Ransomware: How attacker's effort, victim characteristics and context influence ransom requested, payment and financial loss

Tom Meurs
IEBIS
University of Twente
Enschede, Netherlands
t.w.a.meurs@utwente.nl

Marianne Junger
IEBIS
University of Twente
Enschede, Netherlands
m.junger@utwente.nl

Erik Tews
EEMCS
University of Twente
Enschede, Netherlands
e.tews@utwente.nl

Abhishta Abhishta
IEBIS
University of Twente
Enschede, Netherlands
s.abhishta@utwente.nl

***Abstract*—In recent years, ransomware attacks have led to disastrous consequences for victims, not just due to the payment ransom amount but also due to the recovery costs associated with these attacks. So far only a few empirical studies have analysed the financial impact of ransomware attacks. This study aims to understand the expected financial gains for attackers and financial losses of victims after a ransomware attack. To do so, we build a dataset based on 453 ransomware attack investigation reports in the Netherlands reported to the Dutch Police between 2019 and 2022. Using rational choice model of crime (RCM) and crime scripting we hypothesise that the effort of an attacker, victim characteristics and context variables influence not only the ransom requested by an attacker but also the financial losses reported by victims. We use generalised linear models to evaluate and quantify this influence. Our results show that attacker's efforts such as using Ransomware-as-a-Service (RaaS) and victim characteristics such as industry sector, contribute to the ransom requested by attackers and financial losses reported by victims. We also show that the availability of recoverable backups explains the likelihood of victims paying the ransom. A limitation of the present study is the interpretation of the results due to selection bias of victims willing to report to the police. Despite this limitation, we argue that our methodology and results lay the groundwork for future large-scale empirical studies and add to our understanding of attacker and victim behaviour.**

***Index Terms*—ransomware, financial loss, payment, rational choice model, routine activities theory**

I. Introduction

Europols annual Internet Organised Crime Threat Assessment Report mentions ransomware as top priority [20]. Ransomware (ransom software) is a subset of malware designed to restrict access to a network, system or data until a requested ransom amount from the attacker is paid [5]. Financially motivated attackers see large sums of ransom paid for victims to decrypt and retrieve their systems and files during a ransomware attack. The paid ransom is often only a small part of the financial loss for the victim after a criminal attack [2], [3]. Since the IT infrastructure is down, business continuity is often a problem. Therefore, downtime could be an important factor for financial loss. Furthermore, recovery costs, like buying new hardware and software and hiring specialists to clean and recover the systems, could also be an important contributor to financial loss.

Usually, the aim of a ransomware attack is to obtain a ransom, however, using stolen data from ransomware, attackers can also accomplish various other goals [39]. [39] also describes how stolen data can be used to blackmail the victim: (1) by incrimination, for example by reporting the victim to data protection authorities, (2) by threatening with reputational damage/lost revenue by exposure of sensitive data on the dark web, leading to loss of trust of customers and additional victimisation, (3) by threatening with exposing intellectual property, and (4) by fear of humiliation, for instance by exposing embarrassing information about customers or employees [39]. This data can also be used to derive information to support new attacks, e.g., selling email addresses for phishing campaigns [44].

The rational choice model (RCM) of crime [33], [34] assumes that attackers and victims are rational actors, who weigh the costs of their actions against the benefits in order to make a rational choice. It should be noted that RCM defines the weighing of costs and benefits as rational and this assumption helped in understanding behavioural decisions by malicious actors in different types of crimes, like car-theft [64] and burglary [65]. Using RCM, we hypothesise that increase in effort put in by the attacker in an attack increases their ransom demands. At the same time, victims who are not prepared for a ransomware attack (e.g. do not have appropriate back ups) are more likely to pay the ransom.

In this study our goal is to empirically determine the factors that explain the expected financial gains for ransomware attackers and financial losses of victims after a ransomware attack. Therefore, we state main research question as follows: *what are the factors that contribute to the ransom requested by attackers and financial loss of victims?* To answer this question, we focus on three sub-questions:

1) Which factors influence the amount of ransom requested by attackers for the decryption key?
2) Which factors influence the likelihood that victims will pay the ransom?

3) Which factors influence the financial losses of victims reported to the police after an attack?

We analyse 453 Dutch Police Investigation reports of ransomware between January 2019 and July 2022 to collect information on the effort invested by attacker, characteristics of the victim (e.g., yearly revenue, industry sector) and the contextual information regarding the attack (e.g., year and season). To systematically include the factors that contribute to attacker's effort we propose a crime script for a generic ransomware attack. Using generalised linear models (GLM) we test the impact of factors related to rational choice model of crime on the demanded ransom and the likelihood of victims to pay ransom. Our key contributions are:

1) We annotate and analyse 453 Dutch Police Investigation reports describing different ransomware attacks.
2) We show that the amount of effort put by the attacker and yearly revenue of the victim influence the amount of ransom requested by the attacker;
3) We find that along with cost & attacker's effort related variables, the payment of the ransom is determined by the victims being able to recover the encrypted data with backups after an attack;
4) We evaluate the factors that influence the financial loss reported by the victim. We find that factors such as ransom paid, the yearly revenue of the victim and use of RaaS (Ransomware-as-a-Service) by an attacker are statistically significant factors in determining the financial loss reported by a victim after an attack.

The structure for this paper is as follows: We discuss past literature related to use of cyber crime theories and evaluation of ransomware attacks in §II. We introduce the proposed crime script and state our hypotheses in §III. Then, we explain composition of our dataset and methodology for analysis in §IV. Finally, after showcasing our results in §V, we discuss our conclusions and future work in §VI and §VII.

II. Background Work

Previous work on ransomware has focused mostly on the technical aspects of ransomware [4], [5], [8]. Technical aspects include forensic analysis [47], network detection of command-and-control communication during ransomware attacks [4] and reverse engineering [48]. Several countermeasures have been proposed [47], [48]. E.g., [4] mentions taking advantage of weak encryption techniques used by attacker and improving user awareness to prevent phishing attacks.

A current trend is to study ransomware from other scientific fields, like crime science and economics [5], [6]. Crime science research on ransomware has focused mostly on qualitative impact on victims [1], [36]. [1] surveyed 50 organizations in the UK and North America and studied the factors contributing to the severity of an attack, measured by asking how severe an attack was: low, medium or very severe. The authors did not find a difference in severity between ransomware attacks through phishing versus using exploits as an initial access vector. However, the targeting of victims resulted in more severe attacks than the opportunistic choice of victims.

Economic research on ransomware takes a more theoretical approach [10]–[12], [22]. [10] assumes that attackers want to maximize profit and therefore request a ransom which is the trade-off between the probability of a victim paying and maximizing the ransom and therefore profit. One of their results is estimating a demand function of buying the decryption key, where a percentage of the victims would pay a certain price or ransom for returning their files. The authors argue that attackers could maximize profits by estimating the demand function as realistically as possible and subsequently set a ransom which maximizes profits. In this case, it would be beneficial for the attacker to research the victim to estimate the willingness to pay. Their seems to be anecdotal evidence that this happens in practice [37]. However, the authors do not mention which specific factors explain a high or low willingness to pay, except for the ransom requested [10].

Other studies used game-theoretical models to understand willingness to pay [12], [22]. [12] found data exfiltration to be an important determinant for willingness to pay. Attackers extort the victim by releasing sensitive information online if they do not pay. This gives the victim an incentive to pay the ransom to prevent publication, even if they could recover encrypted files from a backup. [21] explain why victims would pay or not: a cost-benefit analysis of victims between the financial costs of not paying, which is related to downtime costs, and ethical concerns of paying criminals. The author mentions that the most important factor for victims to pay is having recoverable backups or not [21].

In sum, both crime science, as well as economic research, emphasized that attackers' behaviour can be described with a rational choice model of crime (RCM) [33], [34]. The costs and benefits calculations are made by attackers to determine the ransom to request to maximize profit.

Besides RCM, crime science also proposed opportunities as a factor that guides attackers' behaviour [34]. Opportunities are characteristics of the targets. Target or victim characteristics influence costs and benefits calculations. Target or victim characteristics influence costs and benefits calculations. For example, in line with routine activity theory [51], [52], we assume that wealthier victims, that is victims with a high yearly revenue and a large staff, constitute more attractive victims as they are likely to be able to pay a higher ransom. Also, victims with a cyber insurance are relatively attractive as they may not care to pay. The main asset in order to avoid paying for victims is to have a backup that can be restored easily. What may help reduce the damage is hiring an incident response company to avoid paying or pay less. Engaging in lengthy negotiations may also help reducing the ransom that has to be paid. Having one's infrastructure in the cloud also helps to reduce the final ransom. Taken together, victim characteristics could influence attackers' behaviour.

