Assignment 1 - Economics of cyber security

An analysis of the darknet market The Real Deal

Group 6
Vishan Baldew - 4180992
Martijn Cligge - 4152220
Stephan Kool - 4151895
Christian Veenman - 4495705

1. Introduction

With the introduction of The Onion Router (TOR) users are now able to create a peer-based encryption layer allowing for anonymous navigation and hosting of any website. TOR and related anonymization networks resulted in the creation of the darknet. Illegal marketplaces on the darknet, commonly referred to as darknet markets or cryptomarkets, have seen a rise in popularity (Owenson 2015). The first of its kind, Silk Road, gained attention in popular media and has been shut down in 2013. Current active darknet markets such as The Real Deal and AlphaBay are used to offer products and services, ranging from drugs and child pornography. Lately these sites are also offering cybercrime products that are focused on the IT area. Examples of such products are spam, source code files, worms or zero day-exploits. In this report we call the IT related products as cybercrime assets.

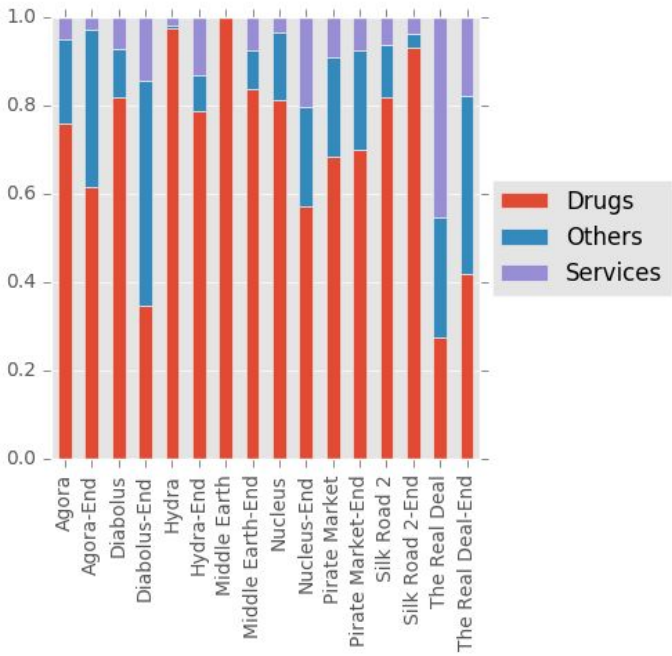

Figure 1 Type of products in darknet markets

In this research we will focus on the security issues (with a focus on cybercrime assets) that have arised with the introduction of the darknet and we provide some metrics that can show the level of security viewed from different actors perspectives. We will do this by analyzing the actor field, the ideal metrics for security decision makers, the metrics that are already being used in practice and by analyzing a dataset. This dataset contains data about the darknet markets. For this analysis, we will focus only on the Real Deal market, since this market has the highest amount of IT related services (see figure 1).

1.1. Security issues that have arised with the introduction of darknets

Zero-day exploits and other similar cybercrime assets have been available for a long time but the accessibility of such assets was low since they were only accessible to a specific group of cyber criminals which were part of a cybercriminal network. With the introduction of darknet websites these products can be bought by anybody without having to connect with cybercriminal networks. This is why the availability and exchange of various cyber tools on

darknet websites form a significant security issue for private companies, but also for governmental institutes who make use of the software. This is an increased security because of a: (McGraw 2004)

- Higher availability of cybercrime assets - With the introduction of the darknet, sellers of cyber security exploits have now a platform to sell cybercrime assets to anybody.
- Higher accessibility of cybercrime assets - Anybody can just download TOR and go to a darknet website to buy a cybercrime asset.
- Anonymity to buy cybercrime assets - TOR provides anonymity both for sellers and buyers of cyber security exploits. You can pay anonymously with Bitcoins because no bank is involved.
- Easiness to pay for cybercrime assets - It is relatively easy to pay with Bitcoins and bitcoins obfuscates the identify of the buyer and seller.

2. The actors involved and their ideal metrics

The security issue as presented above involves a number of actors. We have limited the scope to the operations within the crypto market, an actor attacking the crypto market (the police task unit) and a victim of the operations on the crypto market (a software company). The objectives of the security decision makers within this scope are analyzed and visualized through an objectives-tree, which is a modeling technique for identifying metrics (De Haan et al., 2009).

2.1 Actors

As indicated previously, the actor field is scoped to the operations on the crypto market, an attacker and a victim. Inside the crypto market, we assume three roles: the sellers of goods and services on the market, the buyers of goods and services on the market and the operator of the crypto market, which facilitates the trade between the buyer and the seller. Furthermore, we assume the involvement of a specialized police task unit that assumes an attacking role and a generic software company that is the victim of cyber security exploits being sold on the market. This field of actors is illustrated in the formal chart below.

