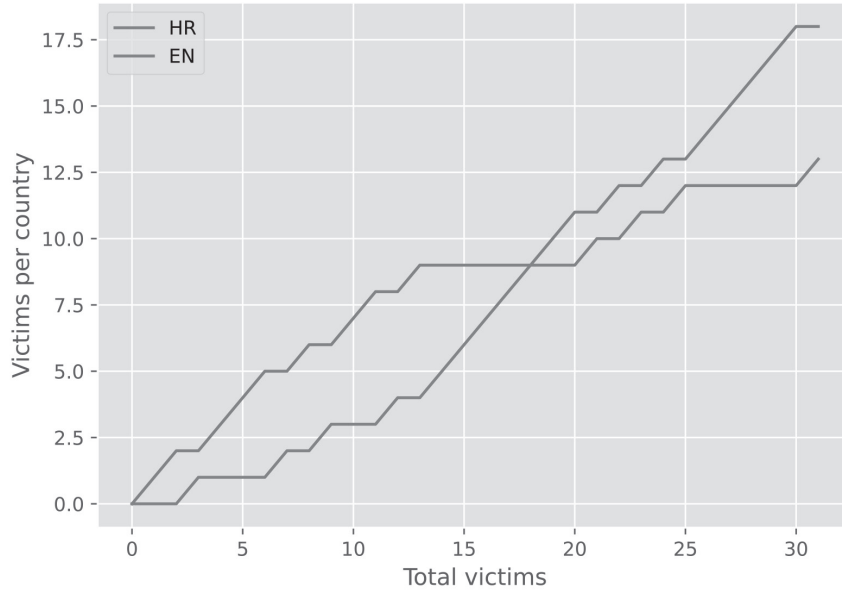

Figure 18 Growth of the number of victims for the first 30 victims

Figure 18 displays the growth of victims for the first 30 victims.

The graph indicates that the first victim comes from Croatia. Two iterations later, the first victim from England appears.

For the first 30 cases, all the victims come from either Croatia or England.

In order to better understand the growth of the number of victims, the second plot is shown:

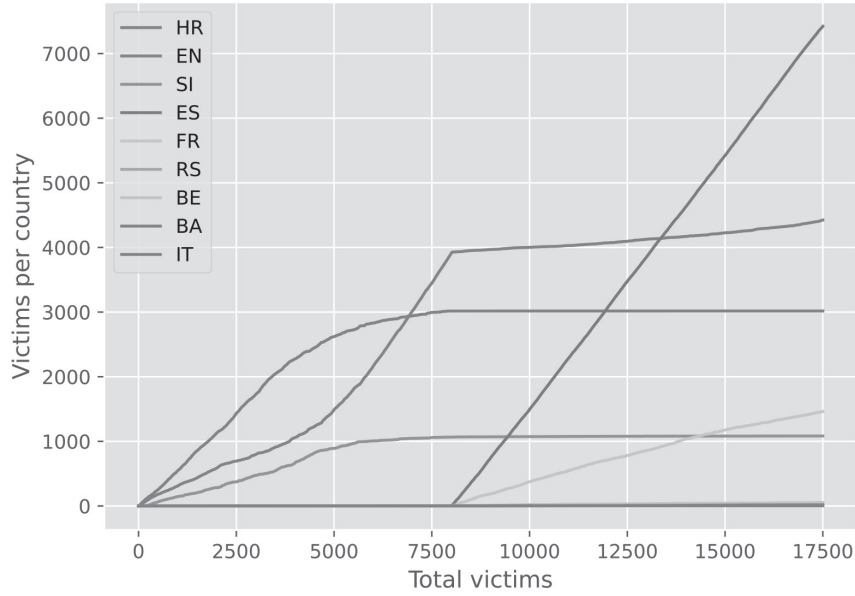

Figure 19 Growth of the number of victims for the first 17500 victims

Figure 19 displays the growth of victims for the first 17500 victims.

The growth of the number of victims in Croatia has flattened out to about 3000 victims. The same is true for Slovenia, with around 1000 victims. The first victim from Spain appears at around 8000 total victims. The growth in Spain seems very strong compared to other countries.

The first victim from France also appears around the same time as the first victim from Spain. This indicates a correlation, but there is insufficient evidence to establish any hypothesis or determine causation.