Next to attackers' effort and the context of the attacks might also influence the ransomware attacks [23], [29], [32]. Besides focusing on wealthy victims in order to be able to request large ransoms, other aspects may play a role. As companies become more and more dependent on their digital assets for their busi-

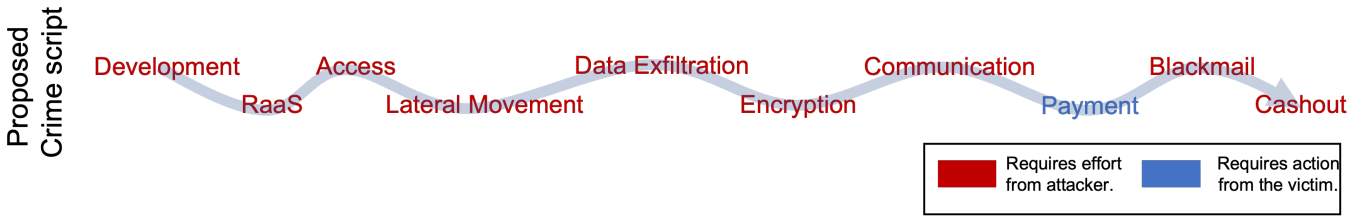

Fig. 1. The steps of the crime script of a ransomware attack used in this study to structure the data.

ness continuity this may lead to an increasing vulnerability, which may lead to a trend of requesting larger ransoms over the years and accordingly, a 'willingness' or need to pay larger ransoms over the years [23]. Second, cybercriminals might be more willing to attack in different seasons [32]. For example, [29] found that seasonality influenced fraud against businesses. This might also occur within the ransomware landscape.

In conclusion, the literature on ransomware attacks has been mostly based on theory, relatively small samples or more qualitative descriptions of ransomware attacks. There is little quantitative empirical research on the risk factors/determinants ransom requested by attackers and financial loss reported by victims after a ransomware attack [35], [39]. This is the focus of the present study. To structure the data and analyse the concepts of opportunity and effort, we propose a crime script of ransomware in the next section.

III. Proposed Crime Script and Hypotheses

As described in the previous section, we hypothesize that ransomware attacks are the result of crude cost-benefit calculation by attackers and their assessment of where the good opportunities lie. To understand the costs, risks and opportunities, it is important to consider the ransomware crime script. From the field of crime science, crime scripting is a way to systematically study the procedures, actions and decisions when performing a crime [26]–[28]. Similarly, in computer science several authors described a kill chain in which the various steps on performing an attack were described [38], [39]. Previous research on ransomware described taxonomies that sometimes included a series of steps [4], [5], [40] as well as different actors and roles [35]. In the present study we propose to describe ransomware as a simplified step-wise process: ransomware is a complex crime involving many steps, often involving a group that probably comprises several members and sometimes also involves collaboration with other groups. Based on insights from previous research [23]–[25], [60], we propose the following global ransomware script (see Figure 1):

1) Development: To start with, it is important to organise the infrastructure and develop the malware beforehand [45]. The infrastructure is needed to deliver the malware and to obfuscate network traces from the system of the victim to the attacker [24], [45].

2) RaaS: RaaS and collaboration with other groups. When individuals or groups lack expertise they can make use of Ransomware-as-a-Service (RaaS). In practice, this means hiring the ransomware from other cyber criminals [35], [46]. The term affiliate has been used to describe the actor hiring the ransomware. RaaS enables affiliates with relatively low-technical skills to use advanced ransomware and this makes the attack much easier to launch [4]. RaaS was described as a way to 'democratize crime' [35]. The advantage for the affiliates is, obviously, that it becomes much easier to execute ransomware attacks: all actors involved in an attack could specialize in a specific part of the attack. For example, obtaining credentials from a victim's network or developing malware [46]. However, extra effort may lie in coordinating their work with the RaaS developer and the possibility to have to share a part of the profit. Each actor, ransomware developer or affiliate, do what they can do well, and do not need to do the other party's work. Accordingly, we believe it is a reasonable hypothesis that, overall, RaaS requires less total work for the involved actors.

3) Access: Gain access to a victim's computer or network and maintain that access. To gain access to the victim's system, attackers need to distribute the ransomware. Reference [5] described how this is usually done. Mostly, attackers send a phishing email that contains a malicious file or a link (33%) or they sand spam (8%). Other options are malicious apps, to infect mobile phones (13%); drive-by-download e.g., malicious advertisements (10%); exploit kits (15%) or a Remote Desktop Protocol (RDP: 8%). Vulnerabilities in the victim's platform such as in operating systems, browsers, or software can also be used by ransomware attackers as infection vectors (10%).

4) Lateral movement: It is moving to other computers on the network with the goal to get an impression of the files and gain control over the entire network.

5) Data exfiltration: Although many groups state they exfiltrate data, probably to put pressure on the victim, most ransomware groups do not actually do this. Data exfiltration is a big risk for the victim, as, for instance, their data may end up on the dark web on a 'sucker's list', be sold, or become visible on the open web for everyone to see [12], [39]. Data exfiltration is considered to take more effort than no data exfiltration.

6) Encryption: Performing encryption of the victims files is of course, key to the entire process.

7) Communication: The attackers need to communicate and possibly negotiate with the victim. They also need to provide

payment credentials and determine the size on the requested ransom. To this end the attackers can send a ransom note to the victim: they want to first have contact with the victim before informing them what ransom requested is. This gives them the opportunity to change the ransom, depending on victim characteristics like yearly revenue [10], [37]. A personalized ransom note is considered to be more effort for the attackers than a standard ransom note for all victims. Furthermore, the mode of communication also influences the attackers effort: some attackers communicate with their victim through e-mail, others use a self-made *TORchat* application. Using a self-made TOR application requires more work on the attackers' side. Although within the RaaS ecosystem affiliates often do not need to make the *TORchat* themselves, we would hypothesize that the overall effort increases. Developers of the ransomware might ask for larger ransoms requested by affiliates due to their extra effort put into building the TOR-chat.

8) Payment: At this stage the victim needs to think about paying or not paying. If the victim does not want to pay, for instance because he has a good backup, the attack could stop here. But victims often do not have a useable backup. According to [43] restoring backups is often difficult: 85% fail during restoration attempt. Consequently, at this stage the victim usually starts communicating with the attackers about the ransom. The victim may be willing to pay, but think the ransom is too high, sometimes he is not allowed to pay the ransom, such as some public organizations. To that end, the victim might engage an incident response company that helps negotiating and the payment of the ransom. The ransom after negotiations may depend not only on the requested ransom, but also on the negotiating skill of the victim and/or the incident response company that the victim hired. The experience of the Dutch police is that attackers have an incentive not to take too much time to negotiate: longer negotiations may lead to a lower final ransom and they want to have their money quickly. Asking a ransom that is unrealistically high may increase the negotiation time [13].

9) Blackmail: Different additional extortion methods can be used to put additional pressure on the victim: perform DDoS attacks on the victims website and/or calling or e-mailing clients or employees of the victim's company [49]. It is important to note that the publication of data on a leakpage is also a type of blackmail, but in this study is categorized as *5) Data exfiltration.*

10) Cash-out: Getting the money, laundering it through different mixers or money mules [7]. Additionally, provide the decryption keys to the victim and possibly helping the victims with decryption of their files.

We emphasize that this crime script is a rough description of a ransomware attack and serves the purpose of this study. Further research might generalize this crime script to include more different types of ransomware attacks which are outside of the scope of this paper.

The crime script presented above is a brief overview of the steps of an complete version of a ransomware attack: not all attacks include all steps. Some ransomware groups are known to perform some of the steps described above. For example, eCh0raix is strain which targets solely Network Attached Storage (NAS) devices [31]. The group(s) behind this strain are known to be RaaS, so affiliates can buy the ransomware in exchange for a share of the profits. Furthermore, these attacks are characterized by only encrypting the NAS device and then leaving a ransom note with a fixed ransom for all victims and bitcoin address in exchange for the decrypter keys. So, from the crime script, only steps 1), 2), 3) and 9) are performed. Similarly, the group(s) behind the ransomware strain Conti [30], is known for being RaaS, and perform almost all steps of the mentioned crime script, except for step 9) Blackmail. As we will illustrate in this study, eCh0raix requests smaller amounts of ransom than Conti, as expected from our reasoning above.