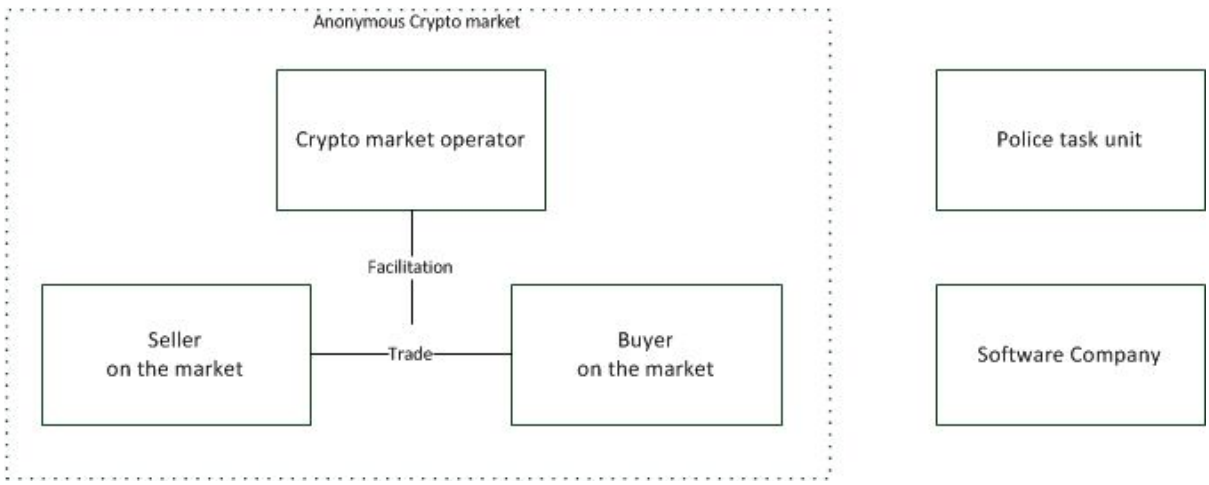

Figure 2 Overview of the involved actors

2.1.1 Security decision makers

In this actor field, we identify the crypto market operator, the police task unit and the software company as security decision makers relevant to the security issue of an available, accessible, anonymous and easy-to-trade crypto market. Security decision maker in this context is similar to the concept of stakeholder. As such, the crypto market operator is a stakeholder because this actor is creating and improving the market that creates the security issue. The police task unit is a security decision maker in an offensive sense: attacking the crypto market with the aim of taking it down. This actor also has to allocate resources for security police, but in an offensive manner. Finally, there is the software company that is victimized. The company needs to allocate resources on defensive actions against the exploits available. Actions within the crypto market directly affect the choices of the software company.

2.2 What would be the ideal metrics for security decision makers?

To answer this question, we use the objectives tree modeling technique from De Haan et al. (2009). It puts the objective of the actor on top and consequently being broken into its different aspects, which are consequently operationalized into measurable metrics. Some of these metrics have been simplified because of the scope of this assignment. For example, the anonymity of the web is dependent on a very large number of factors, both technical and social. This requires a huge simplification to the point where it is simplistic as a stand-alone aspect. Yet on a system level, the operationalization of this aspect the the number of buyers and sellers caught annually gives provides sufficient simplification for analysis and at the same time sufficient complexity to be able to draw meaningful conclusion.

2.3 Objective tree crypto market operator

The objective of the crypto market operator is modeled as constantly creating a better crypto market. We have divided these into four aspects, namely more profit from the crypto market, better availability, better accessibility and better anonymity of the crypto market (see figure 3).

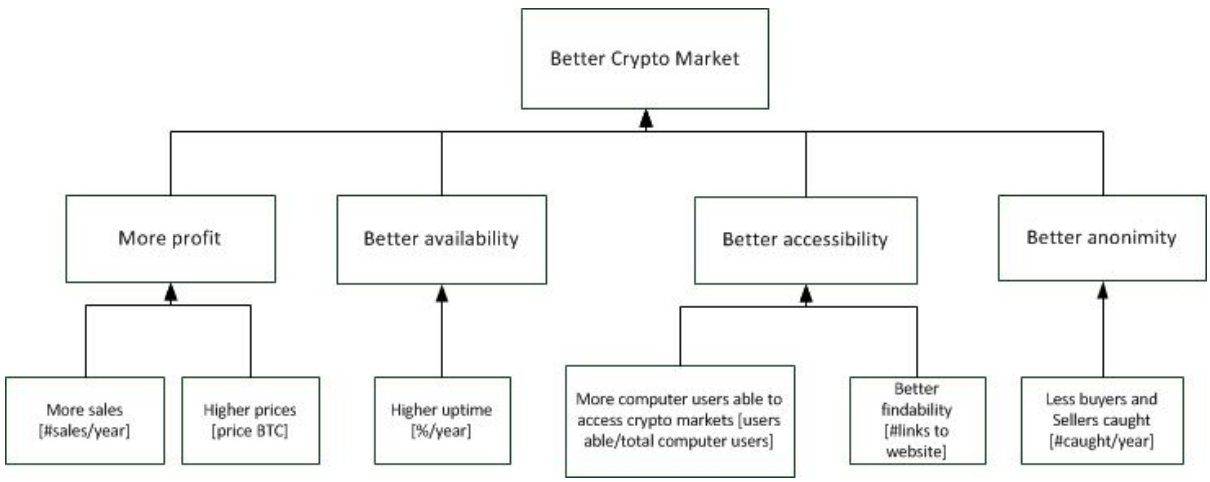

Figure 3 Objective tree crypto market operator.

