CyBOK

ADVERSARIAL BEHAVIOURS KNOWLEDGE AREA

(DRAFT FOR COMMENT)

AUTHOR: Gianluca Stringhini – Boston University

EDITOR: Awais Rashid – University of Bristol

REVIEWERS:

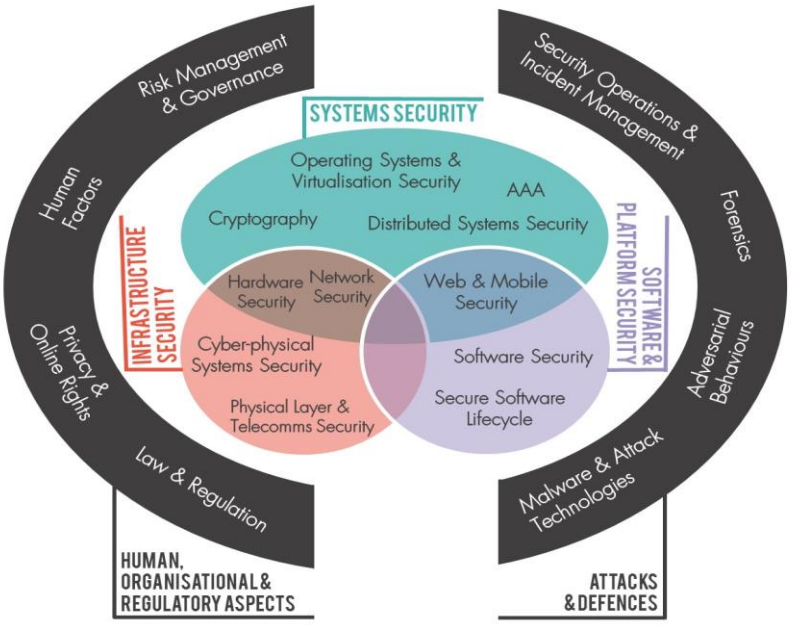
Adam Joinson – University of Bath

Damon McCoy – New York University

Paul Taylor – Lancaster University

Tom Holt – Michigan State University

Following wide community consultation with both academia and industry, 19 Knowledge Areas (KAs) have been identified to form the scope of the CyBOK (see diagram below). The Scope document provides an overview of these top-level KAs and the sub-topics that should be covered under each and can be found on the project website: https://www.cybok.org/.



We are seeking comments within the scope of the individual KA; readers should note that important related subjects such as risk or human factors have their own knowledge areas.

It should be noted that a fully-collated CyBOK document which includes issue 1.0 of all 19 Knowledge Areas is anticipated to be released by the end of July 2019. This will likely include updated page layout and formatting of the individual Knowledge Areas.

Adversarial Behaviours KA

Gianluca Stringhini

April 2019

INTRODUCTION

The technological advancements witnessed by our society in recent decades have brought improvements in our quality of life, but they have also created a number of opportunities for attackers to cause harm. Before the Internet revolution, most crime and malicious activity generally required a victim and a perpetrator to come into physical contact, and this limited the reach that malicious parties had. Technology has removed the need for physical contact to perform many types of crime, and now attackers can reach victims anywhere in the world, as long as they are connected to the Internet. This has revolutionised the characteristics of crime and warfare, allowing operations that would not have been possible before.

In this document, we provide an overview of the malicious operations that are happening on the Internet today. We first provide a taxonomy of malicious activities based on the attacker's motivations and capabilities, and then move on to the technological and human elements that adversaries require to run a successful operation. We then discuss a number of frameworks that have been proposed to model malicious operations. Since adversarial behaviours are not a purely technical topic, we draw from research in a number of fields (computer science, criminology, war studies). While doing this, we discuss how these frameworks can be used by researchers and practitioners to develop effective mitigations against malicious online operations.

CONTENT

1 A Characterisation of Adversaries

In this section, we present a characterisation of adversaries who perform malicious actions. This characterisation is based on their motivation (e.g., financial, political etc.). Although alternative characterisations and taxonomies exist (e.g., from the field of psychology [1]), we feel that the one presented here works best to illustrate known attackers' capabilities and the tools that are needed to set up a successful malicious operation, such as a financial malware enterprise. This characterisation also follows the evolution that cybercrime has followed in recent decades, from an ad-hoc operation carried out by a single offender to a commoditised ecosystem where various specialised actors operate together in an organised fashion [2, 3]. The characterisation presented in this section is driven by case studies and prominent examples covered in the research literature, and as such is not meant to be complete. For example, we do not focus on accidental offenders (e.g., inadvertent insider threats). However, we believe that the set of crimes and malicious activities presented is comprehensive enough to draw a representative picture of the adversarial behaviours that are occurring in the wild at the time of writing. We begin by defining two types of cyber offences as they have been defined in the literature, cyber-enabled and cyber-dependent crimes, and we continue by presenting different types of malicious activities that have been covered by researchers.

Cyber-enabled and cyber-dependent crimes

One of the main effects that the Internet has had on malicious activity has been to increase the reach of existing crimes, in terms of the ease of reaching victims, effectively removing the need for physical proximity between the victim and the offender. In the literature, these crimes are often referred to as *cyber-enabled* [4].

According to Clough [5], criminals have five main incentives to move their operations online:

1. Using the Internet, it is easier to find and contact victims. Email lists are sold on underground markets [6], while online social networks have search functionalities embedded in them, allowing criminals to easily identify potential victims [7, 8].
2. By using the Internet, criminal operations can be run more cheaply. Sending emails is free, while scammers previously had to pay postage to reach their victims. This also allows criminals to increase the scale of their operations to sizes that were previously unthinkable.
3. Compared to their physical counterparts, the Internet allows crimes to be performed faster. For example, emails can reach victims in a matter of seconds, without having to wait for physical letters to be delivered.
4. Using the Internet, it is easier to operate across international boundaries, reaching victims located in other countries. In this setting, often the only limitation is language, with criminals only targeting victims who speak a language that they are familiar with (e.g., people in English-speaking countries) [9].
5. By operating over the Internet, it is more difficult for criminals to get caught. This is mainly due to the transnational nature of cybercrime, and the fact that the problem of harmonising the appropriate laws of different countries is far from being solved [10]. In addition, research shows that online crime is often under reported, both because victims do not know whom to report it to (given that the offender might be located in another country), as well as the fact that they believe that they are unlikely to get their money back [11].

Cyber-dependent crimes, on the other hand, are crimes that can only be committed with the use of computers or technology devices [4]. Although the final goal of this type of crime often has parallels in the physical world (e.g., extortion, identity theft, financial fraud), the Internet and technology generally enable criminals to give a new shape to these crimes, making them large-scale organised endeavours able to reach hundreds of thousands, if not millions, of victims.

In the rest of this section we analyse a number of cyber-enabled and cyber-dependent criminal schemes in detail.

Interpersonal offenders

The first category that we are going to analyse is that of *interpersonal crimes*. These crimes include targeted violence and harassment, directed at either close connections (e.g., family members) or strangers. While these crimes have always existed, the Internet has made the reach of harassers and criminals much longer, effectively removing the need for physical contact for the offence to be committed. As such, these crimes fall into the cyber-enabled category. In the rest of this section, we provide an overview of these adversarial behaviours.

Cyberbullying. Willard [12] defines cyberbullying as 'sending or posting harmful material or engaging in other forms of social aggression using the Internet or other digital technologies'. While not always illegal, cyberbullying often occupies a grey area between what is considered a harmful act and a criminal offence [13]. This practice has become a serious problem for young people, who are often

targeted by their peers not only in real life, but also on online platforms [14]. While the practice of bullying is nothing new, the Internet has changed the dynamics of these harassment practices significantly. What used to be a harmful practice limited to school hours now can be perpetrated at any time, effectively exposing victims to non-stop harassment [15].

One aspect that makes cyberbullying different from traditional, physical harassment is that people online can be anonymous, and do not have their name or face attached to the abusive activity that they are carrying out [16, 17]. Researchers found that interacting with people online creates a *disinhibition* effect wherby personal traits are accentuated (i.e., negative people become meaner and positive people become nicer) [18]. This disinhibition effect can have the effect of making some people more likely to engage in abusive activity than they would do in the offline world [17]. Another aspect that contributes to disinhibition is the fact that online content distributed on certain platforms (e.g., snapchat, 4chan) is ephemeral, in the sense that it is deleted after a certain period of time [19]. As such, harassers feel that their actions have no adverse consequences since there will be no hard evidence of it in the future.

Doxing. Another type of online harassment is the practice of *doxing*, an attack where the victim's private information is publicly released online [20]. This operation is usually part of a larger harassment campaign, where the release of sensitive information is used as a way of embarrassing the victim or facilitating further harassment, even in the physical world (for example, by releasing information at the workplace or the home address of the victim).

The practice of doxing has become increasingly popular in recent years as a way of polarising online discussion and silencing people. A prominent example is the `#GamerGate` controversy, where women activists were often attacked and had their personal information posted online [21]. Doxing has been a primary vehicles for coordinated hate attacks run by polarised online communities such as 4chan's Politically Incorrect board (/pol/) [19]. As part of these attacks, anonymous users post information about their targets online (e.g., social media pages, phone numbers, physical addresses), and then invite other people to carry out loosely coordinated attacks (called *raids*) against those people. These attacks usually consist of hate speech and other abusive language.

While prominent in the online harassment space, the practice of doxing is also used by other offenders. For example, it is one of the techniques used by hacktivist groups such as Anonymous to put their targets on notice. We will discuss the other techniques used by hacktivists, together with their motivations, later in this section.

Cyberstalking. Another harmful activity that has been facilitated by the Internet is stalking. Cyberstalking is the practice of using electronic means to stalk another person [22, 23]. Broadly speaking, we can identify two types of cyberstalkers: those who use the information that they find online to help them stalk their victim in real life (e.g., monitoring social media to know their whereabouts), and those who use the means offered by online services to stalk their victim purely online. Further, the stalkers who operate online are divided into those who act purely passively, without any interaction with the victim, and those who perform interactions, for example, by sending her messages on a social network platform [24]. To counter cyberstalking, legislation has recently been introduced in many countries, including the 2012 Protections of Freedoms act in the UK and the 2000 Violence Against Women Act in the US.

Sextortion. An emerging crime that has risen to relevance is *sextortion*, where a criminal lures victims to perform sexual acts in front of a camera (e.g., a webcam in a chatroom), records those acts, and later asks for a monetary payment in order not to release the footage [25]. Sextortion is becoming such a relevant threat that crime prevention agencies such as the National Crime Agency (NCA) in the UK are launching dedicated awareness campaigns against it.[1]

Child predation. Another crime that is facilitated by the Internet is child predation [26]. Online

[1] http://www.nationalcrimeagency.gov.uk/crime-threats/kidnap-and-extortion/sextortion

services are a fertile ground for criminals to find victims, whether on chats, online social networks, or online gaming platforms. Compared to the corresponding offline offence, online sexual predation has two main differences: first, the victim and the perpetrator almost never know each other in real life. Second, the victim demographics are more skewed towards adolescents than young children, because the age at which kids start going online is slightly higher [27].

Cyber-enabled organized criminals

In this section, we focus on cyber-enabled crimes that are carried out by career criminals. In particular, we provide two prominent examples of cyber-enabled crimes that have received significant attention by the research community: *advance fee fraud* and *drug dealing*. These crimes are not usually carried out by single offenders, but rather by multiple criminals who act together in small organisations [28] or in actual structured criminal organisations [29]. We acknowledge that other crimes exist that have seen increased reach because of technology. However, these crimes have yet to be studied in depth by the research community and, therefore, we decided to focus on the one which the research community has a better understanding of.