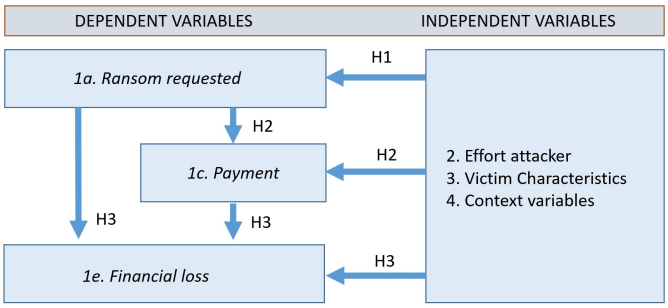

Fig. 2. Hypothesis within this study. Effort of attacker, victim characteristics, context variables determine ransom requested (H1). Combined they influence whether a victim pays (H2). Financial loss of a victim is determined by ransom requested, paid, effort of attacker, victim characteristics, and context variables (H3).

Based on the Rational Choice Model of crime (RCM) and the crime script, it is assumed that increasing the costs of an attack must be balanced by larger rewards and/or easier opportunities, and/or smaller risks, otherwise, attackers will not be interested in investing more time and effort. We therefore hypothesize, that when more effort is put into the attack, the result should be a larger ransom requested and larger financial loss for victims. Specifically, we hypothesize (see Figure 2):

1) The ransom requested (RR) is the result of a costs-benefits calculation by the attackers, considering opportunities and context (H1). It is expected that more attackers effort leads to larger RR.
2) The decision to pay the ransom is the result of the RR and the costs and benefits of the victim (H2). It is expected that victims who have back-ups and attacks where data has been exfiltrated, leads to larger probability of paying.
3) The losses by the victim are the result of RR, payment and attackers' effort, victim characteristics and context variables (H3). It is expected that large RR, effort by attackers, large companies and payment lead to larger financial loss after a ransomware attack.

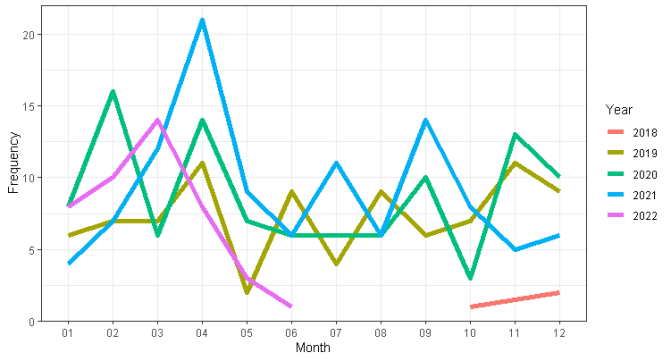

Fig. 3. Frequency of ransomware monthly attacks based on date of encryption in reports. 3 reported attacks were from 2018, while nearly a 100 attacks are from 2019, 2020 and 2021 each. 44 attacks are reported since beginning of 2022.

IV. Data and Methodology

Between 1 January 2019 and 1 July 2022 453 ransomware attacks were reported to the Dutch Police. Attacks were collected by searching the police file systems for the keyword 'ransomware'. Subsequently we collected and coded the data using the variables shown in Table I below. We show the step-by-step methodology in Figure 4. Of these 453 police investigation reports, 13 were attempted ransomware attacks where no files were encrypted, in 6 cases no ransomware was found and 81 ransomware attacks on individuals were outside the scope of this study. Combined, investigation reports for 353 attack remain in our dataset and were used for further analysis. Figure 3 shows the monthly distribution of reports. [41], [43] state that the number of ransomware attacks have increased substantially in the last few years. In our data only 3 of the reported attacks were from 2018, while nearly 100 attacks are from 2019, 2020 and 2021 each. Since we use reports made after January 1st, 2019, we do not see a substantial change in the number of reported attacks in these three years.

Next, we describe the variables coded in police investigation reports. The three **dependent variables** in our study (see Table I) are:

1a. *Ransom requested:* The ransom attackers request for the decryption of the victims files, is a good estimation of how much financial gain they hope to make with the attack [10]. This is the ransom requested in the beginning of the ransomware process, before the negotiations (in euro). The reason is twofold: 1) the ransom after negotiations also depends on the negotiating skill of the victim and/or the incident response company that the victim hired and 2) attackers have an incentive to not take too much time to negotiate, because they want to have their money quickly. Asking a ransom far from the financial gain they hope to make, might increase the negotiation time [13].

1c. *Payment:* This variable is whether victims would pay or not (categories: yes = 1 / no = 0). In our data set, 21% of victims paid. This is different from the willingness to pay [14]. Some victims might be willing to pay, but think the ransom is too high or they are not allowed to pay the ransom, like public organizations.

1e. *Financial loss:* This is the total financial loss reported by the victim (in euro). Some victims specified different aspects of the costs, e.g., repair costs, reputation costs, liability, and payment of ransom. Nevertheless, most victims only gave a rough estimate of the total costs.

These three dependent variables, are log-transformed. This transforms the non-linear distribution to get an approximately normal distributed variable, as is common in social-empirical studies [19]. The logarithm base 10 is chosen to increase the readability of figures.

The **independent variables** in this study (See Table I) are:

1b. *Ransom requested end negotiations:* To understand if ransom requested influences payment, it is important to consider the amount of ransom which was requested after negotiations (in euro), since this is the amount the victim needs to pay.

1d. *Ransom Paid:* To study the factors influencing financial loss, the ransom paid to the attackers has been used as a dependent variable (in euro). This was constructed as a function multiplying payment (*1c.Payment*) and final ransom (*1b.Ransom requested end negotiations*).

2. Effort attacker. To measure effort information was collected on several variables.

2a. *Data exfiltration:* Exfiltrated of data measured whether data from the victim were exfiltrated (categories: yes = 1 / no = 0). Although many groups state they exfiltrate data, probably to put pressure on the victim, most ransomware groups do not. We reported a confirmed data exfiltration when analysis of the network logs has been performed and unusual large amount of data uploading was found or when the victims data has been published on a leak page and the data is identified of being from the victim. Data exfiltration is considered more effort than no data exfiltration.

2b. *Targeted ransom note:* We noted whether the criminals wanted to first have contact with the victim before informing them what ransom they requested, which we define as targeted ransom note (categories: yes = 1 / no = 0). Yes means that first contact with the attackers was required to obtain information about the ransom. No means that the ransom was stated on the ransom note. A personalized ransom note is considered more effort than a standard ransom note for all victims. To our knowledge did a personalized ransom note not yet lead to identification of the attackers.

2c. *RaaS:* Collaboration with other criminals, measures whether the attackers made use of Ransomware-as-a-service (RaaS) [11] or whether they collaborated with other groups to perform the attack (categories: yes = 1 / no = 0). RaaS is considered to be more effort than groups who do not perform RaaS.

2d. *Strain:* The name of the ransomware strain found on the victims encrypted files. Often this is the extension. The

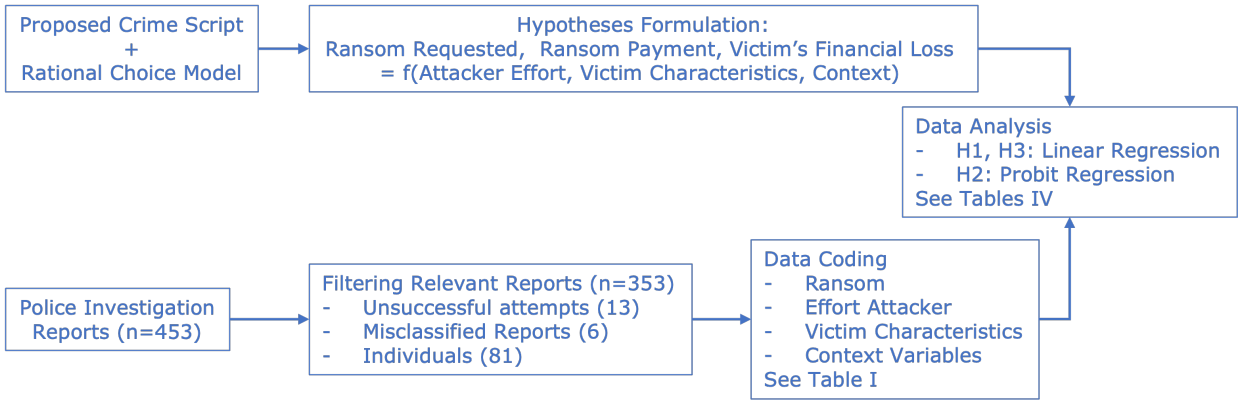

Fig. 4. Methodology used to analyse police investigation reports.

attacks from a specific strain that executed the attack were included if more than 5 attacks were observed, the rest was aggregated to the variable 'Others'. This is due the sensitivity of the data. We assume that groups behind strains vary in the amount of effort used in attacks, and therefore might also vary in the required ransom. This variable accounts for all variance due to factors not labelled in this study but are different between strains or groups.

2e. *NAS:* Network Attached Storage measures whether attackers targeted a Network Attached Storage device (NAS, categories: yes = 1 / no = 0).