The respective metrics are more sales, higher prices, higher uptime, more users able to access the market, better findability of the market and less buyers and sellers being caught. The findability metric has been expressed in links as the TOR websites do not have a public search engine such as Google with the normal web does have. Findability in the TOR-web is done similarly as before Google Search: websites linking to each other. Having more links to the website therefore results in being found more easily.

2.4 Objective tree police task unit

The objective of the police task unit is modeled as aiming for having less illegal trade through crypto markets. This objective is then split into reducing the ease of paying on a market to harm the trade, the lower availability of the market, the lower accessibility of the market and the reduced anonymity on the market.

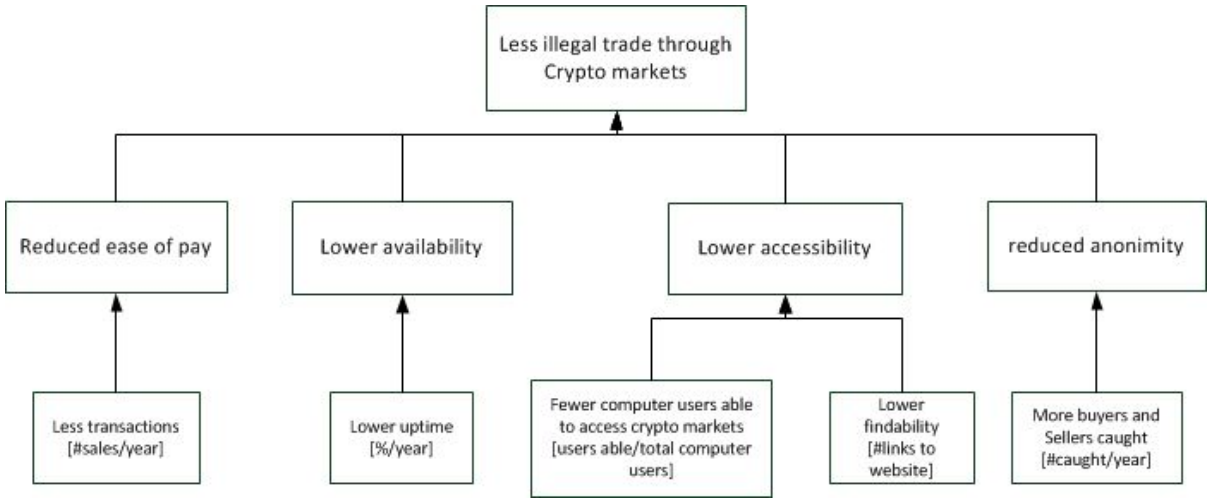

Figure 4 Objective tree police task unit

The respective metrics to these aspects are: less transactions, lower uptime, fewer users able to access the market, lower findability of the market and more buyers and sellers caught (see figure 4).

2.5 Objective tree software company

The software company is the victim of the activities on the crypto market and in this analysis the company will be scoped to this purpose. It is debatable whether better software or profits are the ultimate goal of a software company, but for this specific scope we have identified better software as the ultimate goal. Profit is made an aspect of better software, as the company needs to be profitable to be able to develop better software. The other aspect is better security, which has taken inspiration from the FAIR-framework (Jones 2004) as in that it has three aspects: vulnerability, loss magnitude and attacker strength. The attacker strength is modeled as the attractiveness of an attack on the software of the company.