With around 8000 total victims, there is a sharp decline in the growth of victims in England. This marks the point at which Facebook's anti-malware algorithm detected and banned a certain number of applications from the random rotator served to English audiences. Continuous growth for England indicates that not all the applications were detected, so the total number of victims continues to increase.

Lastly, the spaghetti plot is applied to the entire dataset:

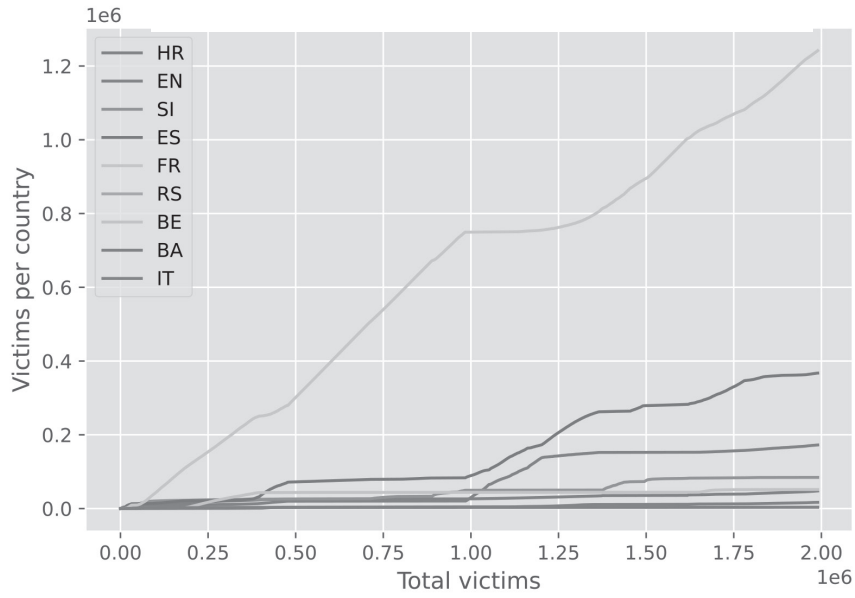

Figure 20 Growth of the number of victims for the entire dataset

Figure 20 displays the growth of victims for the entire dataset.

Figure 19 (earlier) indicated that for the first 17500 victims, the number of victims in Spain is growing at a much higher rate than the number of victims in France. However, when the plot is applied to the entire dataset, the number of victims in France is greater than the sum of the number of victims in every other country. There is no clear conclusion as to why the number of victims in France is dominant.

Second on the list is Spain, followed by England in the third position.

With around 1 million total victims, there is a sharp decline in the growth of victims in France. This marks the point at which Facebook's anti-malware algorithm detected and banned a certain number of applications from the random rotator served to French audiences. However, the attacker managed to regain momentum, as the total number of victims in France nearly doubled after the applications were removed. For the next 250 000 victims, there is hardly any growth in France, as most of the new victims are appearing in England. At around 1.3 million total victims, the attacker regains momentum and manages to attack another 500 000 users.

The last chart indicates how difficult it was for Facebook to protect its users from attackers who used sophisticated methods to bypass the anti-malware algorithm. It also shows the resilience of the attackers and their ability to recover and regain momentum after the detection and removal of malware applications.

The attacker's recovery was achieved mainly through creating new applications and adding them to the random application rotator. The posts published under the removed applications continued to be displayed in the Facebook news feed. Still, without the functional

application behind it, those posts did not lead anywhere, thus cutting the potential victim from the content and redistribution of malware to a further audience.

However, for as long as a single functional application was undetected and remained functional, a connection between the victim and the random application rotator was possible, and with it, the possibility for the victim to get infected and serve the malware to further audiences.

By randomly choosing the applications and serving the audiences with the content delivered in their native language, the attacker managed to manipulate the environment, recover from any malware detection, and continue attacking Facebook users.

5.5. Hypotheses Testing

Based on the stated research, the formulated hypotheses H1 and H2 were tested using the Chi-squared test. Statistical tests were performed at a five per cent level of risk; that is, the probability of occurrence of test statistics less than 0.05 was considered statistically significant.

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Where O is the observed value, and E is the expected value.