Advance free fraud. In this type of scam, the victim is promised a reward (financial or otherwise), but in order to obtain it has to first pay a small fee to the fraudster. After the payment takes place, the victim often does not hear from the scammer again, while sometimes the relationship lasts for long periods of time and the victim is repeatedly defrauded of large sums of money [30].

The archetypal example of advance fee fraud comprises so-called 419 scams [31]. Named after the section of the Nigerian Criminal Code dealing with fraud, these scams became popular in the 1980s, when victims would receive physical letters from an alleged Nigerian prince, looking to transfer large amounts of money outside of the country. To initiate the process, the victim is required to transfer a small amount of money to the fraudster (e.g., to cover legal fees). As it can be imagined, the Internet allowed this type of fraud to flourish, by enabling criminals to instantly reach a large number of potential victims.

Another example of advanced fee fraud is consumer fraud perpetrated on classifieds websites such as Craigslist [32]. As part of this fraud, victims respond to a classified advertisement for a desirable item (e.g., a used car or a rental property) which has much better conditions (such as a lower price) than similar posts. The fraudster responds that they will need a small upfront payment to deliver the goods. After receiving it, the victim will not hear from the fraudster again.

A final example of advanced fee fraud is the online romance fraud. Taking place on online dating sites, this type of fraud usually consists in criminals posing as attractive individuals looking to start a relationship with the victim. Unlike the 419 scam, these online relationships often last for months before the fraudster demands money, for example, to help their family or to open a business [28]. By that time, the victim, who is likely emotionally involved with the persona impersonated by the criminal, is likely to comply. Previous research reported that victims of this crime typically lose between £50 and £240,000 [30]. Unlike other types of advanced fee fraud, however, the psychological damage of losing the fictional relation can be much greater than the financial one.

A common element of every type of advanced fee fraud is the need for criminals to build an enticing narrative that will lure victims into paying the fraudulent fee. To this end, criminals often target specific demographics and impersonate specific personas. For example, previous research showed that romance fraudsters often pretend to be members of the military stationed abroad [33]. By doing so, the fraudsters can build a credible narrative as to why they cannot meet the victim in person, and they can build an emotional connection with the victim, which will increase the chances of theif falling for the scam. Often, fraudsters pretend to be widowed middle-aged men who target widowed women in the same demographic, in an attempt to establish an emotional connection with their victim [28]. In other cases, fraudsters employ psychological tricks to win their victims over, such as applying time pressure or remarking that they specifically selected the victim because of their high moral

characters [31].

More cynically, Herley argues that fraudsters are incentivised to build the most absurd narratives possible, to make sure that only those who are gullible enough to believe them will reply, and that these people will be the most likely to fall for the scam [34]. His argument is that responding to the first boilerplate message is expensive for the fraudster, while sending the first copy to as many victims as they wish is free. For this reason, it is in their interest to rule out those who are not likely to fall for the scam as soon as possible.

Drug dealing. Another category of crimes for which the Internet has offered opportunities is the drug trade. Thanks to anonymising technologies such as Tor [35] and cryptocurrencies [36], online marketplaces have emerged where drug users can purchase illicit substances and have them delivered directly to their home. Research has shown that this business is thriving, despite the instability of these marketplaces, which are often taken down by law enforcement [37, 38]. Online drug markets provide an interesting paradigm switch for drug users, because they remove the need for the buyer to interact with criminals in a physical and potentially unsafe setting. Recent work has shown, however, that the inception of the online drug market has not changed the worldwide drug trade ecosystem: the big players who produce and dispatch drugs remain broadly unchanged, while what has changed is the 'last mile' in the delivery (i.e., how local dealers and drug users get in touch and do business) [29].

Cyber-dependent organized criminals

In this section, we describe crimes that have a financial goal and are carried out using complex technical infrastructures (e.g., botnets [39]). Unlike the cyber-enabled crimes described in the previous section, where the criminal is essentially replicating a physical criminal operation and using the Internet to enhance his/her reach, in the case of cyber-dependent crimes criminals have to set up complex technological infrastructures to achieve their goals. The complexity of these operations has prompted a compartmentalisation in the criminal ecosystem, where each malicious actor specialises in a specific part of a cybercriminal operation (e.g., infecting computers with malware or performing money laundering) and works together towards achieving a common goal. In this section, we provide some examples of cyber-dependent crimes that have been studied by the research literature in recent years. Then, in Section 2, we cover in detail the various elements that criminals need to put in place to make their operations successful.

Email spam. Email spam has been a major nuisance for Internet users for the past two decades, but it has also been at the forefront of very successful criminal operations, who have managed to monetise the sale of counterfeit goods and pharmaceuticals by reaching billions of potential customers through malicious messages [40]. Email spam is defined as *unsolicited bulk email*; this definition highlights the two main elements of the problem: the fact that the messages received by victims are unsolicited (i.e., they were not requested in the first place), and that they are sent in bulk to reach as many victims as possible.

Although the very first spam email was recorded in 1978 [41], email spam rose to prominence in the 1990s, when criminals set up small operations, not dissimilar from the advance-fee fraud ones described in the previous section [42]. The goal of these operations was to sell goods online, which could span from diet supplements to Nazi memorabilia [42]. At this stage, relying on their own expertise and on the help of a small number of associates, criminals would carry out all the activities required to set up a successful spam operation: (i) harvesting email addresses to send the malicious messages to, (ii) authoring the email content, (iii) sending the spam emails in bulk, (iv) processing the orders from people who wanted to purchase the advertised items, (v) reacting to raids by law enforcement (e.g., the seizure of an email server). Although they were still rudimentary compared to the spam operations that came during the next decade, these criminal endeavours prompted the development of legislation to regulate unsolicited bulk emails, such as the Directive on Privacy and

Electronic Communications in the EU,[2] the Privacy and Electronic Communications Regulations in the UK[3] and the CAN-SPAM Act in the US.[4] These pieces of legislation helped prosecute some of those early-day spammers. In 2004, America Online (AOL) won a court case against Davis Wolfgang Hawke, who was selling Nazi gadgets through spam emails. Hawke was sentenced to pay a 12.8M USD fine.

The technical advancements of the early 2000s, and in particular the development of botnets, networks of compromised computers controlled by the same cybercriminal [39], gave unprecedented opportunities to criminals who want to engage in email spam today. Email spam is not a one-person operation anymore, rather it is supported by thriving criminal ecosystems. Spammers can rent botnets from criminals who are specialised in infecting computers with malware [6], purchase lists of target email addresses from specialised actors [43] and sign up to an affiliate programme [44, 45], which will provide the spammer with a way of advertising, as well as taking care of shipments and payments.

The arms race connected to spam mitigation has been going on since the 1990s, with a number of mitigations being proposed [46]. Currently, anti-spam techniques ensure that the vast majority of malicious emails will never reach their victims' mailboxes. To solve this issue, criminals have to send tens of billions of emails [6] to keep their operations profitable. Another issue is that, out of the victims reached by those spam emails that make it through, only a small fraction will purchase the advertised goods and turn a profit for the criminals. Researchers performed a case study for the Storm botnet [47], showing that out of 469 million spam emails sent by the botnet, only 0.01% reach their targets. Of these, only 0.005% of the users click on the links contained in the emails, while an even lower number ends up purchasing items - only 28 users in total out of the 469 million reached, or 0.0004% of the total. Despite this steep drop, McCoy et al. showed that popular spam affiliate programmes were able to make up to 85 million USD of revenue over a three-year period [44]. They also showed that key to this success are returning customers. In fact, spam emails need to reach an interested customer only once, and this person can later keep purchasing on the site without having to worry about spam filters.

Phishing. A particular type of spam is phishing, where criminals send emails that pretend to be from genuine services (e.g., online banking, social network websites) [48]. These emails typically lure users into handing out their usernames and passwords to these services by presenting them with a believable email asking them to visit the website (e.g., to retrieve their latest account statement). By clicking on the link in the email, users are directed to a website displaying fake but realistic login pages. Once they have input their credentials, the criminals gain access to them and they will be able to later log in to those services on behalf of the users, potentially making money directly or selling the credentials on the black market.

For the criminal, a key component to the success of phishing pages is setting up web pages that resemble the original ones as much as possible. To facilitate this task, specialised cybercriminals develop and sell so-called *phishing kits* [49], programs that can be installed on a server and will produce an appropriately-looking web page for many popular services. These kits typically also provide functionalities to make it easier for the criminal to collect and keep track of the stolen credentials [49]. Another element needed by criminals to host these pages is servers under their control. Similar to spam, criminals, researchers, and practitioners are involved in an arms race to identify and blacklist phishing Web pages [50], therefore it does not make economic sense for criminals to set up their own servers. Rather, criminals often host these websites on compromised servers, for which they do not have to pay [51].

After stealing a large number of credentials, criminals can either exploit these accounts themselves

[2] https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32002L0058

[3] https://en.wikipedia.org/wiki/Privacy_and_Electronic_Communications_(EC_Directive)_Regulations_2003

[4] https://en.wikipedia.org/wiki/CAN-SPAM_Act_of_2003

or sell the usernames and passwords on the black market. Previous research has shown that often these criminals log into the accounts themselves and spend time evaluating their value, for example, by looking for financial information in webmail accounts.[9, 52].

Financial malware. Another popular criminal operation is financial malware. In this setting, criminals aim to install malware on their victims' computers and steal financial credentials such as credit card numbers and online banking usernames and passwords. This trend started with the Zeus malware, which criminals could purchase on the black market and use to set up their operations [53]. Once installed on a victim computer, Zeus would wait for the user to visit a website on a pre-configured list of interesting ones that the criminal could specify. It would then record usernames and passwords as the user typed them in, and send them to the command and control server set up by the criminal.

A more sophisticated information stealing botnet was Torpig [54]. Unlike Zeus, Torpig used a botnet-as-a-service model, where a single specialised criminal was responsible for hosting the botnet infrastructure, while other criminals could run their campaigns to infect victim computers, pay a fee to use the torpig infrastructure and later retrieve the stolen credentials. Researchers showed that, in 2009, the Torpig botnet was able to steal 8,310 unique bank account credentials and 1,660 unique credit card numbers over a ten-day period. [54].

To monetise their operations, cybercriminals have two choices: the first is to sell the stolen financial information on dedicated underground forums [55]. The price that criminals can ask for these credentials varies based on the type of records that they were able to steal. For example, on the underground market there are two types of credit card records that are traded: *dumpz*, which contain the information that allows a criminal to clone a credit card (i.e., card number, expiration date, security code), and *fullz*, which also contain the billing address associated with the card. Fullz are worth more money on the black market, because they allow miscreants to purchase items online.

A related type of crime that is becoming more popular is card skimming [56]. In this cyber-enabled crime, criminals install devices on ATM machines which collect details of the cards inserted into the machines by unwitting users. The criminal can then collect the devices to retrieve the stolen financial credentials. While this type of crime is serious, it is also a good example of the limitations of physical crime compared to their online counterparts: the need for physical action by the criminal limits the scale of the operation, while financial malware operations can affect much higher numbers of victims. For example, the Torpig malware was installed on over 100,000 computers [54].

