RAMSES

Internet Forensic platform for tracking the money flow of financially-motivated malware

H2020 - 700326

D4.1 Findings on economic modelling of malware as business model

Lead Author: Darren Hurley-Smith (UNIKENT)

With contributions from: Julio Hernandez-Castro, Edward Cartwright and Anna Stepanova (UNIKENT)

Reviewer: Andrea Continella (POLIMI)

Deliverable nature:	Report (R)
Dissemination level: (Confidentiality)	Public (PU)
Contractual delivery date:	31/05/2017
Actual delivery date:	30/05/2017
Version:	1.0
Total number of pages:	31
Keywords:	Malware, ransomware, economics, modelling, predictive analysis

Abstract

This report is intended for public release, as agreed by the members of the RAMSES consortium. As the initial deliverable of WP 4 – Modelling Ransomware from the Point of View of Economic Theory and Applications, this report presents research regarding the economic modelling of ransomware. The aim is to outline the current state-of-the-art in ransom strategy and to compare current criminal strategies with optimal profit models. This will provide a foundation for the development of a predictive economic model of ransomware in the subsequent work as a part of RAMSES.

Focusing on optimal ransoming strategies as a starting point, this report compares mathematically backed examples of existing strategies to those used in ransomware. Through this comparison, the current state of ransomware from an economic perspective can be derived. This allows us to comment on the likely developments in the ransom process, such as pricing strategy and negotiation, as ransomware continues to be optimised.

Executive summary

- This report is intended for public release.
- The research undertaken is intended to provide a foundation for further development of WP 4 – Findings on economic modelling of malware as business model. Existing examples of malware are compared with economic models and comparable strategies in hostage taking and related activities.
- It is a precursor for D4.2 – Economic Improvements over the Existing Models. It provides a baseline for innovative research that will further explore the ransomware-specific economic attributes that determine the profitability of a given strategy.
- This report also informs D4.3 – Estimated Evolutionary Path from Present to Future. By developing an initial overview of economic factors that determine which ransomware profits and the factors that might lead to failure, the evolutionary forces acting on the domain can begin to be defined. By drawing comparisons with hostage taking and other traditional means of exerting control over items of value for profit, we can report on the similarities and differences between the strategies employed and their efficacy in both traditional ransom and contemporary ransomware scenarios.
- This deliverable is based on research undertaken by Julio Hernandez-Castro, Edward Cartwright and Anna Stepanova of the University of Kent. A research paper approved by the consortium has been generated alongside this deliverable by the previously named academics, and approved for publication by the consortium. Ongoing research by these individuals informs all aspects of this report. A literature survey and historical overview of ransomware is provided.
- The report has been collated and edited by Darren Hurley-Smith of the University of Kent. Andrea Continella of the Politecnico of Milan and Tatiana Silva of Treelogic have provided invaluable feedback.
- The key elements of this report are:
    - A cost analysis of ransomware, which highlights the cost of doing business in a ransomware setting. Set up costs, maintenance and the cost of extracting cash value from such activities are discussed. Relevant economic forces are highlighted and mathematically modelled. This provides an overview of the technical and financial requirements of groups wishing to use ransomware for profit. Technical understanding doesn't need to be exceptionally high, due to the emergence of ransomware as a service. Cost is largely a matter of scale and duration of a ransomware activity: botnets and other infection vectors may be purchased, along with the expertise to use them effectively. The impacts of premature discovery of targeted vulnerabilities and patching on profit are discussed.
    - Economic analysis of ransomware, focusing on pricing strategies and profit maximisation techniques. Fixed pricing, targeted ransoms and negotiation are discussed. It becomes apparent that in many cases, the obvious path is not the optimal profit bearing strategy (for example, negotiation lowers the maximum profit per ransom that can be extracted in many scenarios). This will inform future work that defines the evolutionary forces acting on ransomware, both from competition within the domain and friction caused by mediums of exchange (cryptocurrencies) and victims (the likelihood of victims adopting defensive strategies or choosing not to pay because of a failure in the chosen ransom strategy).
    - A survey of willingness to pay in a range of scenarios. This part of the report focuses on a survey undertaken by University of Kent researchers. A questionnaire was circulated to two groups of participants, who were presented with one of two ransom scenarios. The answers to the questionnaire provide an initial insight into the effect of a given ransom strategy on the willingness of an individual to pay. This work tests earlier academic observations and further develops the statement made regarding the effects of threat, negotiation and context on the willingness of victims to pay ransoms.

- This report will be referenced throughout WP 4 to tie all further work to a common source. It will also be iteratively developed upon ion subsequent deliverables, with new findings being applied to those outlined in this work, where they challenge or further develop concepts.
- The foundation that this report forms will be used to inform the development of:
    - A predictive model of ransomware, modelling the effects of economic forces on individual examples of ransomware.
    - Further research outputs discussing the domain, and the likely evolutionary paths that we will see as ransomware moves beyond its nascent stages and begins to adopt a variety of strategies to overcome the issues associated with competition for limited resources (victims who will pay).
    - An optimal model system, which will collect all work done in previous items and apply it to the development of a software system. This will be used by LEAs to profile ransomware based on the economic and profit-driven factors outlined in this and future work.

Document Information

IST Project Number	700326	Acronym	RAMSES
Full Title	Internet Forensic platform for tracking the money flow of financially-motivated malware		
Project URL	http://www.ramses2020.eu		
EU Project Officer	Nada Milisavljevic		

Deliverable	**Number**	D4.1	**Title**	Findings on economic modelling of malware as business model
Work Package	**Number**	WP4	**Title**	Modelling Ransomware from the Point of View of Economic Theory and Applications

Date of Delivery	**Contractual**	M09	**Actual**	M09
Status	version 1.0		final ■	
Nature	prototype □ report ■ demonstrator □ other □			
Dissemination level	public ■ restricted □			

Authors (Partner)	University of Kent (UNIKENT)			
Responsible Author	**Name**	Darren Hurley-Smith	E-mail	d.hurley-smith@kent.ac.uk
	Partner	UNIKENT	Phone	+447870806745

Abstract (for dissemination)	This report is intended for public release, as agreed by the members of the RAMSES consortium. As the initial deliverable of WP4 – Modelling Ransomware from the Point of View of Economic Theory and Applications, this report presents research regarding the economic modelling of ransomware. The aim is to outline the current state-of-the-art in ransom strategy and to compare current criminal strategies with optimal profit models. This will provide a foundation for the development of a predictive economic model of ransomware in the subsequent work as a part of RAMSES. Focusing on optimal ransoming strategies as a starting point, this report compares mathematically backed examples of existing strategies to those used in ransomware. Through this comparison, the current state of ransomware from an economic perspective can be derived. This allows us to comment on the likely developments in the ransom process, such as pricing strategy and negotiation, as ransomware continues to be optimised.
Keywords	Malware, ransomware, economics, modelling, predictive analysis

Version Log			
Issue Date	**Rev. No.**	**Author**	**Change**
12/03/2017	0.1	Darren Hurley-Smith (UNIKENT)	Initial draft of contents page updated with content from the first publication by UNIKENT.
20/04/2017	0.2	Darren Hurley-Smith (UNIKENT)	Internal review with UNIKENT members Julio Hernandez-Castro and Edward Cartwright. Content from a further study and survey results are added.

09/05/2017	0.3	Darren Hurley-Smith (UNIKENT)	CONSOLIDATED DRAFT SUBMISSION: In preparation for the submission for consortium (POLIMI/TREE) review, UNIKENT performed an internal check of readability, spelling, grammar and the academic rigor of the provided content.
30/05/2017	1.0	Darren Hurley-Smith (UNIKENT)	FINAL VERSION: Acting on the insightful and helpful comments of Andrea Continella (POLIMI) and Tatiana Silva (TREE), the report has been improved for final submission.

Table of Contents

List of figures

List of tables

Abbreviations

IoT: Internet of Things

MR: Marginal Revenue

RaaS: Ransomware as a Service

WTA: Willingness to Accept

WTP: Willingness to Pay

Definitions

Bitcoin: The most iconic cryptocurrency currently in existence. Developed by Satoshi, Bitcoin represents the first successful cryptocurrency and has mass media exposure. This is a common currency for ransomware to request as payment, both due to its media exposure (making it likely that a victim will have heard of it) and the wealth of information and sources that make purchasing Bitcoin less arduous than other cryptocurrencies.

Blockchain: A fundamental technology that supports cryptocurrency. Effectively a distributed ledger, the blockchain provides a record of all transactions that have been agreed by a consensus among trusted nodes on a network. The most common example of Satoshi's Bitcoin blockchain, but other examples are common.