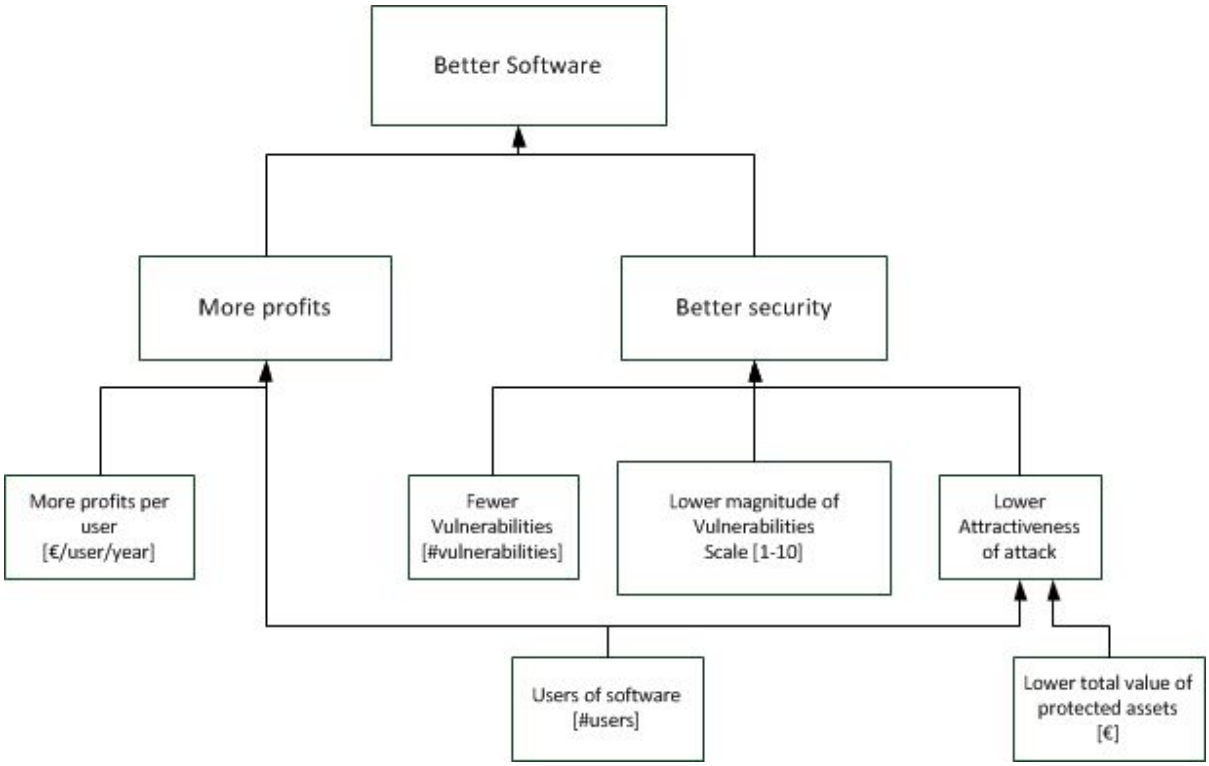

Figure 5 Objective tree software company

The metrics of the software company are: more profits per user, fewer vulnerabilities, lower magnitude of vulnerabilities, lower total value of protected assets and a trade off between both more and fewer users (see figure 5). The more users the company has, the more profits it will make (generally speaking). Yet having a high number of users also increases the attractiveness of an attack, lowering security.

2.6 List of metrics

To summarise, the ideal metrics for the three security decision makers are:

- Number of sales on Crypto markets per year [# / year]
- Average price of products on crypto market [EUR]
- Uptime in of crypto market [% per year]
- Computer users able to access crypto market [% users / total users]
- Number of links to crypto market [#]
- Number of sellers and buyers identified per year [#/year]
- Profit of software company per user per year [EUR/user/year]
- Vulnerability of software [#current vulnerabilities]
- Average magnitude of Vulnerabilities, scale [1-10]
- Users of software [#users]
- Total value of protected assets [EUR]
- (Bitcoin) transactions in crypto markets per year [#/year] (this is equal to the first metric)

3. Metrics in Practice

In practise it is often hard to measure the ideal metrics because the data and measurements needed for those metrics are often not available. However several data sources exist that

contain different metrics and measurements that can be used in practise to relate the vulnerabilities being available to the risk they cause.

3.1 Calculating Ideal Metrics in Practise

The ideal metrics that were defined in the previous section are very difficult to calculate, because data is lacking. Some of the ideal metrics that were defined in the previous section can be calculated while some others cannot:

- **The number of sales on a crypto market per year.** Even though we can scrape darknet markets and see which products appear and disappear we cannot say with certainty if those products have actually been sold or where just simply removed by the seller because of other reasons. However the darknet market operator might have access to this information depending on if the market stores information about products being sold.
- **The average price of products on the crypto market in Euro's** can be calculated by simply scraping all products from the website and calculating the average price by parsing the price of each product.
- **Uptime of the crypto market in percentages per year**. This is possible by creating an automatic system that sends HTTP GET requests to see if the website is still up. By comparing the amount of replies on which the website was down with the amount of replies of the website being up we can calculate the amount of uptime of the crypto market.
- **Computer users able to access a crypto market.** This is hard to calculate because it requires us to determine how many computer users there are and how many users have access to software such as TOR and other anonymization networks.
- **Number of links to crypto market**. Because the darknet is generally unindexed (meaning that darknet search machines cannot find them) it is impossible to scan all the websites to find all the links that would possibly lead to the crypto market. It is possible to scrape indexed websites or scan recursively into the darknet for these links but it requires a lot of effort and does not guarantee to find all the links that lead to the darknet market.
- **Number of sellers and buyers identified per year.** This information is generally only accessible to the police who tries to identify the buyers and sellers of a darknet market. Other actors in the system can only depend on public reports of the police or make guesses on the amount of accounts being abandoned, but the conclusions you can draw from this are very subjective.
- **Profit of a software company per user per year.** This is only known to the software company and the police (if they have sufficient reasons to obtain this knowledge) unless it is a public company and this information is released in a public report. The other actors could calculate this by dividing the total profit of the company by the amount of users of the product but this might not be perfect because a company might also earn money unrelated to the user of that piece of software (like other software being produced by that company).
- **The amount of vulnerabilities of software**. Software vulnerabilities are not always discovered and therefore it is impossible to measure the amount of vulnerabilities in software - some vulnerabilities are simply not yet discovered. Doing penetration tests

on the system might discover more vulnerabilities but does still not guarantee that there are not more vulnerabilities in the system.