Click fraud. Web advertisements are the main way the Web is monetised. A Web administrator can decide to host advertisements on his/her website, and whenever visitors view them or click on them they receive a small fee from an advertiser. To mediate this interaction, specialised services known as *ad exchanges* have emerged. Because of their easy monetisation, Web advertisements are ripe for fraud. In particular, criminals can host advertisements on their own websites and then generate 'fake' clicks (e.g., by using bots). This results in an ad exchange paying criminals for ad impressions that were not 'genuine,' eventually defrauding the advetiser.

Once again, criminals are involved in an arms race with ad exchanges, who are interested in keeping fraud on their services minimal. To help criminals generate large numbers of clicks and remain under the radar by gaining access from large numbers of IP addresses, so-called *click fraud* botnets have emerged. An example is Zeroaccess [57], which was active in 2013. On an infected machine, this malware would act like a regular user, browsing websites and clicking on advertisements that its owner chose. Researchers showed that this botnet was responsible for losses to the advertising industry of approximately 100,000 USD per day [57].

Cryptocurrency mining. With the increasing popularity of cryptocurrencies, a new opportunity has opened up for criminals: using infected computers to mine currency. In 2014, Huang et al.revealed this threat, showing that botnets were used to mine Bitcoin [58]. While revealing this new monetisation for malware, the authors also concluded that these operations did not appear to be making much money, totaling at most 900 USD a day.

A more recent study, however, showed that cryptocurrency mining by botnets could be much more rewarding than previously thought. Pastrana and Suarez-Tangil showed that by mining Monero and using a number of techniques to increase their chances of mining currency (e.g., using mining pools) criminals could make up to 18 nillion USD over a two-year period. [59].

Another emerging trend in cybercrime comprises leveraging Web browsers to mine cryptocurrencies. Instead of installing malware on victim computers and using them for mining, miscreants add scripts to webpages and have their visitors mine cryptocurrencies. This type of malicious activity is called *cryptojacking*. Although using these scripts is not necessarily illegal (i.e., Web administrators can legitimately install them on their webpages in a similar way to advertisements), criminals have been caught adding them to compromised websites on multiple occasions. Konoth et al. showed that a malicious campaign can make #£31,000 over a week [60], while Rüth et al. [61] showed that 1.18% of the mined blocks in the Monero blockchain can be attributed to `Coinhive`, the most popular cryptojacking library.

Ransomware. The newest trend in malware is *Ransomware*. As part of this operation, criminals infect their victim systems with malware which encrypts the user's personal files (e.g., documents) and sends the encryption key to the criminal, who then asks for a ransom in exchange for giving the user access to her data again [62]. The idea of malicious software that uses public key cryptography to hold the victim's data hostage is not new, and it was theorised by Yung in 1996 already [63]. In 20 years, however, the technological advancements on the malware delivery end have made it possible to reach large numbers of victims, and the introduction of anonymous payment methods such as Bitcoin has made it safe for criminals to collect these payments.

Ransomware is, at the time of writing, the golden standard for cybercriminals. This type of malware operation has solved the monetisation problems that were so important in other types of cybercriminal schemes: the criminal does not have to convince the victim to purchase a good, like in the case of email spam, or to fall for a fraud, like in the case of phishing. In addition, the victim is highly incentivised to pay the ransom, because the probability that the criminals have encrypted files that the user will need (and for which she has no backup copy) is high. In fact, recent research was able to trace 16 million USD in payments on the Bitcoin blockchain that can be attributed to ransomware campaigns [64].

Although the most sophisticated ransomware campaigns involve encrypting the victim's files, Kharraz et al. showed that it is not uncommon for malware authors to use other techniques to lock the victim out of his/her computer [65]. These techniques include setting up a password-protected bootloader and not giving the password to the user unless he/she pays. While these techniques are likely to yield a profit for the criminal, they are also easier to mitigate, as the victim's files are safe on the computer and a simple clean up of the malware (and restoring the original master boot record) can fix the problem.

Denial of service. A feature that all Internet-connected devices have is network connectivity. A criminal can leverage the bandwidth of an infected device to perform a distributed denial of service (DDoS) attack against a target. Criminals can simply use the bandwidth generated by the botnet, or leverage *amplification attacks* (i.e., network traffic generated by misconfigured network devices) to enhance the power of their DDoS attacks [66].

The criminals can then set up services where they offer DDoS for hire. To hide the illicit nature of their business, these services often advertise themselves as 'stress testers', services that a Web administrator can use to test how their Web applications perform under stress [67]. In reality, however, these services do not check whether the customer purchasing a DDoS attack is actually the same person who owns the target domain.

Hacktivists

While criminals driven by profit are a big threat, not all adversaries are driven by money. In particular, we define the act of computer crime motivated by a political goal as *hacktivism* [68]. These crimes can take various forms, from denial of service attacks [68] to compromising computer systems with the goal of releasing sensitive information to the public [69]. There is an ongoing debate among scholars on whether actions by hacktivists fall under political activism (e.g., civil disobedience) or cyber terrorism [70]. Holt et al. studied cyber attacks carried out by far left groups in the US and found that there was an increase in online attacks during periods that observed a decrease in physical violence from those same groups [71].

Denial of service. The practice of hacktivism started in the 1990s with *netstrikes* [72]. As part of this practice, Internet users would connect to target the websites simultaneously to deplete their resources and make them unresponsive. This was often done to protest against actions and policies by government agencies and corporations. Twenty years later, with the increased sophistication offered by technology, hacktivist groups such as Anonymous [73] took the idea of netstrikes and makde it bigger in size. This collective became popular for launching denial of service attacks against organisations that were guilty of performing actions that did not match their moral stance, such as governments linked to the repression of the Arab Spring, credit card companies who would not make donations to entities such as Wikileaks or radical religious organisations.

To perform their attacks, Anonymous would ask its sympathisers to install a computer program, called LOIC (Low Orbit Ion Cannon), which would act as a bot in a botnet: their controller would use the computer's bandwidth to carry out a denial of service attack against a chosen target. The difference with traditional botnets (and the ones used to carry out DDoS attacks in particular) is that the user is accepted to be part of it by installing the LOIC program.

Data leaks. Another trend that we have been observing in recent years in the area of hacktivism is the release of stolen documents into the public domain, for example, to raise awareness about secret surveillance programs by governments [74]. A prominent example of an organisation that performs these data leaks is Wikileaks [69]. Similar techniques have also been used by Anonymous (e.g., about the identity of 1,000 Ku Klux Klan members).

Web Defacements. The last trend that is typical of politically-motivated actors is *Web defacement* [75]. As part of this activity, miscreants exploit vulnerabilities (ranging from weak passwords to software vulnerabilities) in the websites of organisations they disagree with, and use them to change the home page of the website to a politically-charged one. An example of an organisation that is prominently using Web defacements to spread their message is the Syrian Electronic Army [76], a group of hackers close to the Assad regime. Although popular with criminals with a political agenda, Web defacement is not just their prerogative. In fact, Maimon et al. showed that this is a popular way for early career cybercriminals to prove their worth [77].

State actors

Another type of malicious actor involved in adversarial behaviours online comprises nation states. In the past few years, we have observed an escalation in the use of computer attacks by state actors to achieve their goals. Broadly speaking, this type of attack differs from those performed by financially motivated cybercriminals for two reasons:

1. Commodity cybercrime needs to gather as many victims as possible to maximise their profits. For instance, criminals setting up a botnet to steal financial information from their victims will want to reach the highest possible number of victims to improve their revenue. This means that the cybercriminal's attacks need to be either generic or diversified enough to cover a large population of devices (e.g., by using exploit kits, as explained in Section 2). In a state-sponsored

attack, on the other hand, there is no need to make money, and usually the victim is well defined (e.g., a specific organisation or a person of interest). In this setting, the attack can be tailored to the victim; this increases the chances of success, because of the time that can be spent designing the attack and the fact that the attack will be unique (e.g., by using a zero day attack [78]), and it will be unlikely that existing protection software will catch it.

2. Because of the need to make money, traditional cybercriminals need their attacks to be fast. This is not the case for state-sponsored attacks, where the reward for achieving its goal (e.g., stealing sensitive information from a government) makes it acceptable to wait for long periods of time before finalising the attack.

State-sponsored attacks fall broadly into two categories, depending on the purpose of the attack: sabotage and espionage. In the following, we describe these two types of attacks in more detail.

Sabotage. Modern critical infrastructure can be disrupted by electronic means. Research has shown that it is not uncommon for critical facilities such as power plants to have some sort of network connectivity between the computers controlling the machinery and the ones connected to the Internet [79]. In the case of a state adversary, even having network security appliances to guard the boundary between the two networks is not enough, since, as we said, attacks can be so sophisticated and tailored that off-the shelf solutions fail to detect them [80]. Once a piece of malware manages to get into the control network, it could make the machinery malfunction and potentially destroy it. Even when there is a physical separation between the control network and the wider Internet, attacks are still possible when we are faced with adversaries with virtually unlimited resources [80].

A prominent example is the Stuxnet worm [81, 80], a sophisticated attack performed against the Nathanz nuclear enrichment facility in Iran in 2010. Allegedly, the malware was introduced into the facility by first infecting the laptop of one of the consultants who was maintaining the machinery. Once the malware was in the right environment, it identified the pieces of equipment that it was designed to target and sabotaged the enrichment experiments, making the centrifuges spin out of control. To date, Stuxnet is a textbook example of the lengths to which state-sponsored attackers can go to achieve their objectives, and of the sophistication that their attacks can achieve.

Espionage. Another goal that state-sponsored actors have for their attacks is spying on opponents and prominent adversaries. Research has shown that state actors make prominent use of spearphishing (i.e., targeted phishing) to lure activists into installing malware that is later used to spy on them [9, 82]. In other cases, state actors infect sensitive systems (e.g., servers in large corporations), with the goal of stealing sensitive information [83]. The security industry has dubbed these long-standing, sophisticated attacks *Advanced Persistent Threats*.

2 The Elements of a Malicious Operation

As we showed in Section 1, malicious operations can consist in rather complex infrastructure, particularly in the case of organised crime, which is mostly motivated by two facts. First, the criminal needs these operations to be as efficient as possible (and consequently make the highest possible profit). Second, multiple actors (law enforcement, security companies, the users themselves) are constantly attempting to take down these malicious operations, and the criminal has, therefore, a need to make them resilient to these takedown attempts.

To ensure that the criminals' needs are met in this scenario, in recent years we have witnessed a *specialisation* in the cybercriminal ecosystem, where different actors specialise in a specific element required for the operation to succeed;the miscreants then trade these services with each other on the black market. In this section, we provide an overview of the elements required for a cyber-dependent orgranised criminal operation to succeed, as described in Section 1. Many of the elements discussed, however, also apply to the other types of adversarial behaviours described in that section.

Affiliate Programmes

The main goal of organised crime is to make money from their operations. This requires not only a well-oiled technical infrastructure to make sure that their botnets operate properly but, perhaps more importantly, a working method to collect payments from victims, while making sure that all the actors involved in the operation get paid.

In the cybercriminal world, this is typically done through *affiliate programmes*. An affiliate programme is a scheme where main organisation provides a 'brand' and all the means required to carry out orders, shipments and payments. Affiliates can join the program, direct traffic to the platform, and get a cut of the sales that they are responsible for. Although this scheme exists for legitimate businesses (e.g., Amazon has an affiliate programme), it has been particularly successful for cybercriminal operations. The main difference between legitimate and criminal affiliate programmes is that the second category of operations typically deals with products that are considered illegal in most jurisdictions (e.g., counterfeit pharmaceuticals, gambling, counterfeit designer products) and they typically endorse criminal promotion techniques (e.g., the use of malware or black hat search engine optimisation).