Crypto-Ransomware: A specific form of ransomware, which works by encrypting the contents of the target computer and using possession of the key as a bargaining chip. Differs from some strains of ransomware by focusing on threats of implicit data-loss instead of other forms of control. It may still rely on associated strategies, such as file deletion, but passively threatens to leave files encrypted in a manner that would make them irretrievable without the appropriate key. Generally a cash fee is asked (though some have more exotic requirements such as pyramid schemes).

Cryptocurrency: Digital currency, backed by one of a variety of means and using blockchain technology to provide a means of undisputed exchange of currency. Usually has a focus on anonymity, community and/or niche markets that provide backing (futures, computational power etc.).

Dark Net: A network overlaying the internet. It can only be accessed with specific software and operates using non-standard communication protocols. Usually intended to be private and anonymous, these networks are attractive to criminals and play host to black markets such as Alphabay.

Fiat Currency: A government backed currency, such as the Euro. Cryptocurrency must be changed back into a fiat currency if criminals wish to take their profits into most areas of the global economy.

Malware: Malicious software. Programs that cause damage and/or disruption to a target. May involve deletion of files, spying software, ransomware or one of many other forms of attack. Some extreme examples may focus on destruction of hardware (Stuxnet).

Malware as a Service: The sale of malware and expertise, instead of direct use. Involves technically adept groups and individuals exchanging their outputs for money.

Ransomware: Software that focuses on seizing control of technology, software or data. The control is used as leverage in a demand for the target to exchange a sum of money for the promised (but debatable) return of the seized items. Used colloquially to reference **Crypto-ransomware**, ransomware can mean any software that focuses on seizure of assets, not just cryptographic methods.

Ransomware as a Service: The act of developing, maintaining and providing technical expertise in the use of ransomware for profit. Instead of using their software themselves, individuals and groups performing ransomware as a service sell their talents and outputs to others, who will then go on to use them. This may technically include distribution methods, such as botnets, which are used to send malicious software to potential victims, though this may also fall under malware as a service (depending on the goal of the purchasing party).

1 Introduction

Taking possession of items, persons or information that is of value to others is a well-documented criminal activity, with a strong profit motive. As we have progressed further into the digital and information age, this practice has been adapted to meet the opportunities presented by the internet. Ransomware is particular kind of malware that exploits trust and/or limited technical knowledge to take control of a target computer, encrypting data and holding it to ransom.

How this is achieved varies, but the profit motive of the perpetrators provides a common element regardless of the infection vector. The drive to extract cash value from the victim of a ransomware infection means that economic theory and models can be applied to identify the attributes and efficacy of a given strategy. Threats, analysis of optimal demands to submit to a target, and the evolution of ransomware to exploit new opportunities (such as IoT devices), are all factors that have economic motivators and which may be employed by criminals to increase the profitability of ransomware.

This report focuses on these economic phenomena and analyses the use of strategies that involve threats, negotiation, price discrimination, and more, to help define the current state of ransomware. This will aid in identifying the next optimal steps in ransomware development, serving as the foundation for ongoing work in predicting the evolution of ransomware as a profit-driven criminal enterprise.

Part 2 introduces the business aspects of ransomware, such as operating costs and markets that service the needs of cyber-criminal enterprise. Initial cost and key elements of profit extraction are key discussion points. Part 3 focuses specifically on economic considerations of ransomware, discussing the methods by which profit is maximised in traditional ransom scenarios and applying this to the digital domain. Part 4 shows the results of a survey, in which affected individuals are asked if they paid ransoms, amongst other salient questions. Results are shown to demonstrate the efficacy and attributes of existing strains of crypto-ransomware. Part 5 concludes the report, providing an overview of the discussion and outlining our upcoming steps towards a definition of optimal ransomware strategies and related model.

2 Cost Analysis of Ransomware

Drawing on evidence from Cryptolocker, Cryptowall, Teslacrypt and other major ransomware strains, we present an economic analysis of ransomware. It is important to clearly define the scope of our discussion: ransomware is a term added to the Oxford English Dictionary in 2012, which is applied as a term for any malware that asks for a ransom after infecting a target. This is a very general definition, and unsuitable for this report. All manner of extortion strategy is described by this term, including those which compromise a target by placing illegal materials on their PC and threatening to call the authorities if a ransom is not paid.

This report focuses on the subtler sub-class of ransomware originally known as crypto-virus, but later redefined as crypto-ransomware. This malware works by encrypting the contents of a target computer, and offering a key to decrypt the assumedly important data in return for a ransom. This may be augmented with threats of deletion, dissemination of data.

Adam Young and Prof. Moti Yung presented the first concept of cryptovirus in 1996 (Young and Yung, 1996). Likely building on previous failed attempts to extort computer users by the AIDS malware, Young & Yung implemented public-key cryptography to perform their proposed extortion in a sound and robust manner. Cryptolocker represents one of the first successful attempts to implement Young & Yung's proposed method. Discovered in the wild in 2013, it has precipitated a deluge of ransomware.

Diversification is rife, with many families claiming a common root branching off either technically or in the strategy employed at the negotiation/profit-extraction level. Figure 1 shows a recent (February 2017) map of the current known ransomware families. The FBI suggested in 2016 that this could be a $1 billion enterprise. This explosion in development, deployment and profit-extraction shows that crypto-ransomware is a thriving criminal enterprise.

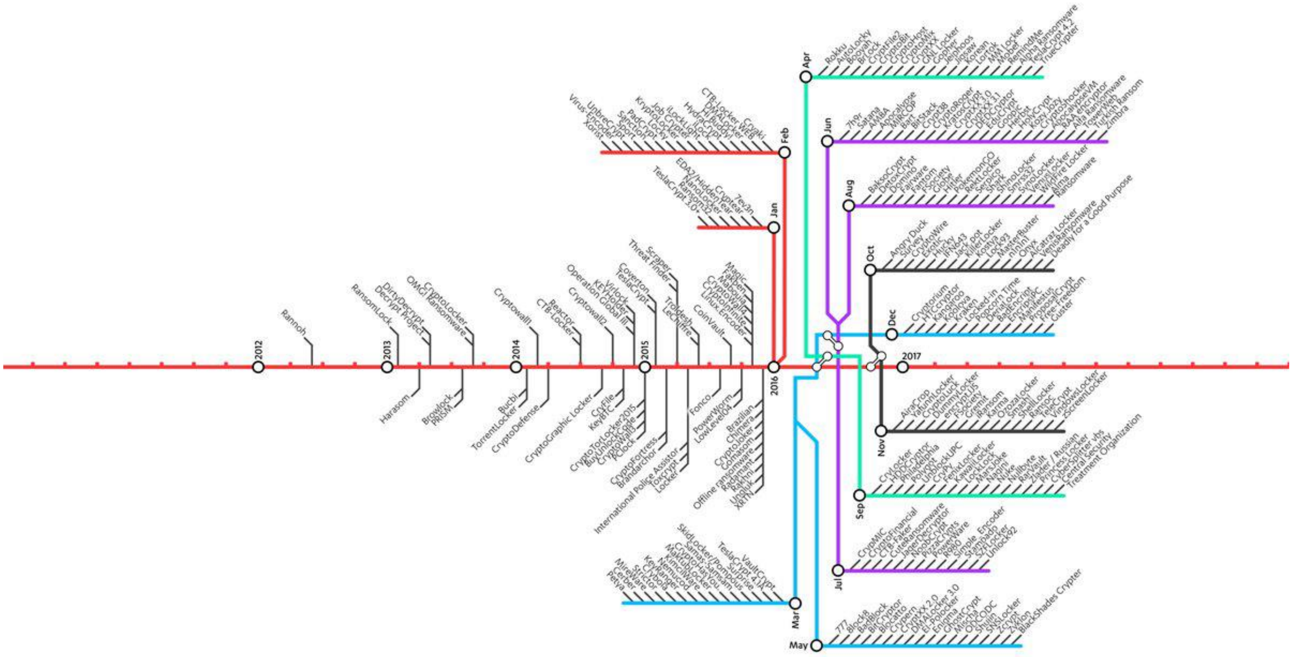


Figure 1 Ramsomware Tubemap from F-Secure 2017.

2.1 Initial costs

In any enterprise, the set-up costs have a significant role in informing the strategies employed to recoup those costs and generate profit. The higher the cost, the higher the profit-per-customer or frequency of custom must be to satisfy the needs of the business.