Table 1 Gender structure of the analysed victims

Variable	Categories of variable	Frequency	Per cent	Valid Percent
Gender	Female	1018337	50,1%	55,4%
	Male	819078	40,3%	44,6%
	Total	1837415	90,4%	100,0%
	Undefined (missing)	194243	9,6%	
	Total	2031658	100,0%	

The first hypothesis (**H1**), according to which Facebook users who defined themselves as female, became victims of the malware application to a greater extent than those users who defined themselves as male, was tested using the Chi-squared test, where the null hypothesis of equal representation of both genders among victims was tested. The result of the test indicates the rejection of the null hypothesis, i.e., according to the test results, the share of male and female user accounts among victims differs statistically significantly, with a higher proportion of victims who defined themselves as female and a smaller proportion of victims who defined themselves as male.

Thus, the first hypothesis that Facebook users who defined themselves as female became victims of the malware application to a greater extent than those users who defined themselves as male is confirmed. The test result of the first hypothesis is shown in Table 2.

Table 2 The result of testing hypothesis H1

Gender	Observed frequencies	Expected frequencies (under the null hypothesis)	Residuals
Female	1018337	918707,5	99629,5
Male	*819078*	*918707,5*	*-99629,5*
$\chi^2_{(1)} = 21608{,}700$			
$p = 0{,}000$			

$\chi^2_{(df)}$ = Pearson's Chi-Square test statistic with associated degrees of freedom; p = probability of occurrence of given test statistics under the null hypothesis (asymptotic statistical significance)

Table 3 The result of testing hypothesis H2

Variable	Categories of variable		Gender		Significance of test statistics
			Female	Male	
Country	Countries of ex YU	OF	55059	66848	$\chi^2_{(1)} = 5548{,}045$ $p = 0{,}012$
		EF	67550	54358	
	Other European countries	OF	962410	751916	
		EF	949920	764407	

OF – Observed frequencies; EF – Expected frequencies; $\chi^2_{(df)}$ = Pearson's Chi-Square test statistic with associated degrees of freedom; p = probability of occurrence of given test statistics, under the null hypothesis (asymptotic statistical significance)

The second test indicates that the share of users who defined themselves as male is dominant in Eastern European countries (member countries of former Yugoslavia). In contrast, the opposite is true for countries of Western Europe.

6. DISCUSSION

6.1. The Gender Gap

The research conducted here shows that most of the victims of these cyberattacks defined themselves as female. There is no clear conclusion as to why the dominant gender is female. This could be due to various cultural, psychological, socio-economic, and/or other factors affecting the genders, but without detailed data describing the victim's age, education, income, health state, and other factors, it is not likely that the causation can be determined. We can speculate, however, that the advertised content of the spam post was simply more interesting to those users who defined themselves as female. The content was titled "*See why this woman killed herself after the wedding*", which could have attracted more users who defined themselves as female due to the use of the word "woman", potentially signalling the target audience.

Secondly, according to the research published by Servaty and Weber (2011), when asked to rate how strongly they agree or disagree with a given statement, 50% of female participants strongly agreed with the statement "*One of my goals is to be married*". In comparison, the percentage of males who strongly agreed with the same statement was 48.5%. Furthermore, when the given statement was "*The principal purpose of marriage is love*", 45.6% of female participants agreed with the statement, while the percentage of males who agreed with the same statement was 34.8%. Finally, when the given statement was "*Personal fulfilment is a purpose of marriage*", 42.6% of female participants agreed with the statement, while the percentage of males who agreed with the same statement was 36.4%. This research signals that females are more likely to be interested and more likely to have stronger opinions in matters regarding marriage than males, which could be another reason for the higher share of users who defined themselves as a female in the pool of victims described in this study.

In the countries of Eastern Europe (member countries of the former Yugoslavia), more victims defined themselves as male than female, which shows a contrast from most of the countries of Western Europe. According to the GDP per capita projection for the year 2021, the four countries with the greater share of victims who defined themselves as male (Croatia, Bosnia, Serbia, and Slovenia) are also the four countries with the lowest GDP per capita of all the countries[8] covered by this study. This signals a potential relationship between the GDP per capita and the representation of genders in the total number of Internet users, but without the education and income data available in the dataset, it is not likely that causation can be established.