Affiliate programmes are popular in the cybercriminal world because they mean affliliates do not have to set up their operations from start to finish, but rather focus on attracting traffic, for example by setting up botnets and sending email spam advertising the affiliate marketplace. The first successful examples of affiliate programmes for cybercrime were centred around email spam, and were advertising counterfeit pharmaceuticals [40, 44, 45]. However, affiliate programmes are present in most types of cyber-dependent crime, an example being the Cryptowall ransomware operation.[5]

In addition to providing the monetisation necessary for cybercriminal operations, affiliate programmes also act as facilitators for criminals to get in contact and trade the services that are needed for the operation to succeed. This is typically done by setting up a forum where affiliates can trade their services [44, 6]. Gaining access to these forums typically requires vetting by the affiliate programme administrators.

Infection vectors

As discussed earlier, the first step required by criminals to perform a malicious activity is often infecting their victims with malware. To this end, the criminals need to first expose their potential victims to the malicious content, and then have them install it on their machines (through either deception or by exploiting a software vulnerability in their system). In the following, we survey three popular methods on delivering malware to victim computers. Note that, while other infection vectors are possible, such as physical access to a network or hijacking a wireless network, to date we are not aware of any large-scale compromise involving these infection vectors, and therefore we do not focus on them.

Malicious attachments. Possibly the oldest method of delivering malware is attaching malicious software to spam emails, disguising it as useful content that the user might want to open. This spreading technique was made popular by email worms in the early 2000s, such as the 'I love you' worm [84], but it is still a popular way of delivering malware to victims [6]. In the commoditised economy described previously, it is often the case that a criminal who wants to spread a malware infection pays another criminal who already has control of a botnet to deliver the payloads [54]. To be successful, the content used for this infection vector needs to convince the user to click on the attachment and install it. To this end, criminals often use deception techniques to make the content look interesting and appealing, similar to the techniques discussed for phishing [48]. This deception falls into the area of social engineering [85].

Black hat search engine optimisation. Search engine optimisation (SEO) is a popular practice whereby webmasters optimise their content so that it is better indexed by search engines and ap-

[5] https://www.secureworks.com/research/cryptowall-ransomware

pears among the first hits for relevant searches. Cybercriminals are also interested in having their malicious Web pages appear high in search results, because this increases the chances that potential victims will find them and click on them. To accomplish this, specialised criminals offer black hat SEO services. As a result of these services, malicious websites are pushed high up in search engine rankings for keywords that are unrelated to the website [86]. This happens particularly often in proximity with popular events (e.g., sports and political events), because people will be more likely to search for keywords related to the event. To achieve effective black hat SEO, cybercriminals compromise vulnerable websites and use them to promote their customers' webpages (e.g., by adding invisible links and text pointing to the target webpage).

Drive-by download attacks. Although deceptively luring users into installing malware works, having an automated method that does not require human interaction is more advantageous for cybercriminals. To this end, cybercriminals have perfected so-called *drive-by download* attacks [87]. As part of one of these attacks, the victim visits a webpage under the control of the criminal (e.g., encountered through black hat SEO). The webpage contains malicious JavaScript code that will attempt to exploit a vulnerability in the user's Web browser or in one of its plugins. If successful, the Web browser will be instructed to automatically download and install the malware.

To host their malicious scripts, cybercriminals often compromise legitimate websites [88]. An alternative trend is purchasing Web advertisement space and serving the malicious content as part of the ad, in a practice known as *malvertisement* [89].

Infrastructure

Another important element that criminals need for their operations to succeed is where to host their infrastructure. This is important for both affiliate programmes (e.g., where to host fraudulent shopping websites) as well as for botnet operations. Law enforcement and Internet Service Providers (ISPs) are continuously monitoring servers for evidence of malicious activity [90], and will take them down if this activity can be confirmed, which would put the criminal operation in jeopardy.

Bulletproof hosting service providers. To maximise the chances of their operations being long-lived, cybercriminals resort to using so-called *bulletproof hosting service providers* [91, 40]. These providers are well known not to comply with law enforcement takedown requests. This is made possible by either being located in countries with lax cybercrime legislation, or by the service provider operators actively bribing local law enforcement [40]. Bulletproof hosting service providers typically charge their customers more money than a regular ISP would. As such, they become a hotspot of illicit activity, since malicious users congregate there because of their guarantees, but legitimate users have no incentive to use them. Despite providing higher guarantees for cybercriminals, bulletproof hosting service providers are not invincible to takedown efforts. In particular, ISPs need to be connected to each other to be able to route traffic, and an ISP that is uniquely hosting malicious content could be disconnected by the other providers without many consequences for legitimate Internet traffic [40].

Command and control infrastructure. A botnet requires a command and control (C&C) infrastructure that infected computers can be instructed to connect to, receive orders and report on progress in the malicious operation. Originally, botnets would use a single command and control server, although this would be a single point of failure. Even assuming that the server was hosted by a bulletproof hosting provider, and could not therefore be taken down, the fact that the server had a unique IP address meant that it could easily be blacklisted by security companies.

To mitigate this problem, cybercriminals came up with C&C infrastructures that are redundant and more difficult to take down. An example is the *multi-tier* botnet infrastructure, where bots are instructed to connect to an intermediary C&C server, which is then responsible for relaying the information to and from a central control server [92]. This infrastructure makes the botnet more resilient, because even if some of the relays are taken down, the central C&C is still operational and additional

relays can be added. In addition, the infected computers never see the IP address of the central C&C server, making it more difficult to locate and take down. A variation of this model is *peer-to-peer* botnets, where infected computers with particularly good connectivity and public IP addresses are 'elected' to act as relays [93]. This infrastructure increases the flexibility that the criminal has and reduces the cost of the operation, because the criminal does not have to spend money to install relays. However, the botnet infrastructure becomes vulnerable to infiltration, whereby researchers can create fake bots, be elected as relays and are thus suddenly able to monitor and modify the traffic coming from the central C&C [47].

Additional techniques used by cybercriminals to make their control infrastructure more resilient are *Fast Flux* [94], where criminals use multiple servers associated with the C&C infrastructure and rotate them quickly to make takedowns more difficult, and *Domain Flux* [95], in which the domain name associated to the C&C server is also rotated quickly. Both methods are effective in making the operation more resilient, but they also make the operation more expensive for the criminal to run (i.e., they have to purchase a more servers and domain names).

Specialised services

In this section, we describe specialised services that help criminals to set up their operations. In addition to these dedicated malicious services, others that have a legitimate use (e.g., VPNs, Tor) are also misused by criminals, for example hosting drug market websites on the Dark Net [96, 38].

Exploit kits. In the previous section, we saw that drive-by download attacks are a powerful weapon that a cybercriminal can use to infect computers with malware without any human interaction. The problem with effectively performing these attacks, however, is that they require an exploit to a software vulnerability in the victim's system. Since cybercriminals want to infect as many victims as possible, it is challenging to find an exploit that can work on the systems of the majority of potential victims. In addition to this issue, exploits do not age well, since software vendors routinely patch the vulnerabilities that they know about. A cybercriminal performing a sustained drive-by download operation, therefore, would need to continuously collate exploits to multiple vulnerabilities, a task that is unfeasible, especially when the criminal also has to run other parts of the business (e.g., the monetisation part). Once a victim visits the exploit kit's webpage, this tool first fingerprints the victim's system, looking for a potential vulnerability to be exploited. It then delivers the exploit to the victim. If successful, the victim's computer is instructed to download the malware of the customer's choice.

These issues have created an opportunity for specialised criminals to provide services for the rest of the community. This has led to the creation of *exploit kits* [97], which are tools that collect a large number of vulnerabilities and are sold on the black market for other criminals to use. An exploit kit is typically accessible as a Web application. Customers can point their victims towards it by compromising websites or using malicious advertisements.

Pay-per-install services. Infecting victim computers and maintaining a botnet is a complex task, and research has shown that malware operators who attempt to do so without the proper expertise struggle to make profits [98]. To solve this issue and satisfy the demand for stable botnets, a new criminal service has emerged called pay-per-install (PPI) services [99]. PPI operators are proficient in setting up a botnet and having it run properly. Other criminals can then pay the PPI operator to install malware on the infected computers on their behalf. PPI services typically offer a good level of choice granularity to their customers, who not only choose how many infections they want to install, but also their geographical location (with bots in developed countries costing more than infections in developing ones [99]).

An advantage of using PPI services is that they make their customers' cybercriminal operations more resilient: if their malware stops working, for example, because law enforcement has taken down the C&C servers that it uses, the criminal can resume operations by asking the PPI operator to install an updated version of their malware on the victim machines. For this reason, this *malware symbiosis*

between PPI services and other botnets is very common in the criminal ecosystem (see, for example, the symbiosis between Pushdo and Cutwail [6], and between Mebroot and Torpig [54]).

Human services

In this section, we discuss the auxiliary services that are needed for an end-to-end cybercriminal operation to succeed. Although these elements are not usually thought to be part of cybercrime, they are as important to the success of a cybercriminal operation as the more technical elements.

CAPTCHA solving services. In some cases, cybercriminals need to set up accounts on online services to initiate their operations (e.g., a spam operation running on social networks [7, 8]). To protect themselves against large-scale automated account creation, however, online services widely use CAPTCHAs, which are notoriously difficult for automated programs to solve. To solve this problem faced by cybercriminals, new CAPTCHA solving services have been established [100]. These services take advantage of crowdsourced workers. Once the CAPTCHA solving customer encounters a CAPTCHA, this is forwarded by the service to one of these workers, who will solve it. This way, the customer can proceed and create the account on the online service.

In other cases, online services require whoever has created an online account to receive a code texted to a phone number and issue that code back to the service. To overcome this issue, cybercriminals can use services that automate this type of interaction [3].

Fake accounts. Since creating fake accounts is time consuming and requires the use of auxiliary services such as CAPTCHA solvers, cybercriminals have started specialising in the creation of fake accounts on multiple online services, and selling them on the black market [101]. Accounts on different services can have different prices, depending on the ease of creating new accounts on the platform and on how aggressively of the service suspends suspected fake accounts.

A problem with newly purchased fake accounts is that they do not have an established 'repuation' on the social network, thus reducing their credibility to potential victims and their reach in spreading malicious messages. This can be mitigated by using 'reputation boosting' services, which help to build a network of contacts for accounts that otherwise would not have any. Examples of these are services offering fake likes on Facebook [102] and luring compromised accounts into following the service's customers on Twitter [103].

Content generation. In some cases, cybercriminals need to set up fake content to send to their victims, whether this is for spam emails, fake websites used for black hat SEO or online social network sites. To generate this content, the criminals can recruit workers on underground forums [104].

Money mules. The main goal of many cybercriminal operations is to make money from their victims. However, extracting money from an operation is not easy. In the case of bank fraud, for example, even if the criminals obtain access to the victim's bank account, they still need to transfer money to accounts under their control without being detected and apprehended.

To facilitate these monetisation operations, criminals take advantage of *money mules* [105]. These are people who are recruited by criminals to perform money laundering operations and make it more difficult for law enforcement to track the money obtained from an illicit operation. In a money mule scheme, the criminal recruits a person to act as a mule and sends them money by using traceable means (e.g., a check or a wire transfer). The mule is then instructed to transfer the money to an account under the criminal's control by using untraceable means (e.g., Western Union). The mule is also told that they can keep a percentage of the amount as a payment. Since these untraceable transactions need to be carried out in person by the mule, they constitute a weak point in the monetisation operation, meaning that law enforcement could identify and arrest the mule before the money is transferred. In fact, even if stolen money is never mentioned, the mule is participating in money laundering when he/she accepts this job.