When considering crypto-ransomware, these initial costs are largely technical. It is not trivial to develop ransomware that is successful; many attempts to do so commit the cardinal sin of using private key cryptography, which is easily reverse-engineered (extracting the private key and making it useless). Such attempts may work on the uninformed, but this greatly reduces the pool of potential victims and significantly decreases the lifespan of the ransomware (as the deficiencies in its design become widely circulated quickly). If a solution that does not involve paying the ransom is found quickly, it is likely that the cost of development will exceed the profit potential of such a strain of crypto-ransomware. TorLocker (2014) is a good example of this, with 70% of observed cases allowing for key derivation despite AES-256 and RSA-2048 encryption. Cryptear, EDA2, and many others make this same mistake, rendering their profit potential low-to-none (Hahn, 2016).

Cryptolocker is a notably different strain, as it is sufficiently well-engineered to force victims to pay. Key derivation is not an option for a typical Cryptolocker implementation, which vastly reduces the alternatives for the victim and maximises the potential-for-profit for the attacker. Botnet distribution, using Gameover or Zeus, maximises the spread of this ransomware, with a ‘blanket’ approach to infection (Hernandez-Castro, Cartwright and Stepanova, 2017). By attempting to infect as many machines as possible, the chances of payment from some of the targets is increased. Sale of bandwidth from these botnets is a recent development, but represents another set-up cost for the ransomware business.

This approach has proven to be worth the potential cost, due to the high conversion rate of victims. The 2016 Kent Cyber Security Survey found that 26% of infected victims acknowledged that they paid the ransom. The 2015 Kent Cyber Security Survey suggested an even higher conversion rate, 35%!

As is common in cybercrime, copycats have emerged to take advantage of the attributes of Cryptolocker. This indicates a successful business model, which is operating more than its initial and operating costs. A recent phenomenon is the emergence of ransomware development kits selling in the black market, and marked increase in Ransomware as a Service (RaaS). Until the cost of setting up and operating such criminal enterprise is made to be larger than the potential profit, such activity will continue to increase in popularity.

2.2 Cryptocurrency

Cryptocurrencies, such as Bitcoin, have played a fundamental role in the success of Cryptolocker and its ilk. Traditionally, payment methods such as UKash and PaySafeCard have been used as a way for individuals to pay ransoms using fiat currency over a digital medium. However, the anonymity and ease of transfer provided by Bitcoin has led to a decline in such methods.

Creating wallets is trivial, and wallets that are not used in conversion to or from fiat currency (using Bitcoin to purchase £ for example) aren’t easily tied to identifying information. The only common identifier for most cryptocurrency wallets is their public key, which is used as a deposit address. This information is provided when demanding a ransom, and multiple addresses are used in many implementations.

Only leaving one step between ransom and cashing out would likely compromise the ransomware operator, as it would be possible to trace the fiat-facing transaction. As a result, a sophisticated laundering process is employed, with near universal steps taken to reduce the opportunity for law enforcement agencies to successfully intercept the entirety of the proceeds of this criminal activity.

After receiving a set amount of currency, these wallets transfer their balance through a series of wallets which commonly only have 2 transactions each. The first of these is the receipt of bitcoin and the second is to split the incoming value between two other addresses. This is known as a ‘peeling chain’ and is common to Bitcoin. Observations of the scale and near-identical timing of such transactions indicates that the laundering of Bitcoin in this manner is automated. This is known as mixing or tumbling, and is an enterprise that demands a 2.5% fee from the laundered balance. This represents an operational cost, as does the cost of transfer incurred when any bitcoin transaction is undertaken.

Sean Sullivan suggests that Bitcoin friction may be ransomware's only constraint. This friction is generated by individuals who are willing to pay but unable to access bitcoin, either due to some constraint (such as lacking sufficient personal identifying information to make the initial transaction from an exchange), or inability to acquire the currency in the (sometimes short) allotted time. The volatility of Bitcoin is another factor, as Bitcoin can rapidly increase or decrease in value on an hourly basis. This puts pressure on ransomware operators to constantly price adjust and provide short periods of time between initial infection and ransom deadlines. These factors reduce the total potential profit of the enterprise, regardless of the ransom strategy itself; they represent technical challenges in the medium of exchange.

2.3 The Cost of Customer Service

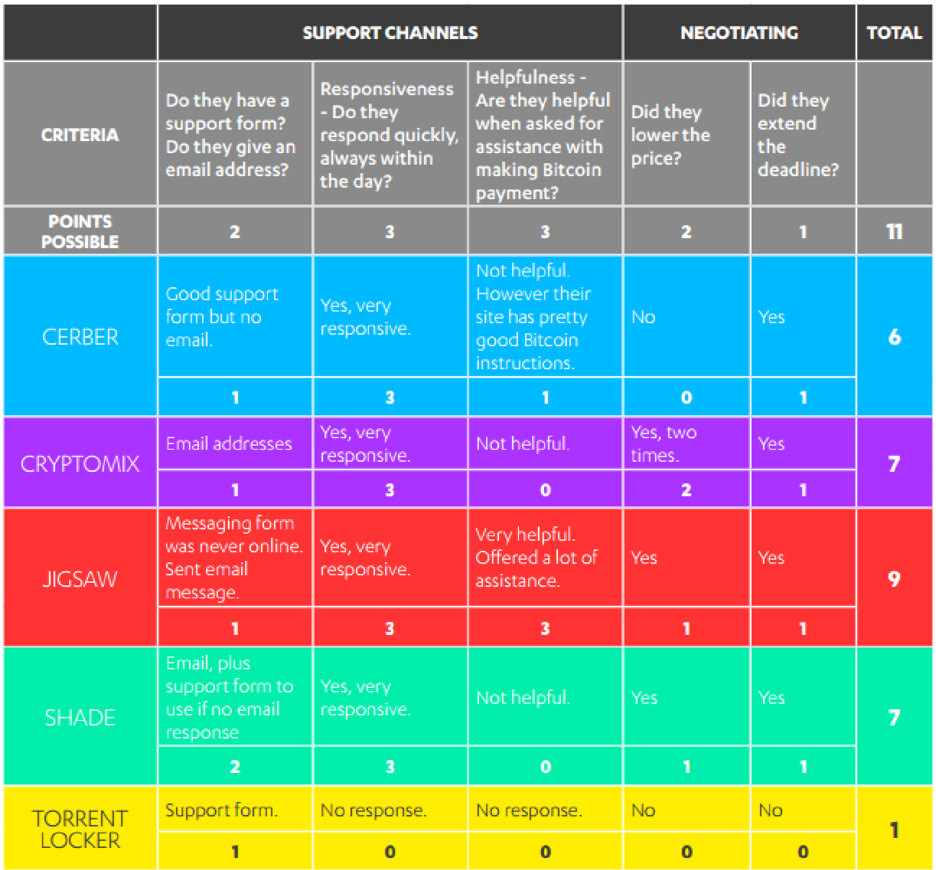
Ransomware is a transaction based system, it requires two parties to come to an agreement. Even if one party is under duress, the other must ensure that every opportunity is given for the willing victim to pay their ransom. Strong arm tactics are effective in showing what happens if people do not submit to the demands of the ransomware operator, but are not conducive to the final stages of a successful ransomware transaction: the payment of the ransom.

F-Secure performed a study in 2016, focusing on the customer-relations aspects of ransomware. Five families of ransomware were analysed, all of which requested Bitcoin payments. Using non-technically inclined participants, the researchers reached out to the operators of each ransomware to pay the ransom, evaluating the customer service elements of the transaction.

Scoring of the ransomware itself was performed on 4 metrics: professionalism, instructiveness, language support and free trial decryption. The look of ransomware can go a long way to providing some feeling of professional integrity, although ransoming is a criminal act the display of professionalism can communicate integrity to a potential payee. The researchers found that a more professional looking ransomware would be seen more favourably by non-technical victims, than text or HTML pages. The development of an appropriate front-end, therefore, represents an overhead cost that may provide returns should the initial outlay be provided by the operator.

As mentioned, non-technical subjects interacted with ransomware, to emulate the likely scenarios that might unfold in an uncontrolled environment. Detailed and informative ransomware was found to provide a better level of customer support than copied or uninformative interfaces. Customers who know how to pay and what it will cost them are generally more cooperative. This represents another upfront cost, though it may just be a matter of operators conducting research and putting time into the wording, design and formatting of their family of ransomware.

Language support was found to be lacking. CERBER was found to have 12 supported languages, but all others supporter 2 languages at most (English in all cases, Russian for SHADE and Dutch for TORRENT LOCKER). By comparison, all but one offered a free trial decryption, which would allow the victim to see that they could recover a file. This demonstrates the fact that although ransomware is itself disreputable, reputation is everything when trying to coerce people into paying.