6.2. Suggestions for Research Improvements

The greatest deficiency of this study is the lack of data that would provide insight into the victim's age, education, and income. If age data were available, victims could be grouped by age, which could provide insight into the victim's life experience or the overall time they could have spent online or on Facebook in particular. If education and income

[8] https://en.wikipedia.org/wiki/List_of_sovereign_states_in_Europe_by_GDP_(PPP)_per_capita

data were available, victims could be grouped by their income, social status, and/or highest completed education, which could provide insight into the victim's general computer-use skills, knowledge of cybersecurity, life interests (Bagić Babac, 2023), available leisure time (Puh and Bagić Babac, 2023), money-management (Puh and Bagić Babac, 2023b), etc.

Suggestions for further research include gathering data that would describe the victim's age, education, and income status, where a more detailed model of a victim could be developed, hopefully describing the typical characteristics of victims. If such a model were developed, researchers and educators could focus on substituting or eliminating the behaviours that could lead to an increased risk of becoming a cyberattack victim.

REFERENCES

1. Albladi, S. M. and Weir, G. R. S., (2020). *Predicting individuals' vulnerability to social engineering in social networks*, Cybersecurity, 3(7). https://cybersecurity.springeropen.com/articles/10.1186/s42400-020-00047-5
2. Bagić Babac, M. (2023). *Emotion analysis of user reactions to online news*, Information Discovery and Delivery, 51(2), 179-193. https://doi.org/10.1108/IDD-04-2022-0027
3. Bagić Babac, M. and Jevtić, D. (2014), *AgentTest: A specification language for agent-based system testing*, Neurocomputing, 146, 230-248 DOI:10.1016/j.neucom.2014.04.060
4. Catanese, S., De Meo, P., Ferrara, E., Fiumara, G., Provetti, A. (2012). *Extraction and Analysis of Facebook Friendship Relations*, Computational Social Networks: Mining and Visualization. Springer Verlag, DOI: 10.1007/978-1-4471-4054-2_12
5. Citigroup (2016). *The Cyber Underground: Facilitating and Monetizing Cyber Attacks*, Citi Bank Online, https://www.citibank.com/tts/sa/emea_marketing/docs/Facilitating_Monetizing_Cyber_Attacks.pdf - accessed 1 June 2021
6. Cvitanović, I. and Bagić Babac, M. (2022). *Deep Learning with Self-Attention Mechanism for Fake News Detection*, Combating Fake News with Computational Intelligence Techniques / Lahby, M., Pathan A. S. K., Maleh, Y., Yafooz W.M.S. (eds.). Switzerland: Springer, pp. 205-229. doi:10.1007/978-3-030-90087-8_10
7. DeLiema, M., Fletcher, E., Kieffer, Christine N., Mottola Gary R., Pessanha, R. and Trumpower, M. (2019). *Exposed to Scams: What Separates Victims from Non-victims?*, Finra Foundation.
8. Fatokun, F.B., Hamid, S., Norman, A. and Fatokun, J.O. (2019). *The Impact of Age, Gender, and Educational level on the Cybersecurity Behaviors of Tertiary Institution Students: An Empirical investigation on Malaysian Universities*, *In International Conference Computer Science and Engineering, Journal of Physics: Conference Series,* 1339, December 2019. IOP Science, https://iopscience.iop.org/article/10.1088/1742-6596/1339/1/012098 - accessed 13 June 2021
9. Griggs, B. (2014). *Facebook goes beyond 'male' and 'female' with new gender options*, CNN Bussiness, February 13, 2014. https://edition.cnn.com/2014/02/13/tech/social-media/facebook-gender-custom/index.html accessed 8 June 2021