2f. *Access:* What type of access was used to infiltrate the victims network (categories: exploit/phishing/different).

2g. *Blackmail:* Whether attackers contacted the victim or clients of the victim to exert additional pressure on the victim to pay (categories: yes = 1 / no = 0).

2h. *Communication victim-attacker:* Whether victim and attacker communicated through e-mail, a self-made *TORchat* application, or differently (categories: TOR/e-mail/different). A self-made *TORchat* application is considered more effort than e-mail.

3. Victim characteristics. To measure opportunity, information was collected on several other variables:

3a. *Yearly revenue:* Yearly revenue victim in euro.

3b. *Staff:* Staff working at victim's company.

3c. *Sector:* Economic sector of the victim's company, as categorized by the Dutch Chamber of Commerce.

3d. *Insurance:* Whether the victim has insurance which covers ransomware attacks (categories: yes = 1 / no = 0).

3e. *Backup:* Whether there were backups and the state of the backups (categories: no = 0, yes but not possible to recover of data = 1, yes but could partially recover data = 2, yes and could fully recover data = 3).

3f. *IR company:* If an Incident Response company helped the victim recover from the attack or/and negotiate with the attackers to get the decrypter (categories: yes = 1 / no = 0).

3g. *Days negotiating :* Amount of days negotiating in logarithm.

3f. *Repeat victimization:* Whether the victim has experienced a ransomware attack before, or another type of cybercrime (categories: yes+ransomware = 2, yes+other cybercrime = 1, no = 0).

3i. *Cloud:* Whether the victim has their IT infrastructure in the cloud (categories: no = 0, yes = 1, partially = 2, mitigating = 3).

4. Context variables. To measure the context of the attack, information was collected on the following variables:

4a. *Year:* Year of encryption[1]

4b. *Season:* Categories: Summer, autumn, winter, spring.

4c. *Time encryption:* Full date and time of encryption.

4d. *Time data exfiltration:* Full date and time of the stealing and exfiltrating of data of the victim.

4e. *Time access:* Full date and time when the first malicious activity on the target network was recorded.

To impute the missing observations we use Multiple Imputation Chained Equations (MICE) method [15]–[17], which is a more reliable than list wise deletion or simple imputation methods [18]. For a good explanation of how MICE works, we refer to [18]. The MICE method still gives reliable estimates with 60% missing variables [16]. We omit variables with more than 70% missing observations from our analysis: repeat victimization (3h), the time between access and encryption (4d), and time between data exfiltration and encryption (4e).

Analysis were conducted using R version 4.0.2, packages *MICE*, *ggplot* and *dplyr*. To test the hypothesis (see Figure 2) a subset of the variables were used, as depicted in the final three columns of Table I:

- H1: The factors influencing ransom requested. The variables in the 'Y1=RR' column were used as independent variables to perform linear regression analysis on the variable *1a. Ransom requested start negotiations*.
- H2: The factors influencing payment. The variables in the 'Y2=Pay' column were used as independent variables in probit regression analysis on the variable *1c. Payment*.

[1]Note that this can be before 1st of January 2019, since encryption occurs before reporting it to the police. We filter based on the date it was reported to the police.

TABLE I

VARIABLES USED IN THIS STUDY AND IN DIFFERENT REGRESSION ANALYSIS. IN THE FIRST COLUMN THE VARIABLES ARE DEPICTED: 1) DEPENDENT VARIABLES, 2) IS CRIMINAL EFFORT, 3) ARE VICTIM CHARACTERISTICS, AND 4) IS CONTEXT. IN THE SECOND COLUMN THE UNITS OR CATEGORIES OF A VARIABLE. IN THE THIRD COLUMN THE AMOUNT OF MISSING OBSERVATIONS PER VARIABLE. FINALLY, THE LAST THREE COLUMNS DEPICT WHICH VARIABLES ARE USED FOR THE REGRESSION ANALYSIS ON RANSOM REQUESTED (Y1=RR), PAYMENT (Y2=PAY) AND FINANCIAL LOSS (Y3=FL).

Variable	**Unit / categories**	**Missing values (%)**	**Y1= RR**	**Y2=Pay**	**Y3=FL**
1a. Ransom requested start negotiations	Euro, Log 10 transformed	196 (55.5%)	X		
1b. Ransom requested end negotiations	Euro, Log 10 transformed	194 (55.0%)		X	X
1c. Payment	Yes = 1 / no = 0	22 (6.2%)		X	
1d. Ransom paid	Euro, Log 10 transformed	26 (7.4%)			X
1e. Financial loss	Euro, Log 10 transformed	184 (52.1%)			X
2a. Data exfiltration	Yes = 1 / no = 0	229 (64.9%)	X	X	X
2b. Targeted ransomnote	Yes = 1 / no = 0	1 (0.3%)	X	X	X
2c. RaaS	Yes = 1 / no = 0	181 (51.3%)	X	X	X
2d. Strain	Lockbit, Dharma, Conti, Phobos, Sodinokibi, ech0raix, Others	87 (24.6%)	X	X	X
2e. NAS	Yes = 1 / no = 0	1 (0.3%)	X	X	X
2f. Access	Phishing, exploits, different	1 (0.3%)	X	X	X
2g. Blackmail	Attacker contacts employees, customers, other type of pressure	2 (0.6%)	X	X	X
2h. Communication victim-attacker	E-mail, TOR-chat, different	64 (18.1%)	X	X	X
3a. Yearly revenue victim	Euro, Log 10 transformed	25 (7.1%)	X	X	X
3b. Staff at victim's company	Log10 transformed	11 (3.1%)			
3c. Sector	Sectors described by Dutch Chamber of Commerce	1 (0.3%)	X	X	X
3d. Insurance	Yes = 1 / no = 0	28 (7.9%)	X	X	X
3e. Backup	No = 0, yes+no recovery = 1, yes+partial recovery = 2, yes+full recovery = 3	11 (3.1%)	X	X	X
3f. IR company	Yes = 1 / no = 0	244 (69.1%)		X	X
3g. Days negotiating	Days, Log10 transformed	45 (12.7%)		X	X
3h. Repeat victimization	Yes+ransomware = 2 , yes+ other cybercrime = 1, no = 0	314 (89.0%)			
3i. Cloud	No = 0, yes = 1, partially = 2, mitigating = 3	22 (6.2%)	X	X	X
4a. Year	2018/2019/2020/2021/2022	4 (1.1%)	X	X	X
4b. Season	Summer/Autumn/Winter/Spring	4 (1.1%)	X	X	X
4c. Time encryption	Date, time (DDMMYYYYhhmm)	4 (1.1%)			
4d. Time data exfiltration	Date, time (DDMMYYYYhhmm)	325 (92.1%)			
4e. Time access	Date, time (DDMMYYYYhhmm)	264 (74.8%)			

- H3: The factors influencing ransom requested. The variables in the 'Y3=FL' column were used as independent variables to perform linear regression analysis on the variable *1e. Financial loss*.

For all three models (Y1=RR, Y2=Pay and Y3=FL), backward stepwise selection was performed to find the best fitting model, using the *step* function in R. Stepwise selection is a method to find the best performing model by iteratively adding and removing predictors [50].

We model the ransom requested a the start of negotiations (Y1) and financial loss (Y3) with a Generalized Linear Model (GLM) with family Gaussian. Payment (Y2) is modeled as a Generalized Linear Model with family Probit. The specific choice for using Gaussian GLM is due to the dependent variable constituting a specific amount of money for (Y1) en (Y3). The Probit GLM is used for Y2 because it has a binary outcome variable. Furthermore, as described in [56] our observations might possibly also have interdependence of events and non-equal mean and variance of the dependent variable. A general model for GLM is defined as follows [58]:

$$Y_i = \beta_i x_i + ... + \delta_i \quad (1)$$

where i refers to the different observations, β_i are the estimated coefficients for x_i, x_i are the independent variables collected for the observations as described in Table I and δ_i is the residual. After the GLM, we group the dummy's of the different nominal variables and perform a Likelihood-ratio test [57] to determine the effect of the different variables. A p-value of 0.05 or lower supports the hypothesis that the variable is a significant predictor for the dependent variable with significance level $\alpha = 0.05$.

V. Data Analysis and Results

We analyse the data and interpret the results using the following 4 steps: First, we give a general overview of the data with the help of descriptive statistics. Second, we present our analysis for the three hypotheses. Third, we identify and discuss factors that contribute to the ransom payment. Last, we examine the financial loss reported by companies after a ransomware attack.