- **Average magnitude of the vulnerabilities**. Because not all the vulnerabilities can be found it is hard to calculate this too. Another question that might be asked is how magnitude is really defined. This often differs per system. For a medical system probably user privacy is the most important aspect, whereas other companies might want to protect their availability the most.
- **Users of software**. This is generally only known to software company but might also be made public by the software company or by other means. For example the market share of web browsers being used can easily be looked up on google.
- **Total value of protected assets in Euro's.** The total value of protected assets in Euro's is something that differs for each actor. The software company assesses this for it's own system while the darknet market operator will estimate this for his market. Because some assets are not easily expressed in terms of money (for example information) this is often a hard to measure/calculate.
- **Bitcoin transactions in the crypto market per year.** The market operator can use this metric to see how successful the market is assuming that the market system tracks which sales have been completed.

3.2 Data sources for calculating metrics

There are several data sources available from which information can be obtained to calculate the ideal metrics as defined in chapter 2. These data sources do not always cover everything but do often help to estimate a given metric.

Netcraft for example contains several metrics about the world wide web from over the years[1] . This includes information about users switching Internet Service Providers but also about the webserver share of the public web. This information could potentially be used to assess the risk when a vulnerability has been made available on a darknet market. Vulnerabilities in web server software that is more often used is likely to increase the frequency of attacks because of the high amount of targets that exist. These information sources are in particular interesting for buyers and sellers of these type of vulnerabilities because it helps estimating the price and magnitude of the exploit being sold. Software companies can also reasonably conclude from this that the higher the amount of users is, the higher the likelihood is that they will be attacked.

Another data source which can help in assessing the risk of a system is given by the TOR Project[2]. The TOR Project exposes many metrics related to the TOR network. For example the amount of TOR Relays over the years provides the darknet market operator with the current security state of the TOR network on which it operates. The more relays being used in the network, the harder it is to track the traffic by the police and governmental institutes. An overview of the number of relays and bridges of 2014-2016 is shown in figure 6.

[1] https://news.netcraft.com/archives/category/web-server-survey/
[2] https://metrics.torproject.org

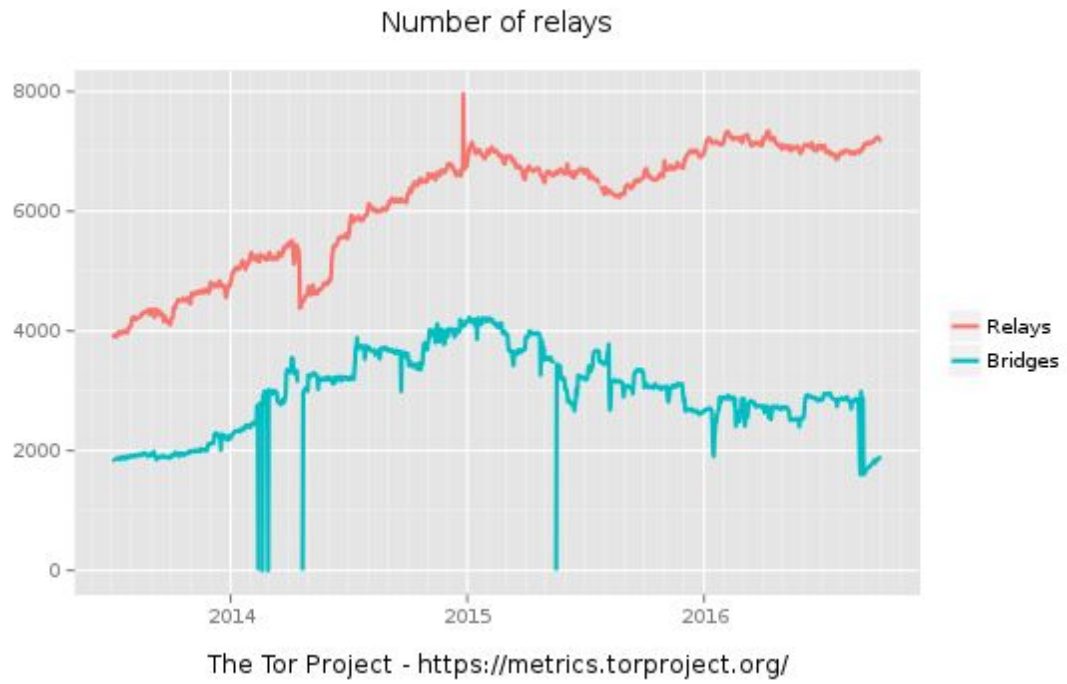

Figure 6 Number of Relays and bridges

Other metrics include the number of unique onion addresses in use (as shown the public relays) and the amount of TOR-users compared to normal internet users, which can be seen in figure 7.