An alternative way of monetising malicious operations, which is used in the case of stolen credit

cards, is *reshipping mules* [106]. In these operations, criminal agencies recruit unsuspecting users advertising a 'shipping agent' job. Then other criminals can recruit the services of these agencies, and purchase expensive items using stolen credit cards (e.g., electronics, designer goods), while sending them to the mule's home address. The mule is then instructed to open the packages and reship the goods to a foreign address, where they will be sold on the black market.

Payment methods

As criminals need to have money transferred to them, they can use a number of different payment methods, each carrying a different level of risk and being more or less familiar to the victims.

Credit card processors. Most transactions online are performed by credit cards. To collect as many customers as possible, cybercriminals tend to accept credit card payments too. McCoy et al. showed that 95% spam affiliate programmes between 2007 and 2012 accepted credit card payments [44], and that DDoS services that did not accept credit cards suffered with regard to the numbers of customers that they were able to attract [67]. Credit card processors keep track of the chargebacks that a company has on its accounts, and too many complaints from customers usually result in the company's accounts being terminated. For this reason, many cybercriminal operations offer 'customer support' to their victims, offering refunds if they are not satisfied with their purchases [107].

A challenge that cybercriminals face is finding banks that are willing to process their payments. Typically, these banks would charge them higher transaction fees (10-20%) to cover the risk of dealing with criminal operations [44]. Despite these increased fees, it is not guaranteed that the criminal operation will be safe: similar to what happens with bulletproof hosting ISPs, banks need to maintain good relations with their peers, otherwise they will be disconnected from the financial network [108].

Paypal. Another payment method that is familiar to users is Paypal. For this reason, Paypal is often accepted by criminals offering illicit services. While user friendly, criminals face the issue that the platform is centralised, and Paypal can keep track of fraudulent payments and terminate the accounts that are found to be in breach of the terms of service [109].

Western Union and other 'untraceable' payments. Other forms of payment offer more anonymity for cybercriminals, and are less risky as well as being not as well regulated. Examples are money exchanges (e.g., Western Union, Money Gram) or pre-paid vouchers (Money Park). These are often used by criminals to transfer funds [110]. To cash the money, these services only require a unique code and an identification document. Depending on the country where the criminal is located, however, the ID requirement might not be very rigorous.

Historically other 'anonymous' payment methods have existed such as Liberty Reserve, Web Money and eGold [3]. These virtual currencies allowed criminals to easily make payments as they took advantage of the loose regulations in their country of origin (e.g., Liberty Reserve was based in Costa Rica). After crackdowns on these payment methods by law enforcement, criminals moved to other payment methods.

Cryptocurrencies. At the time of writing, probably the safest form of payment for cybercriminals is cryptocurrencies. These payments have become popular for multiple types of cybercriminal operations, from ransomware [65] to drug market payments [37]. While research has shown that customers are reluctant to use services that only accept cryptocurrencies [67], this type of payment still works when victims have no choice (e.g., in the case of ransomware) or are very motivated (e.g., in the case of drug markets).

While more anonymous than other payment methods, research has shown that payments made in Bitcoin can be traced [111]. In addition, often cryptocurrencies need to be converted into real money by criminals, and the money ceases to be anonymous at that point. Additional concerns arise from the risks involved in making payments on cryptocurrency exchanges. Moore et al. showed that it is not uncommon for Bitcoin exchanges to suffer breaches that result in losses of currency [112]. Exit

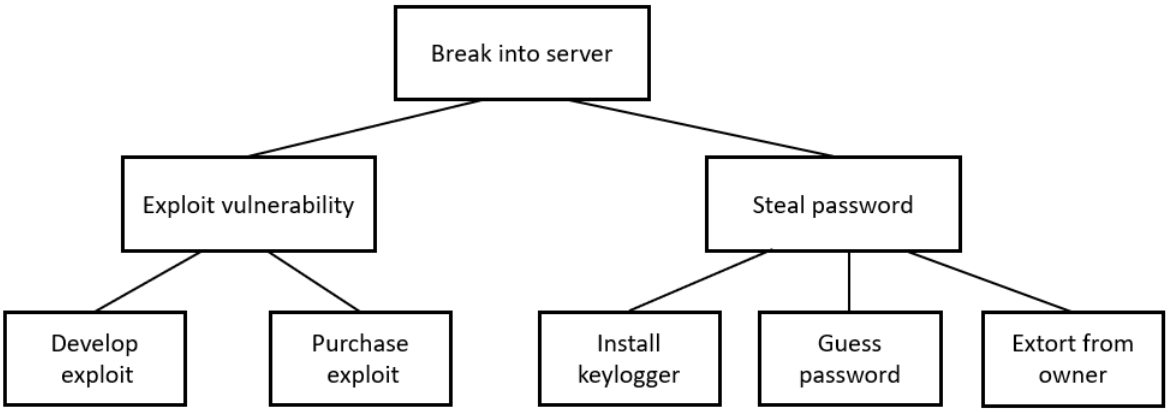


Figure 1: Example of an attack tree describing the action of breaking into a server.

scams, where an exchange vanishes with all the currency stored in it, are also a problem [113].

3 Models to Understand Malicious Operations

As shown in the previous sections, malicious operations can be quite complex and entail multiple technical elements and multiple actors. It is, therefore, necessary for defenders to have the appropriate means to understand these operations, so that they can develop the best countermeasures. In the following, we survey a number of models that have been proposed to model malicious operations. These models come from a number of research areas, including computer security, criminology and war studies.

Attack trees

The first way to model attacks against computer systems involve *attack trees* [114]. Attack trees provide a formalised way of visualising a system's security during an attack. In an attack tree, the root node is the goal of the attack, and its child nodes are the ways an attacker can achieve that goal. Going down the tree, each node becomes a sub-goal that is needed for the attack to succeed, and its children are possible ways to achieve it.

Figure 1 represents an example of an attack tree. In this example, the attackers aim to compromise a server. To do this, they have two choices: they can either exploit a vulnerability or they can obtain the password to the root account and log in using normal means. To exploit a vulnerability, they can either develop the exploit themselves or purchase an already existing one, perhaps through an exploit kit. If the attackers decide to use the account's password to log into the server, they first need to obtain it. To do this, they can either install malware on the server administrator's computer to log the password as she inputs it (i.e., a keylogger), guess the password using a list of commonly used ones or perform a bruteforce attack, and finally extort the password from the owner. The attack graph could then be further refined with the possible ways the attacker could perform these actions (e.g., extorting the password by blackmailing the owner, by kidnapping her etc.).

Attack trees allow two types of nodes, '*or*' nodes and '*and*' nodes. 'Or' nodes represent the different ways attackers can achieve a goal (i.e., the children of any node in Figure 1). 'And' nodes, on the other hand, represent the different steps that all need to be completed to achieve the goal. Once the tree has been created, security analysts can annotate it to assess the system's risk to the attack, for example, by marking the various attack strategies as feasible or unfeasible, by assigning likelihood scores to them or by estimating the cost for an attacker to perform a certain action. The scores can then be propagated along the tree following specific rules [114] to assess the overall feasibility and likelihood of the attack.

Kill chains

Another useful tool that can be used to model and understand attacks is *kill chains*. In the military context, a kill chain is a model that identifies the various phases involved in an attack.[6] In the computer world, Hutchins et al. developed a *Cyber Kill Chain* [115] that models the different steps involved in a malicious operation conducted against computer systems. In their model, Hutchins et al. identify seven phases. The model is designed for operations where the attacker identifies, compromises and later exploits a computer system, and, therefore, not all the phases apply to all the adversarial behaviours discussed in this document. The seven phases are the following:

1. **Reconnaissance**, when attackers identify possible targets. This phase could comprise an attacker scanning the network looking for vulnerable servers or a spammer purchasing a list of victim email addresses on the black market.
2. **Weaponisation**, when an attacker prepares the attack payload for use. This could consist in developing a software exploit against a newly identified vulnerability or crafting an advance-fee-fraud email.
3. **Delivery**, when the attacker transmits the payload to its victim. This could consist in setting up a malicious webserver, purchasing advertisement space to perform a malvertising attack or sending an email containing a malicious attachment.
4. **Exploitation**, when the target's vulnerability is exploited. This phase could entail a drive-by download attack, or the victim being lured into clicking on a malicious attachment through deception.
5. **Installation**, when malicious software is downloaded, thus allowing the attacker to benefit from the victim machine. In their paper, Hutchins et al. considered an attacker wanting to maintain constant access to the victim computer, using a type of malware known as a *remote access trojan* (RAT) [116].
6. **Command and control**, when the attacker establishes a C&C infrastructure and a communication protocol to control the infected computer.
7. **Actions on objectives**, when the infection is monetised. This could entail stealing sensitive information from the victim computer, encrypting the victim's data with ransomware, mining cryptocurrencies, etc.

For each of the seven steps, Hutchins et al. identified strategies to disrupt the malicious operations, following five possible goals (Detect, Deny, Disrupt, Degrade, Deceive). Examples of these techniques include patching vulnerabilities, setting up intrusion detection systems on the network or deceiving the attacker by setting up honeypots [117].

Similar kill chains have been proposed by other researchers over the years. An example is the one proposed by Gu et al. to model botnet infections [118]. In this model, the authors identify five phases where an infection is separated: an inbound scan (similar to phase one in the previously described model), an inbound infection (similar to phase four from the previous model), an 'egg' download (analogous to phase five), a C&C phase (the same as phase six), and an outbound scan. At the time of developing this model, botnets were mostly acting as computer worms as [119], scanning for vulnerable computers, infecting them, and using them to propagate further. While this model correctly depicted early botnets, it ceased to map reality when botmasters started using other methods to install their malware and monetise their infections. Nowadays, worms are almost extinct, with the exception of the infamous WannaCry malware [120]. This example shows that it is difficult to develop models of attacker behaviour that are resilient to changes in the modus operandi of attackers.

[6] https://en.wikipedia.org/wiki/Kill_chain

Environmental criminology

While cybercrime is a relatively new threat, physical crime has been studied by scholars for decades. It is, therefore, interesting to investigate whether this established body of knowledge can be applied to better understand and mitigate the emerging threat of online crime. Environmental criminology, in particular, is a branch of criminology that focuses on criminal patterns in relation to the space where they are committed and to the activities of the actors involved (victims, perpetrators, and guardians) [121]. A particular challenge that arises when we attempt to apply environmental criminology theory to cybercrime is that the concept of 'place' on the Internet is not as well defined as in the real world. In the following, we briefly review the key concepts of environmental criminology, and provide some examples of how they could be applied to mitigating Internet crime.

Routine activity theory. *Routine activity theory* is another commonly used concept in environmental criminology, postulating that the occurrence of crime is mostly influenced by an immediate opportunity for one to commit a crime [122]. In particular, routine activity theory states that for a crime to happen, three components need to converge: (i) a motivated offender, (ii) a suitable target and (iii) the absence of a capable guardian.

These concepts could be useful for better modelling malicious activity online. For example, research has shown that botnet activity reaches a peak during daytime, when most vulnerable computers are switched on and the victims are using them, while it drops significantly overnight [54]. In routine activity theory terms, this can be translated to the fact that when more potential victims are online, the opportunity for criminals to infect them increases and this results in an increase in botnet activity.