A. Descriptive Statistics

Descriptive statistics for a subset of independent variables are shown in Tables II and III below. Table II gives an overview of the different victim characteristics, grouped by sector. Companies within the industry sector *Trade* experienced most ransomware attacks (113 attacks). MAS (Milieu en Agrarische Sector, agriculture) was the fewest with 10 attacks. In the construction sector companies with the largest revenue faced ransomware attacks: 562 million euro. Leisure the least: 6,61 million euros was the average yearly revenue for companies who faced a ransomware attack. However, if we consider the median, education, and government had the largest yearly revenue. Companies in the Leisure sector were most often insured with 15 %, and not reported in MAS, Media, and education with 0%. Finally, the average ransom was largest in the ICT sector, with 1.3 million euros, and lowest for the media, with 11,000 euros on average.

In Table III we present an overview of different attacking strains. Most attacks were performed by group(s) behind Phobos and Sodinokibi (32 times). However, the attacks associated with Conti targets the companies with the largest mean and median revenue: respectively 437 million and 31 million euro. The strain 'Others', contains all other groups. Compared to the groups mentioned here, they target relatively large companies with 327 million euro on average, or 3.7 million euro median.

The final three columns of Tables II and III illustrate the dependent variables: financial loss, payment and ransom requested. It is noteworthy that in Table II the trade and ICT sector have the largest ransom requested, both averages are larger than 1 million euros. Payment is largest in ICT, and transport. Financial loss was largest in trade. This might be due to downtime costs: for companies who sell products or offer services the downtime costs might be highest. Compared to MAS (agriculture) or construction, where work probably could continue without the immediate use of computers.

Considering the different strains in Table III, we find the largest ransom requested by Conti, also the highest financial loss. Dharma has the largest amount payed, perhaps because of the low amount of ransom requested compared to other groups. Ech0raix is the group that targets the smallest companies with 1.27 million euro yearly revenue on average, demanding 750 euro and reported financial loss of 2,620 euros. These results seem to be in line with the relationship which was described in the Introduction: more effort, as defined by the crime script, should lead to larger ransom requested by attackers and financial loss by victims.

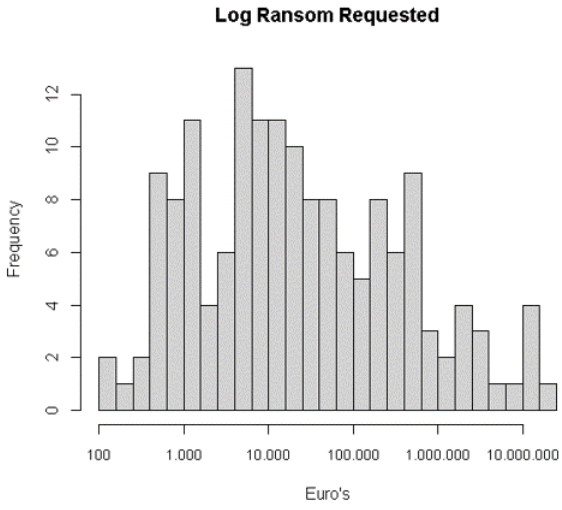

Fig. 5. Distribution of log ransom requested before negotiations.

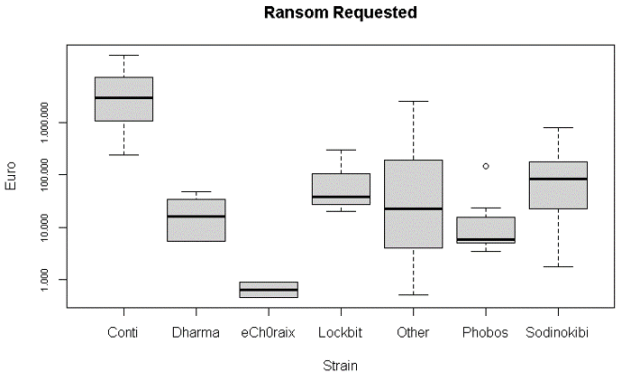

Fig. 6. Boxplot of log ransom requested for each ransomware strain.

B. Hypothesis testing

Next, we use regression analysis to test the three hypotheses introduces in Section III. We use a linear regression model to test **H1** and **H3** and use a probit regression to test **H2**. We discuss the details of our hypothesis testing methodology in Section IV. The likelihood ratios for each of the variables tested for the three hypotheses are show in Table IV.

Based on the GLM likelihood ratios for **H1** we find that variables that capture attacker's effort such as 'Data exfiltration', use of 'RaaS', 'Blackmail' and active 'Communication

TABLE II
DESCRIPTIVE STATISTICS OF VICTIM COMPANIES OF DIFFERENT SECTORS. MEAN AND MEDIAN REVENUE ARE IN MILLION EUROS, INSURED, NO BACKUP, AND PAID ARE PERCENTAGES. FINANCIAL LOSS AND RANSOM IS IN THOUSAND EUROS.

	Sector	Number of attacks	Mean Revenue (Meuro)	Median Revenue (Meuro)	(%) Insured	(%) No Backup	Financial Loss (euro)	(%) Ransom Paid	Ransom Requested (euro)
1	Construction	53	562.84	2.43	10.2	35.3	256,410	27.5	182,840
2	Healthcare	21	37.62	2.33	10.5	42.9	77,690	26.3	23,770
3	Trade	113	133.96	2.84	4.9	38.9	737,610	25.5	1,106,800
4	ICT	60	120.59	3.81	13	30.8	232,580	30.9	1,343,190
5	MAS	12	376.36	0.63	0	18.2	12,500	9.1	13,700
6	Media	20	142.54	3.30	0	52.9	344,800	15.8	11,640
7	Education	14	101.43	19.44	0	14.3	49,800	21.4	555,660
8	Government	10	60.17	18.45	10	20	393,330	0	820,350
9	Leisure	20	6.61	1.08	15	55	27,000	15	81,020
10	Transport	29	389.05	6.00	7.4	34.6	838,85	30.8	529,540

TABLE III
DESCRIPTIVE STATISTICS OF THE DIFFERENT RANSOMWARE STRAINS. MEAN AND MEDIAN REVENUE IN MILLION EUROS, INSURED, NO BACKUP, AND PAID ARE PERCENTAGES. DAMAGE AND RANSOM ARE IN EUROS.

	Strain	Number of attacks	Mean Revenue (Meuro)	Median Revenue (Meuro)	(%) Insured	(%) No Backup	Financial Loss (euro)	(%) Ransom Paid	Ransom Requested (euro)
1	Conti	19	437.43	30.71	28.6	15.8	4,726,280	16.7	6,598,380
2	Dharma	14	7.68	3.98	0	35.7	29,010	50	18,760
3	Others	143	327.26	3.74	4.7	28.6	298,610	27.7	542,330
4	eCh0raix	10	1.27	0.6	11.1	50	2,620	20	750
5	Lockbit	16	23.7	4.57	18.8	43.8	184,380	20	98,980
6	Phobos	32	261.85	1.07	3.2	51.7	167,560	30	21,190
7	Sodinokibi	32	56.43	3.56	16.7	21.9	170,070	18.5	658,010

between attacker and victim' are all significant factors in predicting the requested ransom. The median ransom requested when the communication was made using *TORchat* was 21K euros, whereas the median ransom requested when other communication channels were used was nearly 3.5K euros. Also, factors that increase the perceived benefits for attacks such as 'yearly revenue' of the victim firm, 'industry sector' and 'Insurance' were also statistically significant in predicting the ransom requested. Figures 5 and 6 illustrate respectively the distribution of log ransom requested and the ransom requested when malware strain from a particular group was used.

While analysing the factors that influence the likelihood of victims to pay ransom (**H2**), we find that cost and attacker's effort related variables such as 'ransom requested end negotiations', 'Data exfiltration', 'Targeted ransomnote', 'Blackmail' and 'Days negotiating' were statistically significant predictors. Interestingly, victim characteristic related variable capturing its 'Backup' status was also significant, 28% of the victims with no back ups pay ransom, where as 48% of those whose backups became unrecoverable after ransomware infection pay the ransom. 16% of the victims who has partial recovery of the backups paid ransom, while only 6% of the victims with fully recoverable backups paid the ransom.