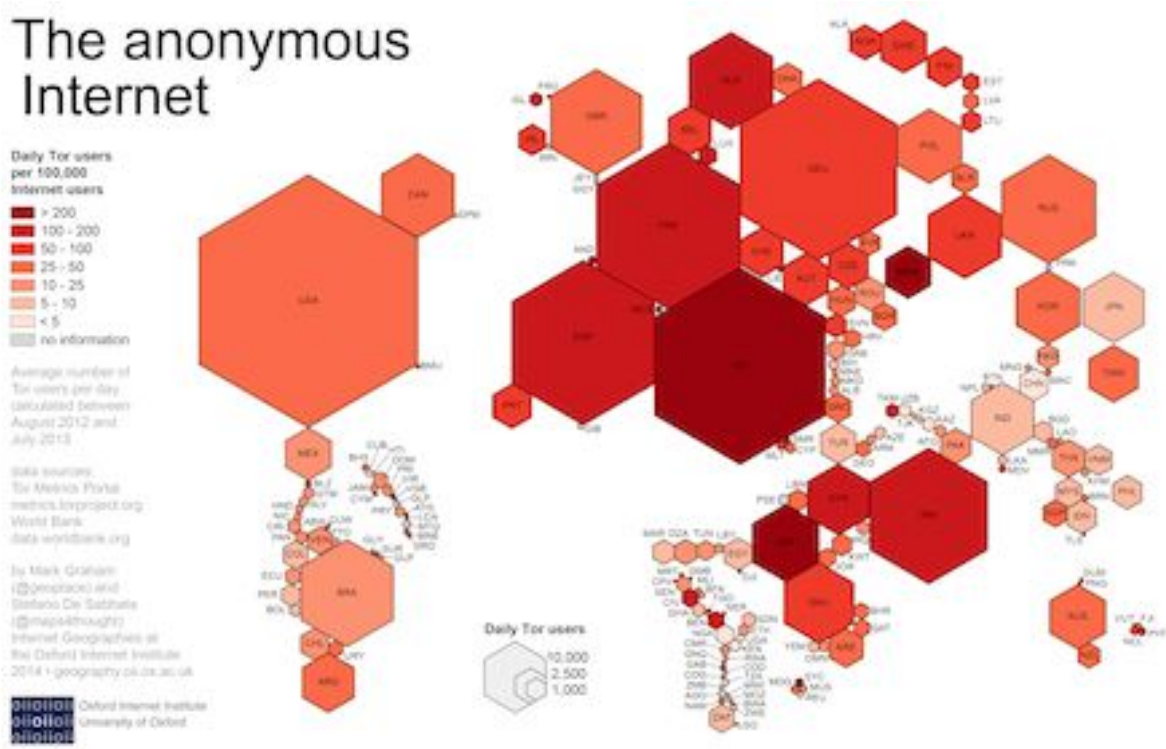

Figure 7 Number of TOR users in the world.

4. Metrics from the Real Deal market dataset.

4.1 The dataset

As described earlier, the dataset contains data from the marketplace on darknet website The Real Deal. The total dataset contains 7980 records. Since we mainly focus on the cybercrime assets we therefore leave other assets like drugs and child pornography out. This resulted in reduced size of 917 records. The dataset contains data about the seller, a description of the offered product, the price and some more details.Between a certain time interval this data was scraped from the web page. There are 917 records with the following columns: Data (date of the scraping), Description (description of the cybercrime asset), URL, Review (no data could be retrieved from this column), Price (in Bitcoin), User (seller of the cybercrime asset) and Producttype (Type of cybercrime asset, e.g. spam or source code). The analysis was done with the software package Rstudio.

4.2 The metrics

In addition to the ideal metrics, which are described in section 2, and the metrics from practice, which are described in chapter 3, we have defined metrics that are derived solely from the dataset. The following metrics are defined:

1. **Number of different cybercrime assets.**

Different type of cybercrime assets are being sold on the Real Deal Market. The main types are Day Private Exploits , Source Code, 0-Day exploits, Accounts, Information, RATs, Databases, Cards, Hardware, Spam. The number of exploits metric shows how many different cybercrime assets are being sold on the darknet website. This can be useful for the actor task *police task unit* and the *software companies.* This provides them an overview of which vulnerabilities cyber criminals sell most often and tell software companies and police task units on whether to invest in counter measurements on certain threats.

2. **Average price per cybercrime asset ($).**

Every cybercrime asset is being sold against different price ranges. This metric provides an overview of the average price of the most expensive cybercrime assets on the Real Deal Market.

3. **Supply per seller.**

This metric shows the seller which sells the most cybercrime assets on the Real Deal Market. This metric might be relevant for the police task unit, to see which criminal is the most active on the platform, and to whom the *police task unit* might need to undertake some action. We need to mention that this is an absolute number of sold items, so there is no correction for duplicates.

4. **Number of cybercrime assets per software company.**

There are different type of cybercrime assets available per software company. This metric providers software companies an overview of how many cybercrime assets there are available for their OS or software package. This allows a specific software company to compare themselves against other software companies. Based on this analysis it might be necessary for a specific company to invest in more counter measurements against cybercrime if they rank high in the number of exploits available. We have to note that in this

metric, there is no correction for the type of cybercrime asset (as discussed in metric 1), so this this the cumulative amount of cybercrime assets available per software company. Future work could focus on the specific number of different type of cybercrime assets per company.

5. Evaluation of the metrics

The data is the Real Deal dataset was fairly limited. To recap: the data contained a category of the product, a description of the product, the link to the product, a name of the seller, an option for reviews and the price of the product in BTC. The data has the following issues:

Firstly, the data contains several snapshots from different days of the same products, creating several duplicates of the same product in the data. For these products, the description and the static URL remained the same. This however, strongly reduces the capability of testing any of the proposed metrics on the data because their is not enough statistical significance.