	SUPPORT CHANNELS			NEGOTIATING		TOTAL
CRITERIA	Do they have a support form? Do they give an email address?	Responsiveness - Do they respond quickly, always within the day?	Helpfulness - Are they helpful when asked for assistance with making Bitcoin payment?	Did they lower the price?	Did they extend the deadline?	
POINTS POSSIBLE	2	3	3	2	1	11
CERBER	Good support form but no email.	Yes, very responsive.	Not helpful. However their site has pretty good Bitcoin instructions.	No	Yes	6
	1	3	1	0	1	
CRYPTOMIX	Email addresses	Yes, very responsive.	Not helpful.	Yes, two times.	Yes	7
	1	3	0	2	1	
JIGSAW	Messaging form was never online. Sent email message.	Yes, very responsive.	Very helpful. Offered a lot of assistance.	Yes	Yes	9
	1	3	3	1	1	
SHADE	Email, plus support form to use if no email response	Yes, very responsive.	Not helpful.	Yes	Yes	7
	2	3	0	1	1	
TORRENT LOCKER	Support form.	No response.	No response.	No	No	1
	1	0	0	0	0	

Figure 2 F-Secure 2016 report, scoring ransomware families on customer service (F-Secure 2016)

Figure 2 shows the second stage of analysis, in which F-Secure evaluate the customer service of each family. Many of the listed families provision for some form of interaction with their victims, with only TORRENT LOCKER failing to provide any means of contact. JIGSAW was found to be the most helpful, with quick response times, plus helpful advice on how to obtain currency and send it. However, JIGSAW also lowered their price and extended deadlines, both good customer service moves, but also moves that lower optimal profits (a more developed insight into this phenomenon is provided in Section 3).

Customer service can be said to have two main types of cost:

- The cost of provision. The can include the development of appropriately professional web presence and communication channels, along with the expertise required to operate them. Cost of maintenance, such as hosting and manning call desks (or equivalents) is also a consideration.
- The cost of over-provision. Being ‘too nice’ can lower income by giving away value on the ransom itself. Lowering ransoms can have the effect of reducing the ceiling for your ransom (individuals talk about the fact they got a lower ransom, so others expect the same). Providing more time may affect the throughput of your customer service team, especially if you provide good service. This may mean you make less money over time, despite having a higher proportion of willing payees. Section 3 considers this phenomenon more closely.

In an activity that requires a reputation to be built on top of a disreputable act, customer experience is critical. However, the cost of provision must be calibrated to maximise willingness to pay, without over-spending. Optimal customer service requires that ransomware operators understand when their efforts offer diminishing returns.

2.4 Shutdown and Obsolescence: The Ultimate Cost

Operation Tovar, led by the US Department of Justice and the FBI, with the assistance of Europol and the UK NCA, Australian Federal Police and other law enforcement agencies, shut down the Gameover/Zeus botnet. This was the one of the main vectors for Cryptolocker distribution, and its takedown represents the end of this strain of crypto-ransomware. This operation recovered a file containing 500,000 infected person's details, and associated keys. This provides evidence that at least this strain possessed keys associated with victims, which could be distributed upon successful payment. It was, in this case, used to decrypt the target computers, denying the ransomware operators any income from those listed in the file. It is important to note that it was not possible to decrypt the affected machines by other means; possession of this file allowed decryption, but the ransomware itself was proof against reverse engineering to extract key details.

Operations like this show that the operation of ransomware can be disrupted, but significant effort is required on the part of law enforcement agencies. Cryptolocker extracted a large volume of money through use of a cryptovirus, with many successors following it due to the profits associated with this activity. The diversification of approaches doesn't just include technical innovations, but also the economic methods employed, particularly in the extraction of ransoms. Cryptolocker was relatively unsophisticated in this regard, and modern variants have far more sophisticated means of generating wealth from ransoms, be it through increase likelihood of payment, higher value per payment or more complex means of calibrating demands to encourage the optimal frequency and value of pay outs. Shutdown representing the ultimate cost, in so far as one must go and redevelop one's ransomware, distribution network and base of victims, may become inevitable if sufficient state sponsored countermeasures are brought to bear, but only on a strain-by-strain or distribution method basis. Ergo, the continued diversification in technical, distribution and ransom strategies to extract maximum value before likely obsolescence.

2.4 Conclusion

The costs associated with ransomware depend on the following key factors:

- Technical cost of development
- Cost of dissemination
- Cost of doing business

The technical cost of ransomware depends largely on the strategy selected when choosing to launch such an enterprise. The development of ransomware is costly in terms of skills and time: the individuals involved in the development of ransomware possess technical capabilities that they may be able to better employ for greater profit elsewhere, at least in the short term. The purchase of ransomware is usually cheaper, and therefore more attractive to a wider audience of prospective ransomware operators.

The cost of dissemination represents the cost of owning, running and developing the hardware for dissemination of ransomware. It may also represent the cost of offloading this to botnets, such as Zeus or Gameover. This is a cost that is incurred if the operators wish to continue spreading their ransomware implementation.

A key operational cost is the need to provide customer service. Ransomware is a criminal act, and so by their nature operators are disreputable (and are likely to be perceived at such). However, they must convince their victims that they will honour ransoms and return encrypted data, essentially trying to build a reputation. Although a scatter-shot approach, in which ransomware operators do not communicate with their marks after infection, may work; it has been found that customer service increases the chances of successful payment (F-Secure 2016). This is a dual-purpose expense on the part of operators; the less technically adept can be guided into payment, and the previously unwilling but communicative can be convinced to pay. However, this places additional overheads and running costs on the operators, as they must provide an appropriately professional interface, and potentially provide responsive, accessible points of contact for queries and negotiation.

Finally, the cost of doing business. To extract value from ransoms, while avoiding the ultimate cost of identification and arrest by law enforcement agencies, one must launder the proceeds. A common way of doing this is to use cryptocurrencies, such as the well-known Bitcoin. Costs incurred by using Bitcoin include the transaction fees for moving Bitcoins between wallets, and the 2.5% fee demanded by most laundering services. Where the operator runs their own laundering enterprise, the transaction fees and cost of developing the require autonomous systems for rapid movement of currency take the place of laundering fees. As a result, some cost in the profit-extraction process is always incurred.

It is evident that the cost of implementing and operating ransomware is below the potential for profit, as the diversity and frequency of ransomware attacks has been on the rise for years now. The rise of RaaS indicates that offloading technical and development work onto those willing to sell their software instead of running their own ransomware operation is also becoming increasingly common. As a result, there is compelling evidence that ransomware operators, largely through trial and error, are optimising their costs.

Having identified the costs associated with ransomware, we analyse the economic attributes of ransomware. By identifying the strategies employed to extract value from victims, and the associated risks, costs and potential proceeds, we can begin to identify the current state-of-the-art and likely optimal path for further development of crypto-ransomware.

3 Economic Analysis of Ransomware

As previously discussed, profit is the most likely attribute that cyber-criminals will seek to maximise. The willingness of victims to pay a ransom plays the largest role in the value of ransoms as an aggregate of all demands for payment. There are many factors that affect this value:

- How much the victim values their files
- Ethical considerations, such as willingness to pay money towards criminals
- The level of trust, whether the victim believes that the criminals will honour their word

These elements can be combined into the simpler attribute: *willingness to pay*. This is unique to each victim, but by abstracting to this single variable, each individual (i) can be included in a set of victims (v). This can be summarised as:

$$\Pi \ = \sum N\ i = 1\ (pi\ - c)1i\ - F$$

N is the number of victims. *P* represents payment, which is a set of values representing individual payment values, representing the amount paid by a given victim. 1_i is an indicator variable that take value 1 if $pi \leq vi$ and 0 otherwise. *F* is the fixed cost of operating malware, and *c* is the cost of dealing with any ransom money (laundering fees, for example). In our work, we assume that the ransomware operators will focus on targeting as many people as possible, rather than specifically targeting a demographic or other numerically constrained set of targets. The focus of this work is to discuss the optimal ransom to charge victims, based on the current state-of-the-art.