In analysis of **H3** we evaluate the factors that explained the loss reported by victim firms. Paid ransom formed a significant part of the reported losses. We again find that attacker effort related factors such as 'Data exfiltration', 'Targeted ransomnote' and use of 'RaaS' again significantly affected the amount of reported losses. The median financial losses reported were the lowest when no or full backup was available. This shows

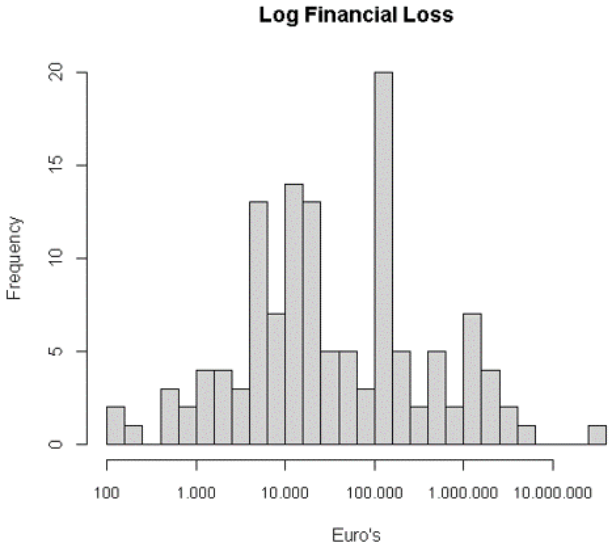

Fig. 7. Distrbution of log financial loss for victims after a ransomware attack.

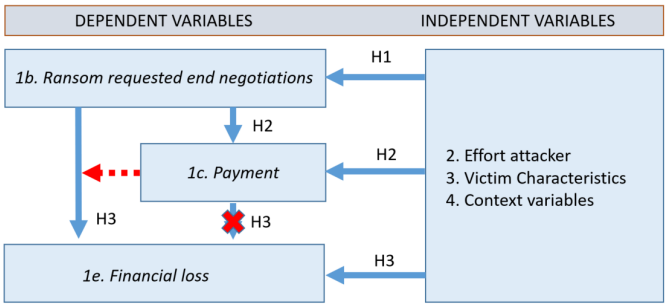

Fig. 8. Compared to Figure 2, our results support the hypothesis that attackers' effort, victim characteristics and context variables influence ransom requested, payment and financial loss. Furthermore, variable '1c. Payment', that has an interaction effect, along with '1a. Ransom requested' is also an important factor for determining the financial loss for victims after a ransomware attack.

that victims that did not already have a backup as part of their resilience strategy, perceived the financial impact of such attacks to be low. While, one with full and recoverable backups were able to hit the ground running without suffering huge losses.

Financial loss of victims (see Figure 7) is influenced by the yearly revenue of the victim, the amount of ransom paid, whether the attacking group is known to be RaaS and whether an Incident Response company helped the victim to recover from the ransomware attack.

Finally, we analysed the direct effect of ransom requested on payment (1), ransom requested, and payment on financial loss (2). Performing a probit regression of ransom requested on payment led to insignificant results ($\beta = -0.1284$, p = 0.33). Regression ransom requested and payment on financial loss, we found ransom requested predicts financial loss ($\beta = 0.82$, p<0.001), but payment variable itself had no effect on financial loss ($\beta = 0.13$, p = 0.38). This result might also be due to how financial loss is operationalized. Victims report to the police financial loss often a couple of weeks after the attack started. At that point, downtime (or other) costs might be not that much different between victims who paid or did not pay.

VI. Discussion and Conclusion

This study set out to examine the relationship between ransom requested, payment, and financial loss. We examined 353 ransomware attacks reported to the Dutch Police. Based on the RCM and ransomware crime script, we argued that attackers' effort, victim characteristics and context variables are important factors to understand the ransom requested by attackers, whether victims paid the ransom and the financial loss reported by the victims after the attack.

For the ransom requested, we found that data exfiltration, RaaS groups, insurance and mode of communication are important predictors for ransom requested. These results support the hypothesis that attackers' effort and victim characteristics are important factors for the ransom requested by attackers. Furthermore, yearly revenue, blackmail and sector had a possible effect on ransom requested. Finally, these results show that effort could be quantified considering a crime script of ransomware.

For payment, the ransom requested after negotiations, data exfiltration, targeted ransomnote, blackmail, backup, days of negotiating best predict whether victims pay the ransom. We find these results in agreement with the rational choice model of crime. Our results do not indicate a difference between victim characteristics like yearly revenue and sector for the probability of paying. Victim behaviour could be more important: did they have a backup and how long did they negotiate? The effort of the attacker also influenced the decision of the victim to pay: Data exfiltration, targeted ransom note and blackmail are positively related to the probability of payment.

For financial loss we found ransom paid, data exfiltration, RaaS, yearly revenue of the victim, backup and season to be important factors contributing to financial loss as reported by the victim to the police. It is important to note that none of the victims was able to indicate the costs of reputation damage and liability or costs in the long term to the police, because they needed to disclose a (realistic) financial loss when reporting the attack to the Police. Nevertheless, these results seem to support the hypothesise that financial loss is determined by attackers' effort, victim characteristics, context and the amount of ransom paid. Interestingly, whether the victims did or did not pay the ransom, did not contribute to the reported financial loss.

Finally, considering the direct relationships between dependent variables (see Figure 8), a direct relationship between ransom requested and financial loss was found. Furthermore, a direct effect of ransom requested after negotiations on payment and payment to financial loss was also found statistically significant.

TABLE IV
RESULTS OF REGRESSION ANALYSIS

Regression			**Regression**		
Variables	Likelihood Ratio	Df	Variables	Likelihood Ratio	Df
Y1=RR			3d. Insurance	11.00	1
2a. Data exfiltration	96.47**	1	3e. Backup	20.72**	3
2b. Targeted ransomnote	0.57	1	3f. IR company	12.00	1
2c. RaaS	32.99**	1	3g. Days negotiating	55.34**	1
2d. Strain	49.81	6	3i. Cloud	567.00	1
2e. NAS	0.9	1	4a. Year	5.46	4
2f. Access	19.61	3	4b. Season	167.00	3
2g. Blackmail	35.31*	1	**Y3=FL**		
2h. Communication victim-attack	114.76**	3	1b. Ransom requested end negotiations	552.00	1
3a. Yearly revenue victim	37.82*	1	1d. Ransom paid	34.01**	1
3c. Sector	162.42*	9	2a. Data exfiltration	6.10*	1
3d. Insurance	76.29**	1	2b. Targeted ransomnote	3.15*	1
3e. Backup	40.93	3	2c. RaaS	5.46*	1
3i. Cloud	0.02	1	2d. Strain	2.14	6
4a. Year	37.05	4	2e. NAS	289.00	1
4b. Season	21.16	3	2f. Access	2.15	3
Y2=Pay			2g. Blackmail	3.63*	1
1b. Ransom requested end negotiations	9.74**	1	2h. Communication victim-attack	2.29	3
2a. Data exfiltration	8.83**	1	3a. Yearly revenue victim	40.14**	1
2b. Targeted ransomnote	4.27*	1	3c. Sector	6.44	9
2c. RaaS	40.00	1	3d. Insurance	88.00	1
2d. Strain	5.21	6	3e. Backup	7.93*	3
2e. NAS	507.00	1	3f. IR company	271.00	1
2f. Access	3.41	3	3g. Days negotiating	44.00	1
2g. Blackmail	19.03**	1	3i. Cloud	1.08	1
2h. Communication victim-attack	3.84	3	4a. Year	4.83	4
3a. Yearly revenue victim	0.00	1	4b. Season	13.22**	3
3c. Sector	13.47	9			

Note. All data is rounded to 2nd significant figure.
*$p \leq .05$. **$p \leq .01$.

VII. LIMITATIONS AND FURTHER WORK

There are different limitations of this study:

1) We collected data based on ransomware reports filed by companies and individuals to the Dutch Police. The nature of this data makes it a challenge to generalize the results to other countries and victims who do not report the ransomware attack to the Dutch Police. These challenges could be tackled by collecting data from multiple Law Enforcement agencies around the world and incident response companies, for example. These companies also help victims who do not report the attack to the police, making it possible to estimate selection bias due to the willingness to report.
2) The crime script of ransomware as described in this paper sets out to understand and structure the collected data. However, it is possible to improve this crime script by including more ransomware attacks and from different countries.
3) The regression models could be biased due to the large number of missing observations. Although the MICE method is, to our understanding, a good way to impute the missing data, the models could be improved by decreasing the amount of missing observations. One way to achieve this is by training police officers to ask for more and specific information when the victim reports a ransomware attack.
4) Due to the sensitivity of the data, only one annotator could code the data. This might result in several types of biases [54]. This problem was important when coding the categorical variables. For example, a company that sells buses, should it belong to trade or to transport? We tried to limit the severity of this limitation by anonymously discussing these issues with experts outside the project and writing down the choices to improve consistency. In this specific case we decided that selling buses belongs to trade, since that is the main objective of the company. Further research could address this issue by asking

permission for multiple researchers to get access to the sensitive data from the start of the research project.