10. Han, C. and Dongre, R. (2014). *Q&A. What Motivates Cyber-Attackers?* Technology Innovation Management Review*,* 4(10), Tim Review, https://timreview.ca/article/838 - accessed 6 June 2021
11. Lipovac, I. and Bagić Babac, M. (2023). *Developing a Data Pipeline Solution for Big Data Processing*, International Journal of Data Mining, Modelling and Management (in press)
12. Liu, W. and Zhong, S. (2017). *Web malware spread modelling and optimal control strategies*, Scientific Reports*,* 7(42308). https://www.nature.com/articles/srep42308 - accessed 13 June 2021
13. Norris, G., Brookes, A. and Dowell, D. (2019). *The Psychology of Internet Fraud Victimisation: a Systematic Review*, Journal of Police and Criminal Psychology*,* 34, 231–245.
14. Redmiles, E. M., Chachra, N. and Waismeyer, B. (2018). *Examining the Demand for Spam: Who Clicks?*, CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 212, 1-10 April 2018. ACM.
15. Servaty, L. and Weber, K. (2011). *The Relationship between Gender and Attitudes towards Marriage*, Journal of Student Research*,* https://minds.wisconsin.edu/handle/1793/53228 – accessed 13 June 2021
16. WP Engine (2020). *Unmasked: What 10 million passwords reveal about the people who choose them* , WP Engine Resource Center, Security. November 20, https://wpengine.com/resources/passwords-unmasked-infographic -accessed 20 May 2021
17. Facebook (2020). *Facebook Platform Terms*, *Facebook for Developers* https://developers.facebook.com/terms/ - accessed 5 June 2021
18. Pennington, N. R. D. (2010). *No consequences: an analysis of images and impression management on Facebook*, M.A. thesis, Kansan, Manhattan : Kansas State University, https://core.ac.uk/download/pdf/5168017.pdf - accessed 1 July 2021
19. Knowak, K. (2017). *How Many Facebook Friend Requests Can You Send Per Day?*, The Autolikes Blog, https://autolikes.com/blog/2017/11/facebook-friend-requests-day - accessed 13 June 2021
20. Schiffer, Z. and Statt, N. (2019). *Facebook suspends 'tens of thousands' of apps from 400 developers over improper data use*. The Verge. https://www.theverge.com/2019/9/20/20876021/facebook-developers-apps-suspensions-data-privacy-cambridge-analytica - accessed 3 June 2021
21. Koolwal, A. (2016). *Introducing the Events From Facebook App facebook.com,* https://about.fb.com/news/2016/10/introducing-the-events-from-facebook-app/ - accessed 4 June 2021
22. Tankovska, H. (2021). *Number of monthly active Facebook users worldwide as of 1st quarter 2021*, Statista. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide - accessed 3 June 2021

23. Peltonen, E., Lagerspetz, E., Hamberg, J., Mehrotra, A., Musolesi, M., Nurmi, P. and Tarkoma, S. (2018). *The hidden image of mobile apps: geographic, demographic, and cultural factors in mobile usage*, MobileHCI '18: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, Article No. 10, ACM Digital Library, https://doi.org/10.1145/3229434.3229474
24. Nakita (2011, February 5). Luda Svadba / Crazy Wedding Video. YouTube. https://www.youtube.com/watch?v=ujjukK1AykI accessed 1 April 2023
25. Puh, K. and Bagić Babac, M. (2023a). *Predicting sentiment and rating of tourist reviews using machine learning*, Journal of Hospitality and Tourism Insights, 6(3), 1188-1204. https://doi.org/10.1108/JHTI-02-2022-0078
26. Puh, K. and Bagić Babac, M. (2023b), *Predicting stock market using natural language processing*, American Journal of Business, 38(2), 41-61. https://www.emerald.com/insight/content/doi/10.1108/AJB-08-2022-0124/full/html
27. Šandor, D. and Bagić Babac, M. (2023). *Sarcasm detection in online comments using machine learning*, Information Discovery and Delivery (in press). https://doi.org/10.1108/IDD-01-2023-0002
28. Whitty, M. T. (2019). *Predicting susceptibility to cyber-fraud victimhood*, Journal of Financial Crime, 26(1), 277-292.

Sažetak

Hrvoje Čemeljić, Marina Bagić Babac

Sprječavanje sigurnosnih incidenata na društvenim mrežama: analiza širenja štetnog sadržaja putem aplikacija

Ova studija opisuje širenje štetnog sadržaja putem zlonamjernih aplikacija implementiranih na društvenoj mreži Facebook. Opisan je sigurnosni incident, s naglaskom na motive i svrhu kibernetičkih napada, kao i opis napadača i žrtava. Detaljno su opisani alati napadača, kao i tehnike koje je koristio kako bi zahvatio mnoge potencijalne žrtve, inficirao ih zlonamjernim programima, monetizirao žrtve i prikrio tragove napada. Izvršena je analiza podataka na skupu podataka koji sadrži informacije o više od dva milijuna žrtava. Fokus analize je modeliranje širenja zlonamjernih programa te određivanje omjera žrtava na temelju spola i zemlje podrijetla. Studija pokazuje znatnu statističku razliku među žrtvama napada na temelju njihova spola.

Ključne riječi: *Facebook, kibernetička sigurnost, kibernetički napadi, zlonamjerne aplikacije, analiza podataka.*