3.1 Uniform Pricing

To maximise profit, an optimal ransom must be identified. This can be difficult when considering diverse demographics with an unknown composition. The operators don't know the average income, target machine value or perceived value of data. To put it simply, they don't know the individual's willing to pay, *vi*. In such an environment, uniform pricing offers a low-effort solution to maximise the likelihood of payment, at the cost of total potential payment from all cases where *vi* exceeds the threshold *p*: the uniform price of ransom. Profit, in this case, can be expressed by:

$$\Pi \ = \ (p - c)Q(p) - F$$

Q(p) represents the number of people willing to pay a ransom *p*. In this context, *Q* can be considered a function detailing demand for this service for all possible prices, solicited though these services may be. The optimal ransom may then be defined as:

$$\frac{p-c}{p} = -\frac{1}{n(p)}$$

$n(p) = \frac{p}{Q(p)}\frac{dQ(p)}{dp}$ describes the price elasticity of demand. This allows operators to measure the sensitivity to changes is the ransom. It is most optimal to price where demand is elastic, as this gives the operators the opportunity to change the ransom to calibrate the profit by raising the ransom in all cases where the rise is less than a proportionate drop off in payments (Pepall, Richards and Norman, 2008). It may be counter-intuitive at first glance, but the optimal ransom will have less than 50% of targets pay. This is because it is likely that as *p* increases, *vi* decreases at a lesser rate. This means that the optimal profit lies below 50%, though there will be a threshold point at which profit falls below the optimal due to non-payment by a significant majority of victims.

Criminals will not know these variables, but they may make general hypotheses about them, with a good likelihood of sufficient accuracy. Suppose that the price of the ransom was £300 and 40% of victims pay: *Q(300) = 0.40N*. Increasing the price to £350 reduces the chance of payment to 35%. If we assume *c* is £10 then the current demand is too inelastic and we'd be best served by raising the ransom. Data can be obtained by varying the ransom over time, to better inform future efforts to optimise profit.

There is currently no evidence that this strategy is in use, and it is (admittedly) counter-intuitive. However, we are confident that in the coming years, examples of this strategy being employed will begin to emerge, as ransomware operators collect sufficient data and expertise to identify the means of seeking optimal profit under uniform pricing schemes.

3.2 Price Discrimination

The ability to distinguish between individual willingness to pay has been, so far, assumed to be beyond the operator. This is, usually, a reasonable assumption, as the means to profile victims is beyond most examples of crypto-ransomware. However, it is possible to gain reasonably detailed information about victims through a variety of sources, such as trojans, keyloggers, and other malware that can access browsing history and personal data. Using such information to build a profile of the victim can aid in the adoption of so-called price discrimination.

Under uniform pricing, one is constrained to a maximum optimal profit that depends of the willingness to pay of all affected individuals. Price discrimination adds value by calibrating the demanded sum to the value expected to be considered worth paying by the victim.

Three types of price discrimination are defined by economists (Varian, 1989). First degree (or perfect) price discrimination involves tailoring prices to individuals. This is highly intensive when considering data collection. There is no evidence of this approach being used in the context of ransomware, at present. It is a 'gold standard', but very resource intensive.

Second or third-degree price discrimination is far more realistic. Second degree price discrimination involves offering a diverse array of services and a price 'menu' to go with those services. This could allow a victim to choose what data they wish to retrieve, lowering their price at the cost of the data they choose to leave behind. However, this requires a lot of control on the part of the ransomware operator, and so may be considered costly and less feasible than third degree price discrimination.

Third degree price discrimination requires only that key attributes of a victim are used to determine a price calibrated to the individual. This means that the number of files, file titles or similar attributes can be used to make informed guesses about the importance of the data to the individual in question. This is made simpler by categorising victims into types, and pricing based on their perceived membership of a given type, rather than trying to make judgements at the individual level. A simple example is outlined in the equation below:

$$\frac{pL - c}{pL} = -\frac{1}{nL(pL)}$$

pL is the ransom set for anyone with many files. nL is the price elasticity of demand for people with many files. The same can be done, replacing L *with* S for few files, making this a two-category third degree price discrimination system.

Charging two different prices increases profitability by making the ransom more appealing to a given group of people. For fewer files, unless quality of data is considered, a smaller ransom is likely to be more appealing. As with uniform pricing, the operators can vary the ransom over time to gain more information about their demographics, and identify price-points at which they can set finer-grained divisions of their victim-base. Effectively, price discrimination allows for a wider pool of victims who are willing to pay, while not reducing the value of the ransom that more willing victims might pay. In this respect, it is a form of value recapture: the operators find the point at which sets of victims will pay and thereby recapture value that would be lost to a uniform price scheme.

Interestingly, some ransomware families are now making use of tool that profile the victims machine to better place them in the operators pricing scheme. Shade uses a Remote Access Trojan (RAT) to spy on victim finances and estimate the ransom that can be paid (Atanasova, 2016). This is a poor way to implement price discrimination, as it is technically demanding, requires a long time to gather appropriate data, and opens the possibility for being traced through one's own trojan.

Fantom ransomware determines ransom amount by the name of the process/executable file (Abrams, 2016). This allows for multiple campaigns to be launched simultaneously, with different price points to appeal to different victims.

One can also target different business or sectors with spam campaigns, so that existing public information can be used to determine what your targets might be willing to pay. This is relatively unsophisticated, but can extract information like the role of the victim in the business; providing information about the potential impact of not paying the ransom and thereby allowing inference of an appropriate price.

It is evident that price discrimination is being adopted as a profit optimisation strategy. We expect the sophistication and efficiency of reconnaissance operations and 'in-line' price discrimination techniques (such as email/business profiling in spam campaigns) to increase, partly as a response to increased competition by other ransomware operators.

3.3 Bargaining

The criminal element of this discussion has been explored, but what of the victims? We have discussed the basic role that the victim plays as the passive actor in the price-setting process, but it is common for victims to play a far more active role in optimal pricing of ransoms. It could be argued that every ransomware infection involves the victim as the active arbiter of whether the attempt is profitable or not, though possibly at great personal cost if the ransom is refused. However, in this case, we would like to discuss the victim as an active role in the negotiation of a price.

The classic model of bargaining assumes that both parties make alternating offers (Rubinstein, 1982). This model is general enough to capture a wide range of potential bargaining scenarios (Muthoo, 1999). If we assume that vi is known by the ransomware operator:

$$pi = \left(\frac{1-\delta_B}{1-\delta_A\delta_B}\right)(vi-c)+c$$

$\delta_A \leq 1$ is the discount factor of the criminals and $\delta_B \leq 1$ is the discount factor of the victim. It may be assumed that the victim has more to lose form a delay in the time to agreeing a ransom, compared with the attacker. This is because the attacker will likely be conducted attacks on a larger scale and over an extended period. This results in the attackers perceived cost of delay (discount factor) being lower than that of the victim. This gives the attackers a strong bargaining position. This can be enhanced by changing the rules of engagement.

Alternating offers may be the default mode of bargaining games, but the attacker may make a take-it-or-leave-it offer, at which point the game changes to an ultimatum game (Binmore et al., 2002). If such ultimatums are made close to vi, then the attacker can extract maximum value from the victim by stopping the bargaining game when they feel that they have reached the point of willingness to pay. However, unreasonable offers are likely to be rejected in an ultimatum game, which will result in the loss of ransom as it is essential that the ultimate deprives the victim of choice to continue (Thaler, 1988). Breaking from this strategy weakens the position of the attacker, possibly in other negotiations if the victim communicates that an ultimatum was backtracked to a further round of negotiation.

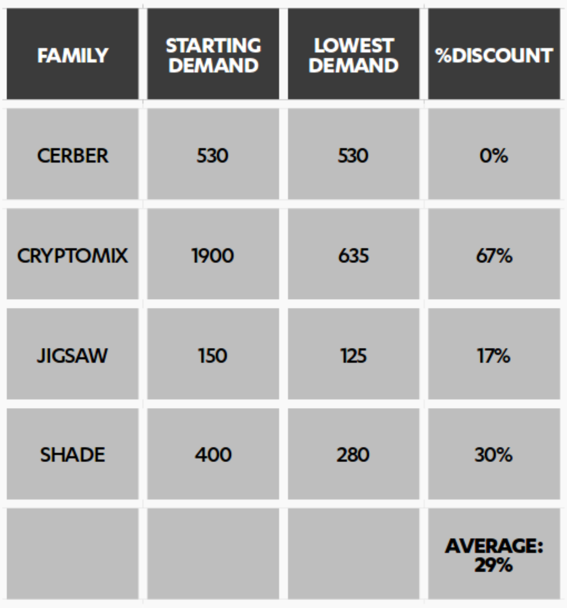
In a realistic scenario, the attacker will not know vi. This means that they will not know where to stop the bargaining phase of negotiations. The Coase conjecture is vital to understanding how negotiations unfold in such an environment (Coase, 1972). The criminals derive an optimal ransom *p* and make a demand at that value. If the victim rejects the demand, should the attackers return with a lower offer?

Intuition suggests that they should, and this is true for one-off interactions. However, in interactions with others that may have communicated with a previous victim who negotiated a better price, it is more likely that the victim will feel that they can also negotiate. The Coase conjecture says that any anticipation of a discount will weaken the attackers bargaining position (Gul, Sonnenschein and Wilson, 1986; Gul and Sonnenschein, 1988). In order to maintain a strong bargaining position, the attacker needs to either refuse negotiations, or have a reputation for toughness and retaliation against over-zealous counter-offers (Kennan and Wilson, 1993).