To understand how sample bias might have affected this study, we compare sample size of ransomware attacks and percentage payments with previous literature. Considering the sample size of ransomware attacks in other studies, [66] examined 623 ransomware incidents in the EU, United Kingdom and United States between May 2021 and June 2022. [67] examined 101 ransomware attacks in 2020 in 81 countries. Comparing these two studies with the present study, we examined a relatively large sample size: around 100 cases within one year in the Netherlands. Furthermore, comparing the sample of this study with research from the industry [61], it seems the sample might contain more cases from individuals and small and medium enterprises, since they perhaps cannot afford incident response services after a ransomware attack.

Considering the percentage payments, [59] indicates 85% of the victims pays, but this is based on 13 observations and was in 2016. [60] surveyed 41 companies in the UK between 2014 and 2018, of which 8 companies (19.5%) paid the ransom. Victims were sampled from UK Police data. The percentages in these two studies are based on small samples. Finding of [60] aligns with our study. As described earlier, it might be that victims who pay are less inclined to go to the Police to report their attack. Perhaps because the Police expresses the strong view to never negotiate or pay (ransomware) criminals. Payment rates from the industry seem to confirm this. According to [61], Kaspersky found that in 2020 52% of ransomware victims paid. It would be interesting to reproduce this study with data from incident response companies and to survey companies when they would go to the Police. As is, this was mostly studied considering other cybercrimes [62], [63] but not ransomware. Taken together, these results indicate that not all companies report to the police and that victims that pay are less willing to report to the police.

One other interesting finding in the present study is that 6 ransomware strains account for almost 50% of the cases. This seems to align with previous offline crime research: there has been a concentration of offending and offenders in time and space [68]. However, other ransomware research did not found such a strong concentration [66], [67].

In conclusion, this study is the first attempt to do a large-scale empirical study. Despite its limitations, the relatively large sample size [59]–[61] made it possible to study the effort of the attacker, victim characteristics and context variables in depth and their influence on the ransom requested, the payment of the ransom and the financial loss reported by the victim.

Furthermore, this study might support interventions by Law Enforcement and policy makers. Law Enforcement could intervene on the factors which influence the ransom requested, to reduce the amount of money attackers make with ransomware attacks. Policy makers could conduct targeted prevention campaigns to companies in specific sectors and large companies, as these characteristics seem to indicate larger ransom requested and therefore more profitable for attackers. These campaigns could be increased during specific seasons, as this was an indicator for the financial loss of victims. Victims who are under attack could be warned and be prepared for potential blackmail strategies and publication of confidential data on leakpages. Finally, prevention campaigns could focus on prevention: make sure that potential victims have reliable backups, which are not accessible through the network by attackers. Backups decrease the probability of paying and therefore decreases the financial gains of ransomware attacks.

VIII. Ethics

We follow the principles from Menlo Report [55] to justify the ethical considerations made in this study:

Respect for persons: Privacy of victims was taken into considerations when writing this paper. By not considering individual cases and only aggregating to strains and sector of victims, we feel confident the privacy of victims is respected.

Beneficence: Information of the police investigations was only available to researcher who had a proper police screening. For the other researchers involved in this project only aggregated results were available. Although this conflicts with the scientific principles of transparency and reproduce-ability, this seemed the only way to conduct a large-scale empirical ransomware study. Furthermore, results presented in this paper should exclude personal identifiable information.

Justice: Selection of ransomware attacks was only on the keyword 'ransomware' in the police systems. In this way, all ransomware attacks got an equal chance to be part of the study. No extra effort was put into attacks which got a lot of media attention.

Respect for Law and Public Interest: An important factor we took into consideration was the information position regarding specific groups and/or strains or the way the Dutch Police operates. These were excluded from the paper.

IX. Acknowledgements

We would like to extend our sincere gratitude to the Dutch Police. In particular, we would like to thank Emma Ratia, Theo van der Plas and Cees van Tent for making the project possible. Furthermore, we thank the Cybercrime Unit East Netherlands and the Ransomware Taskforce for their expertise. Please note that the views expressed in this work are ours alone and do not necessarily reflect those of the Dutch Police. Finally, we would like to thank our shepherd, Laurin Weissinger, and the anonymous reviewers for their suggestions to improve this paper.

References

[1] L. Yuryna Connolly, D.S. Wall, M. Lang, and B. Oddson (2020). An empirical study of ransomware attacks on organizations: an assessment of severity and salient factors affecting vulnerability. Journal of Cybersecurity, 6(1), tyaa023.

[2] C. Simoiu, J. Bonneau, C. Gates, and S. Goel (2019). " I was told to buy a software or lose my computer. I ignored it": A study of ransomware. In Fifteenth symposium on usable privacy and security (SOUPS 2019) (pp. 155-174).

[3] R. Anderson, C. Barton, R. Böhme, R. Clayton, M. J. Van Eeten, M. Levi, and S. Savage (2013). Measuring the cost of cybercrime. In The economics of information security and privacy (pp. 265-300). Springer, Berlin, Heidelberg.

[4] C. Beaman, A. Barkworth, T. D. Akande, S. Hakak, and M. K. Khan (2021). Ransomware: Recent advances, analysis, challenges and future research directions. Computers & Security, 111, 102490.

[5] H. Oz, A. Aris, A. Levi, A. S. and Uluagac (2021). A survey on ransomware: Evolution, taxonomy, and defense solutions. ACM Computing Surveys (CSUR).

[6] M. Humayun, N. Z. Jhanjhi, A. Alsayat, and V. Ponnusamy (2021). Internet of things and ransomware: Evolution, mitigation and prevention. Egyptian Informatics Journal, 22(1), 105-117.

[7] K. Oosthoek, J. Cable, and G. Smaragdakis (2022). A Tale of Two Markets: Investigating the Ransomware Payments Economy. arXiv preprint arXiv:2205.05028.

[8] O. Owolafe, and A.F. Thompson (2022). Analysis of Crypto-Ransomware Using Network Traffic. Journal of Information Security and Cybercrimes Research, 5(1), 72-79.

[9] K. Wang, J. Pang, D. Chen, Y. Zhao, D. Huang, C. Chen, and W. Han (2021). A large-scale empirical analysis of ransomware activities in bitcoin. ACM Transactions on the Web (TWEB), 16(2), 1-29.

[10] J. Hernandez-Castro, A. Cartwright, and E. Cartwright (2020). An economic analysis of ransomware and its welfare consequences. Royal Society open science, 7(3), 190023.

[11] A. Cartwright, and E. Cartwright (2019). Ransomware and reputation. Games, 10(2), 26.

[12] Z. Li, and Q. Liao (2021). Game theory of data-selling ransomware. Journal of Cyber Security and Mobility, 65-96.

[13] Behind the curtains of the ransomware economy - the victims and the Cybercriminals. Check Point Research. (2022). Retrieved July 12, 2022, from https://research.checkpoint.com/2022/behind-the-curtains-of-the-ransomware-economy-the-victims-and-the-cybercriminals/.

[14] A. Cartwright, E. Cartwright, L. Xue, and J. Hernandez-Castro (2022). An investigation of individual willingness to pay ransomware. Journal of Financial Crime, (ahead-of-print).

[15] S. Van Buuren, and K. Groothuis-Oudshoorn (2011). mice: Multivariate imputation by chained equations in R. Journal of statistical software, 45, 1-67.

[16] E. Kontopantelis, I.R. White, M. Sperrin, and I. Buchan, (2017). Outcome-sensitive multiple imputation: a simulation study. BMC medical research methodology, 17(1), 1-13.

[17] K. Heidt (2019). Comparison of imputation methods for mixed data missing at random.

[18] J.R. Van Ginkel, M. Linting, R.C. Rippe, and A. van der Voort (2020). Rebutting existing misconceptions about multiple imputation as a method for handling missing data. Journal of personality assessment, 102(3), 297-308.

[19] F. Changyong, W. Hongyue, L. Naiji, C. Tian, H. Hua, and L. Ying (2014). Log-transformation and its implications for data analysis. Shanghai archives of psychiatry, 26(2), 105.

[20] Europol (2021). Internet Organised Crime Threat Assessment (IOCTA) 2021, Publications Office of the European Union, Luxembourg, Retrieved August 31, 2022, from https://www.europol.europa.eu/publications-events/main-reports/internet-organised-crime-threat-assessment-iocta-2021.

[21] T. Hofmann (2020). How organisations can ethically negotiate ransomware payments. Network Security. 2020 Oct;2020(10):13–7.

[22] R. P. Baksi (2022). Pay or Not Pay? A Game-Theoretical Analysis of Ransomware Interactions Considering a Defender's Deception Architecture. 52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks-Supplemental Volume (DSN-S) (pp. 53-54). IEEE.

[23] N. A. Hassan(2019). Ransomware revealed: a beginner's guide to protecting and recovering from ransomware attacks. Apress.

[24] M. Keshavarzi, and H.R. Ghaffary (2020). I2CE3: A dedicated and separated attack chain for ransomware offenses as the most infamous cyber extortion. Computer Science Review, 36, 100233.