Secondly, some of the attributes are redundant. For example, the description and the link contain the same keywords because the link is based on a 'subject'- description of the product.

And finally, much of the information is ultimately useless. Starting with the review and ratings: these are rarely used - so rarely that a review is a statistical outlier. There is no information about who bought a product or the number of transactions. There is the name of the seller but this is always a nickname chosen to protect identity and there is no limit to the number of accounts a single person can have - so theoretically all the sellers could be the same person. The last remaining attribute is pricing - which is also problematic to extract information from. The selling of exploits has a monopolistic character with very little offer and extreme specificity of the product, the price elasticity is very low. So the price is neither reflective of the value of the exploit nor the result of a classical offer- demand relation curve.

5.1 Reflecting on the metrics form the dataset

1. **Number of different cybercrime assets**

Figure 8 shows the frequency of the different type of cybercrime assets.

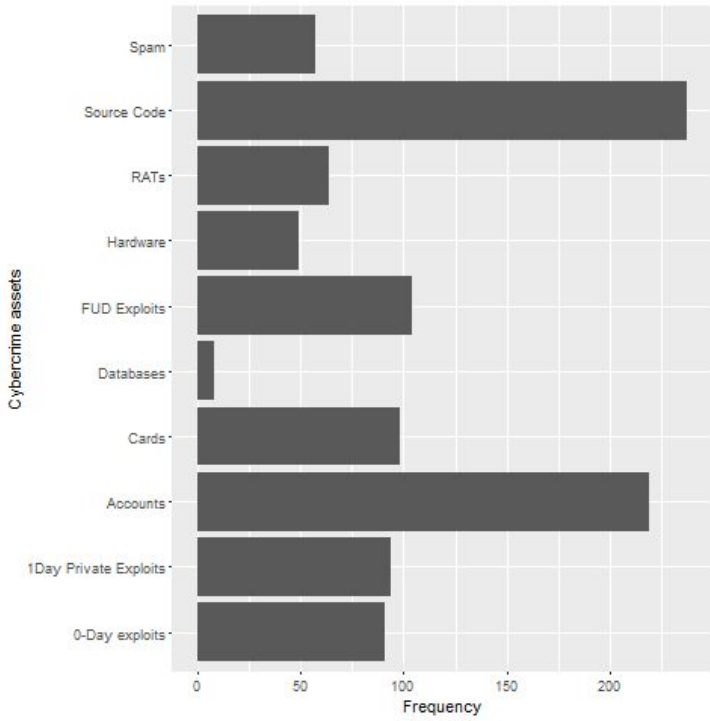

Figure 8 Number of cybercrime assets.

We can see that the most offered products contain either source code files or account information. This tells software companies that they might need to focus more on the

security of their users' account information. Also, they might need to look into more specific software vulnerabilities, which can be exploited by the source code files. The usability of this metric is fine, although it only provides some high level overview of the available products. It would be a nice addition to see some more details. Unfortunately, this bar chart does not show what vulnerabilities are in danger when the source code products are used. This is something this metric cannot show.

2. **Average price per cybercrime asset ($)**

Figure 9 displays an overview of the average price of Cybercrime assets. From this figure can be inferred that 1-Day exploits have the highest average price, followed by Hardware, Databases, and 0-day exploits. Other assets, such as Accounts, Cards and Spam are relatively cheap. This is interesting, because based on the fact that figure X tells us that there are e.g. a large number of accounts available, this might mean that there is (much) more supply than demand, hence the relatively low price. This metric is especially usable for software companies and the police task unit, because it might give an indication of the popularity or criticality of the cybercrime asset.