Jigsaw ransomware creates a very real sense of peril for the victim, by deleting a chunk of data whenever a given period (72 hours in this case) elapses. This accelerates and increases the cost of delay for the victim and strengthens the attackers position. Copycat ransomware that may or may not delete contents, while threatening to do so (Razy 5.0 being an example of one that threatens but doesn't follow through), also exists. These threats are so credible that many business (up to 33% of businesses with 250 or more employees) hold Bitcoin as part of their contingency plans (Parker, 2016).

It is important, for the attacker, that there is not a perceived end date or finite number of attacks. A paradox known as the Chain Store Paradox describes an effect, where the attacker is perceived as likely to give in to negotiations if there's a likely time at which they'd stop attacking (thus losing all potential profit after that date) (Milgrom and Roberts, 1982).

The figure below, shows the profitability of several types of malware. Some of them allow negotiation, others do not. An F-Secure employee using the simple phrase "That's too expensive, I don't really need the files that bad anyway" led to the discounts shown (F-Secure, 2017).



FAMILY	STARTING DEMAND	LOWEST DEMAND	%DISCOUNT
CERBER	530	530	0%
CRYPTOMIX	1900	635	67%
JIGSAW	150	125	17%
SHADE	400	280	30%
			AVERAGE: 29%

Figure 3 Some examples of ransomware open to bargaining (F-Secure, 2017)

According to our economic analysis, engaging in bargaining is, although tempting, a sub-optimal strategy. Cerber, which does not offer the option to negotiate, will eventually generate their authors more profit, and is likely to be adopted into the next generation of ransomware families.

3.4 Determinants of Willingness to Pay

As previously discussed, the success of a given ransomware strategy is dependent on the willingness of the victims to pay. There are, however, several variables that can change between different strains of ransomware.

Ransomware operators can choose whether they return some, all or no files. Trustworthiness factors into whether the victim believes that payment will be rewarded with the return of their data. To maximise profit, criminals should endeavour to return files unfailingly. Being unreliable will have an adverse effect on profitability, as people consider their files forfeit and do not want to be scammed as well as ransomed.

The way in which the ransom demand is framed can also have a significant impact on the likelihood of a pay-out. People are risk loving over losses and risk averse over gains (Pfleeger and Caputo, 2012). If the operator of the ransomware emphasises the imminent threat of losing files, victims are more likely to pay out. To maximise profit, it is best to emphasise that files will be lost for non-compliant behaviour, then to suggest that they will be recovered if you pay. This is further reinforced by evidence that files will be returned (such as stories from previously infected individuals who paid the ransom). With a reputation to returning files, this fact becomes secondary to the need to emphasise the price of non-compliance. This makes it best to impress on victims that their files will surely be lost without payment of the ransom.

Knowledge of the ecosystem in which the ransomware exists, on the part of the victim, has a negative effect on payment. As the number of victims refusing to pay directly reduces the profit of ransomware, potentially preventing further infections due to lack of profit motive, victims knowledgeable of such facts are less likely to pay. There is little an attacker can do about this, as they are unlikely to communicate such information anyway, but lack the means to limit the spread of such information.

The value of files is another victim-controlled element of the equation. Victims with recent backups are unlikely to pay ransoms, simply because the costs associated with restoring files are lower than any likely ransom. It is important to point out that individuals who back up their files benefit the entire affected population. They reduce the impact of the ransomware, forcing expenditure on individuals who are guaranteed to not pay as they have the means to circumvent the ransomware. As this method is shown to be effective, and is communicated, it becomes harder to find victims who will pay a ransom that meets the profit requirements of the ransomware operator. This has the effect of reducing the ransom that can be reliably demanded, and the number of people who will pay it.

It is obvious that backing up files should be undertaken to prevent oneself from falling prey to a variety of malicious and incidental data loss vectors. Less obvious is the role of positive externalities. By encouraging strong social norms to backing up files and not paying a ransom (appealing to a sense of empathy and community by sharing information about how paying just propagates ransomware) the profit potential of cyber criminals can be greatly reduced.

3.5 Game Theory and Kidnapping

The previous sub-sections describe pricing strategy and tactics to maximise profit, focused on the willingness of the victim to pay. However, there are further factors to consider, which may be elegantly addressed using game theoretic modelling of kidnapping scenarios. It may be considered that the files are a kidnap target of value to the victim of ransomware. Ergo, it is possible to apply the rules of a simple game of kidnapping to the problem of ransomware. We explore two models of ransoming in a game theoretic context: the first focusing on optimal ransom (Selten, 1988) in much the same way as our previous discussions. The second model focuses on whether potential victims should act to deter hostage taking (Lapan and Sadler, 1988). In our discussions, kidnapping and encryption of files for ransom are considered equivalent. This work draws on research undertaken by Hernandez-Castro, Cartwright and Stepanova (2017).

Reinhardt Selten (1988) proposes a simple game of kidnapping, which we shall refer to as the kidnapping game. It is conducted in six stages, outlined below:

1. *The criminal chooses whether they will infect the victim's computer. Choosing not to infect the computer*

ends the game with a payoff of 0.

2. *Assuming the machine is infected, then the criminal sets a ransom demand, D ≥ 0. This demand is sent to the victim.*
3. *The victim receives demand D and chooses a counter-offer C ∈ [0,D]. It is not clear if it is in the interests of the attacker to allow a counter-offer (previous discussion suggests strongly against this for optimal profit). It is merely assumed, at this point, that the potential for this action exists.*
4. *Irrational aggression on the part of the criminal may result in the files being destroyed during negotiations or for other reasons. The probability of this is, according to Selten (1988), a = a(1-C/D) where a ∈ (0,1) is a constant. This is a natural event in this game, without preceding motive on the part of any player. The destruction of files results in a payoff of -Y ≤ 0 for the criminal and -W < 0 for the victim.*
5. *If the files are not irrationally destroyed, the criminal must choose between releasing the files and receiving C, or destroying the files and receiving 0. The criminal may be considered as possessing a minimum acceptable offer M. If C ≥ M, the files are released and otherwise they are destroyed. This simple game doesn't account for when criminals take the ransom and destroy the files, which will be discussed later.*
6. *The probability of being caught by the police (q), exists. This probability is considered to be independent of the actions of the criminal. If the criminal is caught, having not destroyed the file, their payoff if -X < 0 if they received a ransom and -Y < 0 is they did not. A harsher payoff of -Z < -X is applied in any case that resulted in the files being destroyed.*

The potential outcomes of this game are summarised below:

Table 1 The payoff to different outcomes in a simple game of kidnapping

Outcome	**Payoffs**	
	Criminal	**Victim**
Criminal doesn't infect computer	0	0
Release of files for *C*	*C*	-*C*
Files destroyed	-*Y*	-*W*
Criminal caught after release of files	-*X*	0
Criminal caught after destruction of files	-*Z*	-*W*

This demonstrates that although it may seem that taking a ransom and running (not returning files) is an optimal strategy as it doesn't require one to expend effort on returning the files, it is in fact quite the opposite. By not returning files, criminals reduce their credibility and increase the potential costs of capture by the authorities. We also previously discussed the phenomenon of positive externalities. From the victim's point of view, a criminal reputed to not return files isn't worth a ransom, and so the initial offer will be refused. As the victim will not pay, there is no incentive to infect to begin with, as this represents effort expended for no gain. This effectively diminishes the criminal's opportunity to profit from their activity.

As with our previous example, this game assumes complete willingness to pay. This assumption doesn't stand in the real world, but as discussed during our analysis of price discrimination, it is possible to derive attributes about individual victims, or sets of victims. This requires more work on the part of the attacker, but may be the product of experience. Longer-lived ransomware operations will accrue greater information about the types of victims that will pay and how profit is maximised, gaining a clearer idea of what the willingness to pay variable is.

Imperfect information does not affect how the game is played, but has an impact of the likelihood of the victim recovering their files. The better the criminal can infer the willingness of a victim to pay, the more likely they are to reach an accord and secure both the value of C for the attacker and the files for the victim.