Rational choice theory. *Rational choice theory* aims to provide a model as to why offenders make rational choices to commit crime [123]. In the case of cybercrime, this model could be useful for understanding the reaction of criminals to mitigation as a rational choice, and help to model the implementation issues introduced by situational crime prevention such as displacement. For example, when a bulletproof hosting provider is taken down by law enforcement, what factors play a part in the criminal's choice of the next provider?

Pattern theory of crime. Another theory, called the *pattern theory of crime*, allows researchers to identify various places that are related to crime. These places are likely to attract offenders (crime *attractors*), they generate crime by the availability of crime opportunities (crime *generators*) and they enable crime by the absence of place managers (crime *enablers*).

Although defining places in cyberspace is not as straightforward as in physical space, thinking in terms of pattern theory can help identify locations that are hotspots for cybercrime, whether they are particularly appealing targets, such as corporations storing sensitive data (attractors), poorly configured systems that are easier to compromise (generators) or online services with poor hygiene that do not react promptly to spam/malware posted on their platforms (enablers). Identifying these hotspots can then be used to design appropriate countermeasures against the malicious activity (e.g., to whom to direct education campaigns).

Situational crime prevention. Situational crime prevention comprises a set of theories and techniques that aim to reduce crime by reducing the opportunities for crime [124]. The ideas behind situational crime prevention are based on three main concepts, which also apply to cybercrime:

- Crime is much more likely to happen in certain places (*hotspots*). This idea applies to the context of cybercrime. As we have seen, criminals tend to concentrate their malicious servers in bulletproof hosting service providers, which provide them with guarantees that their operations can continue for long periods of time. At the opposite end of the spectrum, regarding victims, criminals tend to target computers with vulnerable software configurations, which also constitute hotspots in this acception.
- Crime is concentrated in particular 'hot products'. This also applies to cybercrime, with mis-

creants focusing on whichever operations yield the highest profits (i.e., at the time of writing, ransowmare).

- Repeat victims are more likely to experience crime compared to other people. In the context of cybercrime, the same concept applies. A vulnerable computer that is not patched is likely to be compromised again [119]. Similarly, in the case of advance fee fraud, victims are likely to repeatedly fall for the fraud, because the narrative used by the criminals particularly resonates with them [30]. In addition to the natural predisposition of victims to fall for similar scams again, criminals actively seek to contact past victims of fraud, by compiling so-called *suckers lists* and sharing them with each other [125].

To reduce the opportunities for crime, situational crime prevention proposes five categories of mitigations. In the following, we list them along with some examples of mitigations against cybercrime that have been proposed in the computer science literature and that can be grouped into these categories:

- **Increase the effort of crime.** Mitigations here include deploying firewalls and setting up automated updates for software installed on computers.
- **Increase the risk of crime.** Mitigations here include reducing payment anonymity (e.g., requesting an ID when someone cashes money from Western Union).
- **Reduce rewards.** Mitigations here include blocking supicious payments or parcels, or penalising malicious search results.
- **Reduce provocations.** Examples here include applying peer pressure to rogue ISPs and banks.
- **Remove excuses.** Typical mitigations in this category include running education campaigns or setting up automated redirects to divert victims who would have viewed malicious content, explain to them what happened and urge them to secure their systems.

An interesting aspect of the situational crime prevention framework is that it identifies, for each mitigation, the *implementation issues* that arise when putting the mitigation in place [124]. In the case of cybercrime, the two implementation issues that are most relevant are *adaptation* and *displacement*.

Adaptation embodies the fact that criminals will actively attempt to circumvent any mitigation by making their operation stealthier or more sophisticated. This is a typical arms race that can be observed in computer security research. When researchers started compiling blacklists of IP addresses known to belong to C&C servers, criminals reacted by developing Fast Flux. When making payments through traditional means became more difficult due to increased vetting, criminals moved on to cryptocurrencies. Considering adaptation is important when designing mitigations against cybercrime. In particular, effective mitigations are those which the criminal cannot easily react to, or where adaptation comes at a financial price (e.g., a reduction in revenue).

Displacement represents the fact that once mitigations are put in place, criminals can simply move their operations elsewhere. While in the physical world how far criminals can travel is dictated by practical constraints, on the Internet moving from one 'place' to another is virtually free. Examples of displacement include criminals starting to register DNS domains with another registrar after their preferred one increased the domain price to curb misuse [126], or a multitude of drug markets opening to fill the gap left by Silk Road's takedown [37]. Displacement effects are important when planning action against cybercrime. Generally speaking, a mitigation should make it difficult for criminals to move elsewhere. Conversely, a mitigating action that simply displaces a cybercriminal operation without affecting its effectiveness is probably not worth pursuing.

Researchers have applied Situational Crime Prevention to a number of computer crimes, including organisational data breaches [127] and the mitigation of software vulnerabilities [128]. Following the discussion in this section, however, this framework could be applied to any criminal activity that happens online.

Crime scripting. Another useful technique that can aid the analysis of malicious activities on the Internet from the field of criminology is crime scripting [129]. As part of this technique, researchers extrapolate the sequence of steps performed by an adversary to commit their offences. For example, in a romance scam, fraudsters create a fake account on a dating profile, the identify a suitable victim, go through a grooming phase, followed by the actual fraud when the scammer asks their victim for money. Dissecting the various steps of an offence can be useful to better understand it and to identify potential interventions. Crime scripting is somewhat related to kill chains, although the two techniques were developed in completely independent areas.

Modelling the underground economy as a flow of capital

As discussed in Section 1, many malicious operations are performed by criminals with the goal of making money from their victims. For this reason, following the flow of money is a useful way to better understand malicious operations, and in particular identify bottlenecks that could be leveraged to develop mitigations against them and stop criminals [108, 130].

Thomas et al. presented a model that is designed to keep track of a money flow within a cybercriminal operation [131]. As part of this model, they introduced two elements that are needed for a cybercrime operation to run: *profit centres*, through which victims transfer new capital into the criminal operation, and *support centres*, which can facilitate the criminal operation by providing several services for a fee. Money is introduced into the ecosystem through profit centres, and is then consumed by the various actors involved in it, who provide tools and services for each other. As an example, in an email spam operation, the profit centre would be victims purchasing counterfeit pharmaceuticals from an affiliate programme, while all the services needed by the spammers to operate (e.g., bulletproof hosting providers to host the C&C servers, pay-per-install services to deliver the malware, content generation services to create the spam content) are support centres. This model provides an interesting conceptualisation of how money flows into the cybercriminal ecosystem and how wealth is divided between the different actors there. By cross-referencing it with real world data, it can also help to form an idea of the profit that each criminal is making, and of the revenue of the operation.

Another interesting aspect of tracing the cash flow of cybercriminal operations is that at some point the criminals will want to cash out, which will be done using traditional payment methods (see Section 2). Since these interactions happen in the physical world, it is easier for law enforcement to trace them and potentially apprehend the criminals [130].

Attack attribution

When talking about malicious activities, attribution is important. Law enforcement is interested in understanding what criminals are behind a certain operation, and in particular attributing apparently unrelated cybercriminal operations to the same actors could help build a legal case against them. In similar fashion, governments are interested in identifying the culprits behind the attacks that they receive. In particular, they are interested in finding which nation states (i.e., countries) are behind these attacks.

Attribution, however, is a contemptuous topic in cyberspace. As we discussed previously, the concept of ‘place’ is relative for computer attacks, and attackers can easily route their network connections through proxies or compromised machines in third countries, thus hiding their actual location. It is reasonable to assume that the same actors will follow a similar *modus operandi* in their attacks, and in particular will use the same software exploits to break into their victims’ systems. These exploits could be used to identify state-sponsored groups or other attackers. Unfortunately, this approach has two

main drawbacks. The first is that the commodisation of cybercrime services has enabled attackers to use exploit kits, which contain a large number of exploits and, therefore, increase the likelihood of an attack happening. While advantageous for attackers, this trend means that the exploits used become a less significant signal for identifying attackers, especially those who do not have the sophistication to exploit vulnerabilities in house (e.g., cyber-enabled cybercriminals). The exception to this trend is state-sponsored actors, who unlike traditional criminals usually have very specific targets. For this reason, they can tailor their attacks more carefully, and even develop new exploits to hit a specific victim. Most importantly, they often develop exploits for vulnerabilities that are not publicly known, also known as *zero days attacks* [78]. Being unique to an actor, they could be used to identify who is behind a specific attack. An issue here is that, once an exploit is used, it could be intercepted by the victim (or anyone on the network) and later used against another target affected by the same vulnerability. This would actively mislead attribution. Recent leaks have shown that the CIA has been actively collecting exploits used by other nation states and adding them to their arsenal, thus allowing them to make it look like another country was behind any given computer attack.[7]

Rid et al. proposed a framework to systematise the attribution efforts of cyberattacks [132]. Within this framework, they identified three layers of analysis that are needed to correctly perform attribution: tactical, operational and strategic. The tactical component consists of understanding the technical aspects that composed the attack (the *how*), the operational component consists of understanding the attack's high-level characteristics architecture and the type of attacker that we are dealing with (the *what*), while the strategic component deals with understanding the motivation behind the attack (the *why*).

While this framework was developed with state-sponsored attacks in mind, it could be used to attribute other types of malicious activity. For example, to attribute an online hate attack orchestrated by 4chan's Politically Incorrect Board, [19] one could trace the hate messages reaching the victim (*how*), observe the personal information of the victim on the board (*what*) and analyse the discussion about the victim to understand the motivation behind the attack (*why*).

CONCLUSION

In this document, we presented an overview of the adversarial behaviours that exist on the Internet at the time of writing. We surveyed various types of malicious operations, depending on the attacker's motivations and capabilities, and analysed the components that are required to set up successful malicious operations. Finally, we described a number of modelling techniques from a variety of fields (computer science, criminology, war studies) that can help researchers and practitioners to better model these operations. We argued that having good models is of fundamental importance to developing effective mitigations that are difficult to circumvent.

REFERENCES

[1] M. Kjaerland, "A classification of computer security incidents based on reported attack data," *Journal of Investigative Psychology and Offender Profiling*, vol. 2, no. 2, pp. 105–120, 2005.

[2] R. Leukfeldt, E. Kleemans, and W. Stol, "A typology of cybercriminal networks: from low-tech all-rounders to high-tech specialists," *Crime, Law, and Social Change*, pp. 21–37, 2017.

[3] K. Thomas, D. Huang, D. Wang, E. Bursztein, C. Grier, T. J. Holt, C. Kruegel, D. McCoy, S. Savage, and G. Vigna, "Framing dependencies introduced by underground commoditization," in *Workshop on the Economics of Information Security*, 2015.

[4] M. McGuire and S. Dowling, "Cyber crime: A review of the evidence," *Summary of key findings and implications. Home Office Research report*, vol. 75, 2013.

[5] J. Clough, "Cybercrime principles," 2010.

[7] https://en.wikipedia.org/wiki/Vault_7

[6] B. Stone-Gross, T. Holz, G. Stringhini, and G. Vigna, "The underground economy of spam: A botmaster's perspective of coordinating large-scale spam campaigns." *LEET*, vol. 11, pp. 4–4, 2011.

[7] C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@ spam: the underground on 140 characters or less," in *Proceedings of the 17th ACM conference on Computer and communications security*. ACM, 2010, pp. 27–37.

[8] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in *Proceedings of the 26th annual computer security applications conference*. ACM, 2010, pp. 1–9.