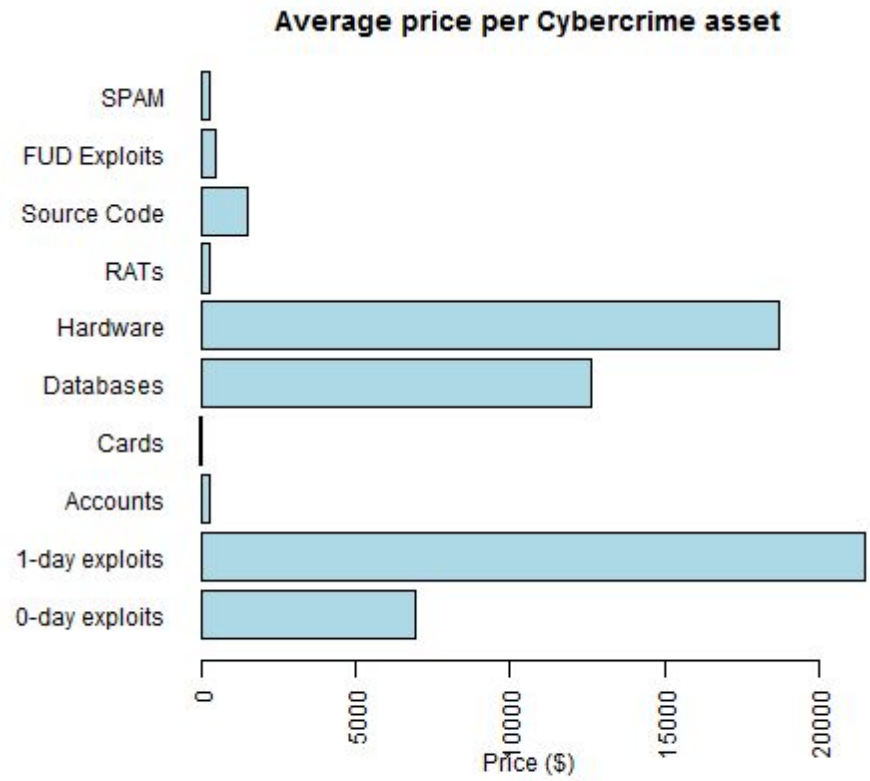

Figure 9 Average price per cybercrime asset

.

3. **Supply per seller**

In table 1 is shown which displays who sells the most cybercrime assets. This can be used by police task units to identify certain users they might want to trace down. The table shows us that there are only three people who sell the most cybercrime assets.

Username	**# cybercrime assets**
bestbuy	480

KnowHowForAll	169
Logger	119
mump	41
roundybit888	26
PlayBit	21
osquare	20
fukwits	19
ExploitFun	15
MrMorgan	13
ForeverPP	11

Table 1 The number of cybercrime assets per seller

4. Number of cybercrime assets per software company

This bart chart, shown in figure 10, the cybercrime assets per company are shown. Four software companies have the most cybercrime have the most cybercrime assets (Microsoft, Wordpress, Google and Adobe). It is easy to conclude that these companies have to take counter measurements to improve their security level, because this metric does not take the userbase into account. One could make the assumption that the more users a platform has, the more cybercrime assets there are provided for this platform. However, in general Apple software is conceived as safe software, so this might be a conclusion that is somewhat correct. We have to mention again, that in the construction of this metric we did not correct for duplicates because of time constraints.

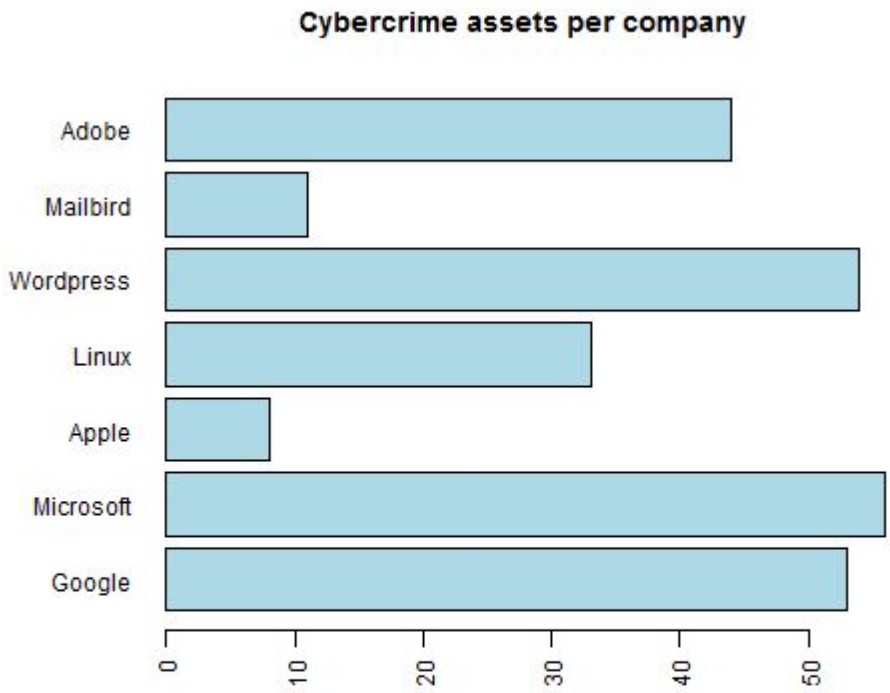

Figure 9 Number of cybercrime assets per software company.

6. Conclusion

This research has shown that there is still a gap between ideal metrics and the metrics in practice, since not all the metrics are easy to calculate. To come up with new metrics based on the dataset was difficult since the dataset was limited for our needs. Nonetheless, four additional metrics that might be convenient for the described stakeholders (Police task unit, software company and the operator) could be derived from the dataset. These metrics include *Number of cybercrime assets per software company*, *Supply per seller*, *Average price per cybercrime asset ($)* and *Number of different cybercrime assets*. A more detailed dataset might be suitable to find more metrics, and due to our time constraint the data analysis part was a little superficial. Also, we only analyzed one market. It might be interesting to compare other markets as well. These limitations might need some improvements, this can be future work.

References

De Haan, A., W. P. A. C. Willemse, P. de Heer, S. C. Vos, and P. W. G. Bots. "Inleiding technische bestuurskunde." (2009).

Jones, J. (2004). *U.S. Patent Application No. 10/912,863*.

McGraw, G. (2004). Software security. IEEE Security & Privacy, 2(2), 80-83.

Owenson, G. H., & Savage, N. J. (2015). The tor dark net. *Global Commission on Internet Governance*, (20).