This brings us neatly to the second game: Lapan and Sadler's (1988) model of whether to bargain. This is a four-stage game that greatly extends our previous definition of the bargaining game in sub-section 3.3:

1. *The potential victim chooses how much to spend deterring an attack. This may represent virus protection, training for staff and greater care about opening unsolicited files. Expenditure is denoted by E ≥ 0.*
2. *The criminal chooses whether to attack. Not attacking ends the game. Attacking creates a payoff of 0 for the attacker and a pay off of -E for the victim.*
3. *If the criminal chooses to attack with a probability* $\theta(E)$ *the attack fails, where* θ *is a continuous, monotonically increasing function of E. With probability* $1-\theta(E)$*the attack succeeds and the victim's files are infected. Failure on the part of the attacker ends the game. The criminal has payoff* $-F\ <0$ *and the victim has payoff* $-A-E$*, where* $-A\ \leq 0$.
4. *If the attack is a success the criminal demands C as a ransom. The victim can pay or not pay. If the victim pays, they regain their files. The victim's payoff is -C-E and the payoff of the criminal is C. If the victim doesn't pay, the files are destroyed. The victim's payoff is -W-E and the payoff for the criminal is* $-L\leq 0$.

In a pure interpretation of the game, the attacker has the choice to kidnap. The victim has the choice of how much to spend on deterrence and whether to pay a ransom in the case of a successful attack. The table below outlines the potential outcomes for this game.

Table 2 The payoff of different outcomes in a game of kidnapping with the possibility of attack deterrence

Outcome	**Payoffs**	
	Criminal	**Victim**
No attack	0	*-E*
Failed attack	*-F*	*-A-E*
Release of files for ransom C	*C*	*-C-E*
Ransom not paid	*-L*	W-E

Incomplete information has a similar effect as that previous described for the game inspired by Selten. This now applies to the victim, as they do not have complete knowledge of attacker capabilities. Zero day threats may circumvent defences that only focus on patching known exploits or enforcing policies that protect against known tactics. This can lead to victims that spend too little on protection, increasing the chance of an attack, or who over-anticipate an attack, which raises their operational costs significantly above the optimal. As a result, both attacker and victim must make hypotheses regarding the benefits of expenditure and ransom values.

Another impact of deterrence is the promotion of previously discussed positive externalities. As deterrence reduces the profitability of an attack overall, it has a beneficial effect on all potential victims, as the ransomware is forced to ask higher prices to remain optimal, or potentially fall below the threshold that the attacker considers it worth continuing to launch attacks. However, low-expenditure victims may effectively 'free-ride' on the higher expenditure of more cautious victims.

The concept of low-expenditure victims as free-riders may be unfair. The game doesn't explicitly account for the victim lowering the cost of W by backing up files. Therefore, it is possible for low-rolling victims to lower their W in large numbers, while high-rollers such as corporations, can spend more on E and augment the positive externalities that both approaches generate. Ensuring a healthy culture of backups and entity-appropriate cyber-security expenditure can greatly diminish the profitability (and motivation) of attackers. This demonstrates that this is a two-party economic issue, with a great deal of diversity within each party (attackers and victims).

4 Results of Survey

To obtain some preliminary estimates of people's willingness to pay to recover their files we conducted a face-to-face survey using standard contingent valuation techniques (Alberini and Kahn 2009). The survey was performed at the University of Kent (UK) during a celebration day on campus. A total of 149 respondents took part (54% male, average age of 24). Note that because we conducted the survey on a celebration day the respondents were primarily alumni or residents of the city (and not students).

Half of those surveyed were asked the following two questions:

1. *Suppose that because of a mistake you made, you have lost access to all the files on your computer. The only way you can recover the files is to pay a private company who are experts in file recovery. What is the maximum amount you would pay to recover those files?*
2. *Suppose that your computer was infected by a virus which means you cannot access any of your files. The criminals responsible have been caught and you are now eligible for monetary compensation. How much money would you want to recompense you for the loss of files? Note that is your request is deemed too high the authorities will use a technique to recover your files and so you will receive your files but no compensation.*

Question 1 directly elicits willingness to pay (WTP) - how much is someone willing to pay to recover their files. We know, however, that people tend to understate willingness to pay. Question 2 addresses this problem by eliciting willingness to accept (WTA) - how much would someone need to be paid to compensate for the loss of files. In principle, WTP and WTA should be identical. Typically, however, one obtains a WTA-WTP disparity in that WTA is significantly higher than WTP (Horowitz and McConnell 2002). This difference can be anything from a factor of two to ten. The true valuation *vi* can reasonably be assumed to lie between an individual's stated WTA and WTP. There are arguments to suggest that the true valuation will be, in fact, closer to WTA than to WTP (Bateman et al. 2005).

As a robustness check, the other half of those surveyed were asked the following questions:

3. *Suppose that because of a mistake you made, you cannot access any of your files. You have an insurance policy that means you are eligible for monetary compensation. How much money would you want to recompense you for your loss of files? Note that if your request is too high a technique will be used to recover your files meaning you will received your files but no compensation.*
4. *Suppose that your computer was infected by a virus which means you cannot access any of your files. The only way you can recover your files is if you pay a fee to the criminals. If you can be certain that your files will be returned, what is the maximum you would pay to recovery your files?*

Note that question 3 elicits WTA while question 4 elicits WTP. Reversing the order of questions allows us to check that respondents stated WTA and WTP is not influenced by the order in which the questions are asked. Moreover, the framing varies between questions 1 and 4, and 2 and 3 in terms of who was responsible for the loss of files and who will be paid or recompensed. Ideally, we would want answers to questions 1 and 4 to be similar, and those to questions 2 and 3 to be similar. This was indeed the case and so, in the following, we shall solely report stated WTP (questions 1 and 4) and WTA (questions 2 and 3).

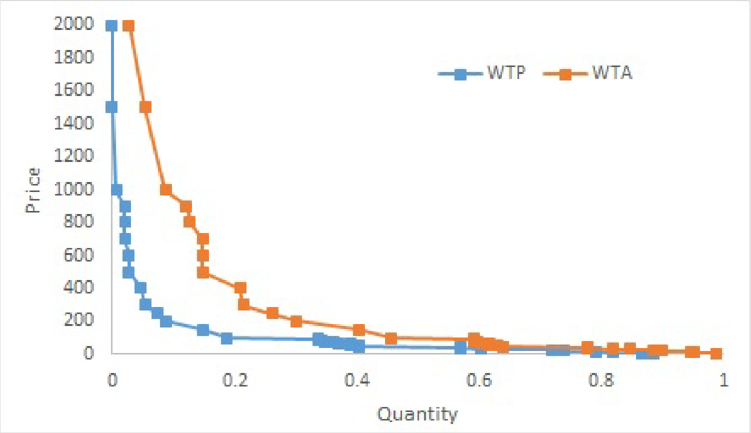


Figure 4 Demand curve elicited using WTA and WTP

In economics, it is conventional to plot the inverse demand curve Q^{-1} *(p)*. Figure 4 plots the inverse demand curve using the elicited values of WTA and WTP with total demand normalized to 1. Note that price is in sterling and at the time of the survey the exchange was around $1.40 per £1.00. As expected, the WTA (mean £547) exceeds the WTP (mean £276) resulting in higher demand when using the WTA as compared to the WTP. For instance, 20% of those surveyed were willing to accept at least £400 and pay up to £100. This is a WTA-WTP disparity of factor four.

As discussed above, elicited WTA is a better measure of true valuation. To calculate the profit maximizing ransom we fitted a (six degree) polynomial to the raw demand function elicited using WTA18. From this fitted demand function, we could then calculate marginal revenue (MR) using equation:

$$MR(Q) = \frac{dp(Q)}{dQ}Q + p(Q)$$

The optimal ransom demand is found where marginal revenue equals marginal cost. It seems reasonable to assume that the marginal cost to the criminals of dealing with an additional victim is near zero. The optimal ransom is, therefore, found by setting *MR(Q) = 0.*

Figure 5 plots the raw demand function (the same as in Figure 1), the fitted demand function and marginal revenue. Note that there are 5 values of *Q* for which *MR(Q) = 0*. Only one of these is the global optimum and it is simple to show that this is the smallest *Q* where *MR(Q)* = 0. This gives an optimal ransom of around £950. It is predicted that, in this setting, around 10% of victims will pay. An interpretation of this finding is that it is in the criminal’s interest to target high value victims that are willing to pay £1000 or above. For instance, at the optimum value the expected profit of the criminals is £99 per victim (because a little over 10% of victims will pay £950). If the ransom is dropped to, say, £150 then over 40% of victims will pay but this only results in a profit per victim of around £60.