[9] E. Bursztein, B. Benko, D. Margolis, T. Pietraszek, A. Archer, A. Aquino, A. Pitsillidis, and S. Savage, "Handcrafted fraud and extortion: Manual account hijacking in the wild," in *ACM SIGCOMM Conference on Internet Measurement Conference (IMC)*. ACM, 2014, pp. 347–358.

[10] J. Clough, "A world of difference: The budapest convention of cybercrime and the challenges of harmonisation," *Monash UL Rev.*, vol. 40, p. 698, 2014.

[11] C. Kershaw, S. Nicholas, and A. Walker, "Crime in england and wales 2007/08: Findings from the british crime survey and police recorded crime," *Home Office Statistical Bulletin*, vol. 7, no. 08, 2008.

[12] N. E. Willard, *Cyberbullying and cyberthreats: Responding to the challenge of online social aggression, threats, and distress*. Research Press, 2007.

[13] Y. Jewkes, "Public policing and the internet," 2010.

[14] D. Chatzakou, N. Kourtellis, J. Blackburn, E. De Cristofaro, G. Stringhini, and A. Vakali, "Mean birds: Detecting aggression and bullying on twitter," in *Proceedings of the 2017 ACM on web science conference*. ACM, 2017, pp. 13–22.

[15] R. Slonje and P. K. Smith, "Cyberbullying: Another main type of bullying?" *Scandinavian journal of psychology*, vol. 49, no. 2, pp. 147–154, 2008.

[16] R. M. Kowalski and S. P. Limber, "Electronic bullying among middle school students," *Journal of adolescent health*, vol. 41, no. 6, pp. S22–S30, 2007.

[17] J. Suler, "The online disinhibition effect," *Cyberpsychology & behavior*, vol. 7, no. 3, pp. 321–326, 2004.

[18] A. N. Joinson, "Disinhibition and the internet," in *Psychology and the Internet*. Elsevier, 2007, pp. 75–92.

[19] G. E. Hine, J. Onaolapo, E. De Cristofaro, N. Kourtellis, I. Leontiadis, R. Samaras, G. Stringhini, and J. Blackburn, "Kek, cucks, and god emperor trump: A measurement study of 4chan's politically incorrect forum and its effects on the web," in *International Conference on Web and Social Media (ICWSM)*. AAAI, 2017.

[20] P. Snyder, P. Doerfler, C. Kanich, and D. McCoy, "Fifteen minutes of unwanted fame: Detecting and characterizing doxing," in *Proceedings of the 2017 Internet Measurement Conference*. ACM, 2017, pp. 432–444.

[21] D. Chatzakou, N. Kourtellis, J. Blackburn, E. De Cristofaro, G. Stringhini, and A. Vakali, "Hate is not binary: Studying abusive behavior of# gamergate on twitter," in *Proceedings of the 28th ACM conference on hypertext and social media*. ACM, 2017, pp. 65–74.

[22] D. Freed, J. Palmer, D. E. Minchala, K. Levy, T. Ristenpart, and N. Dell, "Digital technologies and intimate partner violence: A qualitative analysis with multiple stakeholders," *Proceedings of the ACM on Human-Computer Interaction*, vol. 1, no. CSCW, p. 46, 2017.

[23] D. Freed, J. Palmer, D. Minchala, K. Levy, T. Ristenpart, and N. Dell, ""a stalker's paradise": How intimate partner abusers exploit technology," in *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2018.

[24] A. Mishra and D. Mishra, "Cyber stalking: a challenge for web security," in *Cyber Warfare and Cyber Terrorism*. IGI Global, 2007, pp. 216–226.

[25] B. Wittes, C. Poplin, Q. Jurecic, and C. Spera, "Sextortion: Cybersecurity, teenagers, and remote sexual assault," *Center for Technology at Brookings. https://www. brookings. edu/wp-*

content/uploads/2016/05/sextortion1-1. pdf. Accessed, vol. 16, 2016.

[26] H. Whittle, C. Hamilton-Giachritsis, A. Beech, and G. Collings, "A review of online grooming: Characteristics and concerns," *Aggression and violent behavior*, vol. 18, no. 1, pp. 62–70, 2013.

[27] J. Wolak, D. Finkelhor, and K. Mitchell, "Is talking online to unknown people always risky? distinguishing online interaction styles in a national sample of youth internet users," *CyberPsychology & Behavior*, vol. 11, no. 3, pp. 340–343, 2008.

[28] J. Huang, G. Stringhini, and P. Yong, "Quit playing games with my heart: Understanding online dating scams," in *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*. Springer, 2015, pp. 216–236.

[29] M. Dittus, J. Wright, and M. Graham, "Platform criminalism: The'last-mile'geography of the darknet market supply chain," in *Proceedings of the 2018 World Wide Web Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2018, pp. 277–286.

[30] M. T. Whitty and T. Buchanan, "The online romance scam: A serious cybercrime," *CyberPsychology, Behavior, and Social Networking*, vol. 15, no. 3, pp. 181–183, 2012.

[31] H. Glickman, "The nigerian "419" advance fee scams: prank or peril?" *Canadian Journal of African Studies/La Revue Canadienne Des Études Africaines*, vol. 39, no. 3, pp. 460–489, 2005.

[32] Y. Park, D. McCoy, and E. Shi, "Understanding craigslist rental scams," in *International Conference on Financial Cryptography and Data Security*. Springer, 2016, pp. 3–21.

[33] M. Edwards, G. Suarez-Tangil, C. Peersman, G. Stringhini, A. Rashid, and M. Whitty, "The geography of online dating fraud," in *Workshop on Technology and Consumer Protection*. IEEE, 2018.

[34] C. Herley, "Why do nigerian scammers say they are from nigeria?" in *Workshop on the Economics of Information Security (WEIS)*, 2012.

[35] P. Syverson, R. Dingledine, and N. Mathewson, "Tor: The secondgeneration onion router," in *Usenix Security*, 2004.

[36] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[37] N. Christin, "Traveling the silk road: A measurement analysis of a large anonymous online marketplace," in *international Conference on World Wide Web (WWW)*. ACM, 2013, pp. 213–224.

[38] K. Soska and N. Christin, "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem." in *USENIX Security Symposium*, 2015, pp. 33–48.

[39] M. Abu Rajab, J. Zarfoss, F. Monrose, and A. Terzis, "A multifaceted approach to understanding the botnet phenomenon," in *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*. ACM, 2006, pp. 41–52.

[40] B. Krebs, *Spam nation: the inside story of organized cybercrime-from global epidemic to your front door*. Sourcebooks, Inc., 2014.

[41] S. Hinde, "Spam: the evolution of a nuisance," *Computers & Security*, vol. 22, no. 6, pp. 474–478, 2003.

[42] B. S. McWilliams, *Spam Kings: The Real Story behind the High-Rolling Hucksters Pushing Porn, Pills, and%*@)# Enlargements*. " O'Reilly Media, Inc.", 2014.

[43] G. Stringhini, O. Hohlfeld, C. Kruegel, and G. Vigna, "The harvester, the botmaster, and the spammer: on the relations between the different actors in the spam landscape," in *Proceedings of the 9th ACM symposium on Information, computer and communications security*. ACM, 2014, pp. 353–364.

[44] D. McCoy, A. Pitsillidis, G. Jordan, N. Weaver, C. Kreibich, B. Krebs, G. M. Voelker, S. Savage, and K. Levchenko, "Pharmaleaks: Understanding the business of online pharmaceutical affiliate programs," in *USENIX Security Symposium*. USENIX Association, 2012, pp. 1–1.

[45] D. Samosseiko, "The partnerka—what is it, and why should you care," in *Proc. of Virus Bulletin Conference*, 2009.

[46] N. Spirin and J. Han, "Survey on web spam detection: principles and algorithms," *Acm Sigkdd Explorations Newsletter*, vol. 13, no. 2, pp. 50–64, 2012.

[47] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage, "Spamalytics: An empirical analysis of spam marketing conversion," in *ACM SIGSAC Conference on Computer and Ccommunications Security (CCS)*. ACM, 2008, pp. 3–14.

[48] R. Dhamija, J. D. Tygar, and M. Hearst, "Why phishing works," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006, pp. 581–590.

[49] X. Han, N. Kheir, and D. Balzarotti, "Phisheye: Live monitoring of sandboxed phishing kits," in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM, 2016, pp. 1402–1413.

[50] T. Moore and R. Clayton, "Examining the impact of website take-down on phishing," in *Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit*. ACM, 2007, pp. 1–13.

[51] ——, "Evil searching: Compromise and recompromise of internet hosts for phishing," in *International Conference on Financial Cryptography and Data Security*. Springer, 2009, pp. 256–272.

[52] J. Onaolapo, E. Mariconti, and G. Stringhini, "What happens after you are pwnd: Understanding the use of leaked webmail credentials in the wild," in *Proceedings of the 2016 Internet Measurement Conference*. ACM, 2016, pp. 65–79.

[53] H. Binsalleeh, T. Ormerod, A. Boukhtouta, P. Sinha, A. Youssef, M. Debbabi, and L. Wang, "On the analysis of the zeus botnet crimeware toolkit," in *Annual International Conference on Privacy Security and Trust (PST)*. IEEE, 2010, pp. 31–38.

[54] B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna, "Your botnet is my botnet: analysis of a botnet takeover," in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM, 2009, pp. 635–647.

[55] A. Haslebacher, J. Onaolapo, and G. Stringhini, "All your cards are belong to us: Understanding online carding forums," in *Electronic Crime Research (eCrime), 2017 APWG Symposium on*. IEEE, 2017, pp. 41–51.

[56] N. Scaife, C. Peeters, and P. Traynor, "Fear the reaper: characterization and fast detection of card skimmers," in *27th {USENIX} Security Symposium ({USENIX} Security 18)*, 2018, pp. 1–14.

[57] P. Pearce, V. Dave, C. Grier, K. Levchenko, S. Guha, D. McCoy, V. Paxson, S. Savage, and G. M. Voelker, "Characterizing large-scale click fraud in zeroaccess," in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM, 2014, pp. 141–152.

[58] D. Y. Huang, H. Dharmdasani, S. Meiklejohn, V. Dave, C. Grier, D. McCoy, S. Savage, N. Weaver, A. C. Snoeren, and K. Levchenko, "Botcoin: Monetizing stolen cycles." in *Network and Distributed Systems Symposium (NDSS)*. Intenet Society, 2014.

[59] S. Pastrana and G. Suarez-Tangil, "A first look at the crypto-mining malware ecosystem: A decade of unrestricted wealth," *arXiv preprint arXiv:1901.00846*, 2019.

[60] R. K. Konoth, E. Vineti, V. Moonsamy, M. Lindorfer, C. Kruegel, H. Bos, and G. Vigna, "Minesweeper: An in-depth look into drive-by cryptocurrency mining and its defense," in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM, 2018, pp. 1714–1730.

[61] J. Rüth, T. Zimmermann, K. Wolsing, and O. Hohlfeld, "Digging into browser-based crypto mining," in *ACM SIGCOMM Internet Measurement Conference (IMC)*. ACM, 2018, pp. 70–76.

[62] G. O'Gorman and G. McDonald, *Ransomware: A growing menace*. Symantec Corporation, 2012.

[63] A. L. Young and M. Yung, "Cryptovirology: The birth, neglect, and explosion of ransomware," *Communications of the ACM*, vol. 60, no. 7, pp. 24–26, 2017.