[25] M.N. Olaimat, M. A. Maarof, and B. Al-rimy (2021). Ransomware anti-analysis and evasion techniques: A survey and research directions. In 2021 3rd international cyber resilience conference (CRC) (pp. 1-6). IEEE.

[26] D.B. Cornish (1994). The procedural analysis of offending and its relevance for situational prevention. Crime prevention studies, 3(1), 151-196.

[27] A. Hutchings, and T. J. Holt (2015). A crime script analysis of the online stolen data market. British Journal of Criminology, 55(3), 596-614.

[28] A. Rege, Z. Obradovic, N. Asadi, B. Singer, and N. Masceri (2017). A temporal assessment of cyber intrusion chains using multidisciplinary frameworks and methodologies. In 2017 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (Cyber SA) (pp. 1-7). IEEE.

[29] M. Junger, V. Wang, and M. Schlömer (2020). Fraud against businesses both online and offline: Crime scripts, business characteristics, efforts, and benefits. Crime science, 9(1), 1-15.

[30] Team Cymru (2022). Analyzing ransomware negotiations with CONTI: An in-depth analysis.

[31] S. Hilt, and F. Merces (2022). Backing Your Backup. Trend Micro.

[32] G.J. Falk (1952). The influence of the seasons on the crime rate. J. Crim. L. Criminology & Police Science, 43, 199.

[33] D.B. Cornish, and R.V. Clarke (1989). Crime specialisation, crime displacement and rational choice theory. In Criminal behavior and the justice system (pp. 103-117). Springer, Berlin, Heidelberg.

[34] M. Felson, and R. V. Clarke (1998). Opportunity makes the thief. Police research series, paper, 98(1-36), 10.

[35] P. H. Meland, Y. F. F. Bayoumy, and G. Sindre (2020). The Ransomware-as-a-Service economy within the darknet. Computers & Security, 92, 101762.

[36] H. Borrion, and L. Yuryna Connolly (2020). Your money or your business: Decision-making processes in ransomware attacks. Association for Information Systems.

[37] Leaks of Conti Ransomware Group Paint Picture of a Surprisingly Normal Tech Start-Up (2022). Check Point Research. Retrieved August 31 2022, from https://research.checkpoint.com/2022/leaks-of-conti-ransomware-group-paint-picture-of-a-surprisingly-normal-tech-start-up-sort-of/

[38] Bahrami, P. Nikkhah, A. Dehghantanha, T. Dargahi, R. M. Parizi, K. C. Choo and H. H. S. Javadi (2019). "Cyber Kill Chain-Based Taxonomy of Advanced Persistent Threat Actors: Analogy of Tactics, Techniques, and Procedures." Journal of Information Processing Systems 15(4):865-89.

[39] L. Y. Connolly, M. Lang, P. Taylor and P.J. Corner (2021). The Evolving Threat of Ransomware: From Extortion to Blackmail.

[40] J. DiMaggio (2021). Ransom Mafia. Analysis of the World's First Ransomware Cartel. Analyst1. 7 April.

[41] A. Haymore (2021). We Wait, Because We Know You. Inside the Ransomware Negotiation Economics. Threat Intelligence, Fox-IT and European Research, Threat Intelligence NCCgroup. (from https://research.nccgroup.com/2021/11/12/we-wait-because-we-know-you-inside-the-ransomware-negotiation-economics/)

[42] E. M. Hutchins, M. J. Cloppert and R.M. Amin (2011). Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. Leading Issues in Information Warfare & Security Research 1:80.

[43] P. Leo, Ö. Işik and F. Muhly (2022). The Ransomware Dilemma. Mit Sloan Management Review MIT Sloan Management Review.

[44] G. Stringhini (2019). Adversarial Behaviours Knowledge Area. Vol. Cyber Security Body of Knowledge. The National Cyber Security Centre.

[45] D.Y. Huang, M.M. Aliapoulios, V.G. Li, L. Invernizzi, K. McRoberts, E. Bursztein, J. Levin, K. Levchenko, A.C. Snoeren, and D. McCoy (2018). Tracking Ransomware End-to-end. 39th IEEE Symposium on Security and Privacy, S & P, pp. 618–631.

[46] K. Huang, M. Siegel, and S. Madnic. (2018). Systematically Understanding the Cyber Attack Business: A Survey. ACM Comput. Surv. 51, 4, Article 70 (July 2019), 36 pages. https://doi.org/10.1145/3199674.

[47] I. Kara, and M. Aydos (2022). The rise of ransomware: Forensic analysis for windows based ransomware attacks. Expert Systems with Applications, 190, 116198.

[48] B. Urooj, M. A. Shah, C. Maple, M. K. Abbasi, and S. Riasat (2022). Malware detection: a framework for reverse engineered android applications through machine learning algorithms. IEEE Access.

[49] B. Payne, and E. Mienie (2021). Multiple-Extortion Ransomware: The Case for Active Cyber Threat Intelligence. In ECCWS 2021 20th European Conference on Cyber Warfare and Security (p. 331). Academic Conferences Inter Ltd.

[50] H. Väliaho, and T. Pekkonen, T. (2022). A Procedure for Stepwise Regression Analysis. In A Procedure for Stepwise Regression Analysis. De Gruyter.

[51] M. Felson, and L. E. Cohen (1980). Human ecology and crime: A routine activity approach. Human Ecology, 8(4), 389-406.

[52] E. R. Leukfeldt, and M. Yar (2016). Applying routine activity theory to cybercrime: A theoretical and empirical analysis. Deviant Behavior, 37(3), 263-280.

[53] I. Arghire (2022). QNAP Appliances Targeted in New DeadBolt, eCh0raix Ransomware Campaigns — SecurityWeek.Com. Security-Week - A Wired Business Media Publication. Retrieved August 26, 2022, from https://www.securityweek.com/qnap-appliances-targeted-new-deadbolt-ech0raix-ransomware-campaigns

[54] R. Artstein, and M. Poesio (2009). Bias decreases in proportion to the number of annotators. In Proceedings of FG-MoL 2005: The 10th conference on Formal Grammar and The 9th Meeting on (Vol. 139).

[55] M. Bailey, D. Dittrich, E. Kenneally, and D. Maughan (2012). The menlo report. IEEE Security & Privacy, 10(2), 71-75.

[56] M. Korczynski, S. Tajalizadehkhoob, A. Noroozian, M. Wullink, C. Hesselman, and M. Van Eeten (2017). Reputation metrics design to improve intermediary incentives for security of TLDs. In 2017 IEEE European Symposium on Security and Privacy (EuroS&P) (pp. 579-594). IEEE.

[57] J. Fan, C. Zhang, and J. Zhang (2001). Generalized likelihood ratio statistics and Wilks phenomenon. The Annals of statistics, 29(1), 153-193.

[58] J. J. Faraway (2016). Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. Chapman and Hall/CRC.

[59] K.S. Choi, T.M. Scott, and D.P. LeClair (2016). Ransomware against police: diagnosis of risk factors via application of cyber-routing activities theory. International Journal Forensic Science Pathol, 4:253–8.

[60] A.Y. Connolly, and H. Borrion (2022). Reducing Ransomware Crime: Analysis of Victims' Payment Decisions. Computers Security, 102760.

[61] D. Ndichu (2021) Kaspersky: over half of ransomware victims paid off attackers in 2020, Kaspersky, 4 April, available at: https://gulfbusiness.com/kaspersky-over-half-of-ransomware-victims-paid-offattackers-in-2020/ [Accessed November 2022].

[62] A. Graham, T. C. Kulig, and F. T. Cullen (2019). Willingness to report crime to the police: Traditional crime, cybercrime, and procedural justice. Policing: An International Journal.

[63] S.G. Van de Weijer, R. Leukfeldt, and W. Bernasco (2019). Determinants of reporting cybercrime: A comparison between identity theft, consumer fraud, and hacking. European Journal of Criminology, 16(4), 486-508.

[64] R.V. Clarke, and P.M. Harris (1992). A rational choice perspective on the targets of automobile theft. Criminal Behaviour and Mental Health, 2(1), 25-42.

[65] D. Walsh (2017). Victim selection procedures among economic criminals: The rational choice perspective. The reasoning criminal (pp. 39-52). Routledge.

[66] ENISA (2022). Ransomware: Publicly Reported Incidents are only the tip of the iceberg. Available from: https://www.enisa.europa.eu/news/ransomware-publiclyreported-incidents-are-only-the-tip-of-the-iceberg. [accessed November 2022].

[67] P. Langlois (2020). 2020 Data Breach Investigations Report. Verizon.

[68] G. Farrell (2015). Crime concentration theory. Crime prevention and community Safety, 17(4), 233-248.