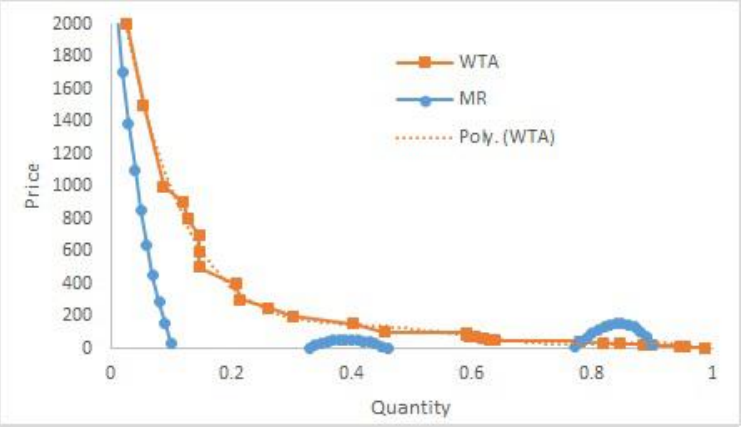


Figure 5 Demand curve elicited using WTA, and marginal revenue (MR)

Clearly, our survey results can be improved, for example with a larger and more representative sample, etc. So our findings are not definitive, and more evidence to support them is needed.

Our results do, however, demonstrate the potential for a method to calculate an optimal ransom. Moreover, we would not be surprised to find that the optimal ransom is much higher than what is currently seen in most ransomware. As our results illustrate, if there are a few people willing to pay a high ransom, then it may well be optimal to set a much higher ransom and accept that many victims will not pay up.

If it is the case that the optimal (uniform) price would result in many victims being unwilling to pay then it is especially in the criminal's interest to price discriminate. Ideally the criminals would want to identify those willing to pay a high ransom and ask a high ransom while making a lower ransom demand to others. This, though, requires being able to discern those willing to pay a high ransom. We find in our data that the willingness to pay is on average higher for women (mean £326) than men (mean £233) and increasing with age (correlation coefficient 0.07). These differences are, however, not statistically significant and so nothing more than suggestive. Again, however, they demonstrate the potential to estimate to what extent willingness to pay depends on discernible characteristics of an individual. This is key if the criminals are to effectively utilize price discrimination.

5 Conclusion

This report has presented an economic analysis of crypto-ransomware. The costs, pricing strategies and path to optimal pay-out has been discussed in detail referencing state-of-the-art literature. Ransomware is a growing, diversifying and highly profitable criminal enterprise. It is evident that it has not converged to an optimal form, and may not do so for a significant period. However, this doesn't mean it is not a high impact and costly issue in the present.

Ransomware has reached a technical level, at which it is no longer able to be reverse engineered to find the key for decrypting a victim's files. This puts a great deal of bargaining power in the hands of attackers who are best served by not tolerating any dissent: they are most likely to achieve a strategy-optimal pay-out if they are uncompromising. This has coupled with the need to cultivate a sense of trust (expectation that files will be returned), encouraging a business-like approach to serving the demand (manufactured though it may be) for the return of files for an agreeable price. Attackers are becoming more sophisticated not only in their use of crypto-ransomware, but in their analysis of potential victims and appropriate pricing strategies.

We have not, in the scope of this report, addressed pyramid-strategies, in which attackers allow victims to pay off their ransom by infecting other machines. This is because such attacks only postpone the ransom phase, effectively deferring the ransom phase to secondary and tertiary victims. Such attacks also imply that the attacker considers unpaid ransoms to be an operational cost so long as the victim spreads the infection further, which essentially means that the attacker is just paying opportunity cost to harness an ad-hoc and likely unreliable botnet, instead of using other methods previously discussed. This also raises the issue of victim empathy, wherein they choose not to perform any action (refuse the ransom entirely), as they are now directly aware that by giving in to either demand they promote further infections on a greater than 1:1 basis. As previously discussed, this is not economically optimal, as it has a diminishing effect on potential return as a greater body of potential victims becomes aware of the implications of their actions.

The costs associated with ransomware have led to the development of black markets that service less technically adept criminals. It is possible to purchase RaaS, botnets and other software and expertise to rapidly set up and deploy a profit-bearing ransomware enterprise. However, this has led to an abundance of uniformly priced ransoms. Uniform pricing represents the least effort approach to pricing, but its optimal pricing falls short of more sophisticated strategies. As a result, competition in this area is fierce, and prices may potentially drop as the factors limiting the expansion of ransomware (Bitcoin friction and willing-to-pay victims) begin to add pressure in an increasingly crowded economic space.

Price discrimination and cursory examination of one's potential victims increase the optimal value of ransoms, as does an uncompromising stance. More sophisticated malware can differentiate high value and low value targets and ask for appropriate ransoms, ensuring that high-rolling victims don't pay less than they are willing to. Herein lies the issue: incomplete knowledge of willingness to pay means that ransomware operations are an ongoing exercise in data collection, profiling and iterative ransom valuation. This indication of a currently unsophisticated criminal enterprise implies a worrying trend towards increasing sophistication as criminals become increasingly aware of how to maximise profit. The profit motive is effectively a self-subsidising movement towards increasing sophistication.

It is our finding that although increased sophistication on the part of attackers may lead to a larger value being attributed to ransomware as a for-profit enterprise, the welfare of individuals affected by ransomware may counter-intuitively rise. This is due to the increasing importance of trustworthiness on the part of attackers: as untrustworthy strains die off; trustworthy strains mean that an increase in ransom payment (and thus value of the enterprise) indicates a greater number of returned files.

This does not reduce the illegality of ransomware and the further criminal acts to which the proceeds may be put. Ransomware of this type is a growing, evolving and profit-bearing activity that is very likely to increase soon. With increasing numbers of privacy focused cryptocurrencies, it may also be possible for Bitcoin friction to be eased, allowing much more rapid growth of ransomware in both the short and long term. Convergence to optimal strategies and higher ransom values, either by design, or simple trial and error, is inevitable.

References

http://www.ramses2020.eu

Atanasova, S. (2016). The Shade Ransomware with New RAT Features to Determine Worthwhile Victims Virus Guides' Computer Security News of August 12, 2016. http://virusguides.com/shade-ransomware-new-rat-features-determine-worthwhile-victims/

Abrams, L. (2016). Fantom Ransomware derives Ransom Amount and Address from Filename Bleeping Computer News of September 21, 2016. https://www.bleepingcomputer.com/news/security/fantom-ransomware-derives-ransom-amount-and-addressfrom-filename/ ransomware-hit-you

Coase, R. H. (1972). Durability and monopoly. Journal of Law & Economics, 15: 143-149.

F-Secure (2017) The Customer Journey of Crypto-ransomware. https://fsecureconsumer.files.wordpress.com/2016/07/customer_journey_of_crypto-ransomware_f-secure.pdf

F-Secure (2017) State of Cyber Security in 2017. https://www.f-secure.com/documents/996508/1030743/cyber-security-report-2017

Gul, F., & Sonnenschein, H. (1988). On delay in bargaining with one-sided uncertainty. Econometrica, 56: 601-611.

Gul, F., Sonnenschein, H., & Wilson, R. (1986). Foundations of dynamic monopoly and the Coase conjecture. Journal of Economic Theory, 39: 155-190.

Hernandez-Castro, J., Cartwright, E. and Stepanova, A., 2017. Economic Analysis of Ransomware.

Hernandez-Castro, J., Cartwright, E. and Stepanova, A., 2017. Ransomware and Game Theoretic Insights on Kidnapping.

Lapan, H. E., & Sandler, T. (1993). Terrorism and signalling. European Journal of Political Economy, 9(3), 383-397.

Lapan, H. E., & Sandler, T. (1988). To bargain or not to bargain: That is the question. The American Economic Review, 78(2), 16-21.

Muthoo, A. (1999). Bargaining theory with applications. Cambridge University Press.

Parker, L. (2016). Large UK businesses are holding bitcoin to pay ransoms. Bravenewcoin.com News, 9 June 2016. http://bravenewcoin.com/news/large-uk-businesses-holding-bitcoin-to-pay-ransoms/

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

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. Econometrica, 50: 97-109. Selten, R. (1977). A simple game model of kidnapping. Lecture Notes in Economics and Mathematical Systems 141: pp 139-155.

Sandler, T. (2003). Terrorism & game theory. Simulation & Gaming, 34(3), 319-337.

Sandler, T., & Arce, D. G. (2007). Terrorism: a game-theoretic approach. Handbook of Defence economics, 2, 775-813.

Selten, R. (1988). A simple game model of kidnapping. In Models of strategic rationality (pp. 77-93). Springer Netherlands.

Varian, H. R. (1989). Price discrimination. Handbook of industrial organization Volume 1, R. Schmalensee and R. Willig (Eds.) Elsevier: North Holland, pp. 597-654.

Young, A., & Yung, M. (1996). Cryptovirology: Extortion-based security threats and countermeasures. Security and Privacy Proceedings IEEE Symposium. IEEE.