[64] D. Y. Huang, M. M. Aliapoulios, V. G. Li, L. Invernizzi, E. Bursztein, K. McRoberts, J. Levin, K. Levchenko, A. C. Snoeren, and D. McCoy, "Tracking ransomware end-to-end," in *IEEE Sym-*

posium on Security and Privacy. IEEE, 2018, pp. 618–631.
[65] A. Kharraz, W. Robertson, D. Balzarotti, L. Bilge, and E. Kirda, "Cutting the gordian knot: A look under the hood of ransomware attacks," in *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA)*. Springer, 2015, pp. 3–24.
[66] C. Rossow, "Amplification hell: Revisiting network protocols for ddos abuse." in *NDSS*, 2014.
[67] M. Karami, Y. Park, and D. McCoy, "Stress testing the booters: Understanding and undermining the business of ddos services," in *Proceedings of the 25th International Conference on World Wide Web*. ACM, 2016, pp. 1033–1043.
[68] T. Jordan and P. Taylor, *Hacktivism and cyberwars: Rebels with a cause?* Routledge, 2004.
[69] B. Brevini, A. Hintz, and P. McCurdy, *Beyond WikiLeaks: implications for the future of communications, journalism and society*. Springer, 2013.
[70] M. Conway, "Cyberterrorism: Hype and reality," 2007.
[71] T. J. Holt, M. Stonhouse, J. Freilich, and S. M. Chermak, "Examining ideologically motivated cyberattacks performed by far-left groups," *Terrorism and Political Violence*, pp. 1–22, 2019.
[72] S. Milan, *Routledge Companion to Alternative and Community Media*. London: Routledge, 2015, ch. Hactivism as a radical media practice.
[73] L. Goode, "Anonymous and the political ethos of hacktivism," *Popular Communication*, vol. 13, no. 1, pp. 74–86, 2015.
[74] G. Greenwald, *No place to hide: Edward Snowden, the NSA, and the US surveillance state*. Macmillan, 2014.
[75] H.-j. Woo, Y. Kim, and J. Dominick, "Hackers: Militants or merry pranksters? a content analysis of defaced web pages," *Media Psychology*, vol. 6, no. 1, pp. 63–82, 2004.
[76] A. K. Al-Rawi, "Cyber warriors in the middle east: The case of the syrian electronic army," *Public Relations Review*, vol. 40, no. 3, pp. 420–428, 2014.
[77] D. Maimon, A. Fukuda, S. Hinton, O. Babko-Malaya, and R. Cathey, "On the relevance of social media platforms in predicting the volume and patterns of web defacement attacks," in *Big Data (Big Data), 2017 IEEE International Conference on*. IEEE, 2017, pp. 4668–4673.
[78] L. Bilge and T. Dumitras, "Before we knew it: an empirical study of zero-day attacks in the real world," in *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM, 2012, pp. 833–844.
[79] M. B. Line, A. Zand, G. Stringhini, and R. Kemmerer, "Targeted attacks against industrial control systems: Is the power industry prepared?" in *Proceedings of the 2nd Workshop on Smart Energy Grid Security*. ACM, 2014, pp. 13–22.
[80] K. Zetter, *Countdown to Zero Day: Stuxnet and the launch of the world's first digital weapon*. Broadway books, 2014.
[81] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security & Privacy*, vol. 9, no. 3, pp. 49–51, 2011.
[82] S. Le Blond, A. Uritesc, C. Gilbert, Z. L. Chua, P. Saxena, and E. Kirda, "A look at targeted attacks through the lense of an ngo." in *USENIX Security Symposium*. USENIX Association, 2014, pp. 543–558.
[83] D. Alperovitch *et al.*, *Revealed: operation shady RAT*. McAfee, 2011, vol. 3.
[84] P. Knight, "Iloveyou: Viruses, paranoia, and the environment of risk," *The Sociological Review*, vol. 48, no. S2, pp. 17–30, 2000.
[85] K. Krombholz, H. Hobel, M. Huber, and E. Weippl, "Advanced social engineering attacks," *Journal of Information Security and applications*, vol. 22, pp. 113–122, 2015.
[86] D. Y. Wang, M. Der, M. Karami, L. Saul, D. McCoy, S. Savage, and G. M. Voelker, "Search+ seizure: The effectiveness of interventions on seo campaigns," in *ACM SIGCOMM Internet Measurement Conference (IMC)*. ACM, 2014, pp. 359–372.
[87] N. Provos, D. McNamee, P. Mavrommatis, K. Wang, N. Modadugu *et al.*, "The ghost in the browser: Analysis of web-based malware." *HotBots*, vol. 7, pp. 4–4, 2007.
[88] M. Cova, C. Kruegel, and G. Vigna, "Detection and analysis of drive-by-download attacks and

malicious javascript code," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 281–290.

[89] A. Zarras, A. Kapravelos, G. Stringhini, T. Holz, C. Kruegel, and G. Vigna, "The dark alleys of madison avenue: Understanding malicious advertisements," in *Proceedings of the 2014 Conference on Internet Measurement Conference*. ACM, 2014, pp. 373–380.

[90] B. Stone-Gross, C. Kruegel, K. Almeroth, A. Moser, and E. Kirda, "FIRE: Finding rogue networks," in *Annual Computer Security Applications Conference (ACSAC)*. ACM, 2009, pp. 231–240.

[91] M. Konte, R. Perdisci, and N. Feamster, "Aswatch: An as reputation system to expose bulletproof hosting ases," in *ACM SIGCOMM Computer Communication Review*, vol. 45, no. 4. ACM, 2015, pp. 625–638.

[92] C. Y. Cho, J. Caballero, C. Grier, V. Paxson, and D. Song, "Insights from the inside: A view of botnet management from infiltration." *LEET*, vol. 10, pp. 1–1, 2010.

[93] C. Kreibich, C. Kanich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage, "Spamcraft: An inside look at spam campaign orchestration." in *LEET*. USENIX Association, 2009.

[94] T. Holz, C. Gorecki, K. Rieck, and F. C. Freiling, "Measuring and detecting fast-flux service networks." in *NDSS*, 2008.

[95] S. Yadav, A. K. K. Reddy, A. N. Reddy, and S. Ranjan, "Detecting algorithmically generated domain-flux attacks with dns traffic analysis," *IEEE/Acm Transactions on Networking*, vol. 20, no. 5, pp. 1663–1677, 2012.

[96] J. Lusthaus, *Industry of Anonymity: Inside the Business of Cybercrime*. Harvard University Press, 2018.

[97] C. Grier, L. Ballard, J. Caballero, N. Chachra, C. J. Dietrich, K. Levchenko, P. Mavrommatis, D. McCoy, A. Nappa, A. Pitsillidis *et al.*, "Manufacturing compromise: the emergence of exploit-as-a-service," in *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM, 2012, pp. 821–832.

[98] J. Iedemska, G. Stringhini, R. Kemmerer, C. Kruegel, and G. Vigna, "The tricks of the trade: What makes spam campaigns successful?" in *Security and Privacy Workshops (SPW), 2014 IEEE*. IEEE, 2014, pp. 77–83.

[99] J. Caballero, C. Grier, C. Kreibich, and V. Paxson, "Measuring pay-per-install: the commoditization of malware distribution." in *USENIX Security Symposium*. USENIX Association, 2011, pp. 13–13.

[100] M. Motoyama, K. Levchenko, C. Kanich, D. McCoy, G. M. Voelker, and S. Savage, "Re: Captchas-understanding captcha-solving services in an economic context." in *USENIX Security Symposium*, vol. 10. USENIX Association, 2010, p. 3.

[101] K. Thomas, D. McCoy, C. Grier, A. Kolcz, and V. Paxson, "Trafficking fraudulent accounts: The role of the underground market in twitter spam and abuse." in *USENIX Security Symposium*. USENIX Association, 2013, pp. 195–210.

[102] E. De Cristofaro, A. Friedman, G. Jourjon, M. A. Kaafar, and M. Z. Shafiq, "Paying for likes?: Understanding facebook like fraud using honeypots," in *ACM SIGCOMM Internet Measurement Conference (IMC)*. ACM, 2014, pp. 129–136.

[103] G. Stringhini, G. Wang, M. Egele, C. Kruegel, G. Vigna, H. Zheng, and B. Y. Zhao, "Follow the green: growth and dynamics in twitter follower markets," in *ACM SIGCOMM Internet Measurement Conference (IMC)*. ACM, 2013, pp. 163–176.

[104] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker, "Dirty jobs: The role of freelance labor in web service abuse," in *Proceedings of the 20th USENIX conference on Security*. USENIX Association, 2011, pp. 14–14.

[105] D. Florêncio and C. Herley, "Phishing and money mules," in *International Workshop on Information Forensics and Security (WIFS)*. IEEE, 2010, pp. 1–5.

[106] S. Hao, K. Borgolte, N. Nikiforakis, G. Stringhini, M. Egele, M. Eubanks, B. Krebs, and G. Vi-

gna, “Drops for stuff: An analysis of reshipping mule scams,” in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*. ACM, 2015, pp. 1081–1092.

[107] B. Stone-Gross, R. Abman, R. A. Kemmerer, C. Kruegel, D. G. Steigerwald, and G. Vigna, “The underground economy of fake antivirus software,” in *Workshop on the Economics of Information Security and Privacy (WEIS)*. Springer, 2013, pp. 55–78.

[108] K. Levchenko, A. Pitsillidis, N. Chachra, B. Enright, M. Félegyházi, C. Grier, T. Halvorson, C. Kanich, C. Kreibich, H. Liu *et al.*, “Click trajectories: End-to-end analysis of the spam value chain,” in *IEEE Symposium on Security and Privacy*. IEEE, 2011, pp. 431–446.

[109] D. Y. Wang, M. Der, M. Karami, L. Saul, D. McCoy, S. Savage, and G. M. Voelker, “Search+ seizure: The effectiveness of interventions on seo campaigns,” in *Proceedings of the 2014 Conference on Internet Measurement Conference*. ACM, 2014, pp. 359–372.

[110] R. Anderson, C. Barton, R. Böhme, R. Clayton, M. J. Van Eeten, M. Levi, T. Moore, and S. Savage, “Measuring the cost of cybercrime,” in *Workshop on the economics of information security and privacy (WEIS)*. Springer, 2013, pp. 265–300.

[111] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker, and S. Savage, “A fistful of bitcoins: characterizing payments among men with no names,” in *ACM SIGCOMM Internet Measurement Conference (IMC)*. ACM, 2013, pp. 127–140.

[112] T. Moore and N. Christin, “Beware the middleman: Empirical analysis of bitcoin-exchange risk,” in *International Conference on Financial Cryptography and Data Security*. Springer, 2013, pp. 25–33.

[113] K. Moeller, R. Munksgaard, and J. Demant, “Flow my fe the vendor said: Exploring violent and fraudulent resource exchanges on cryptomarkets for illicit drugs,” *American Behavioral Scientist*, vol. 61, no. 11, pp. 1427–1450, 2017.

[114] B. Schneier, “Attack trees,” *Dr. Dobb’s journal*, vol. 24, no. 12, pp. 21–29, 1999.

[115] E. M. Hutchins, M. J. Cloppert, and R. M. Amin, “Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains,” *Leading Issues in Information Warfare & Security Research*, vol. 1, no. 1, p. 80, 2011.

[116] B. Farinholt, M. Rezaeirad, P. Pearce, H. Dharmdasani, H. Yin, S. Le Blond, D. McCoy, and K. Levchenko, “To catch a ratter: Monitoring the behavior of amateur darkcomet rat operators in the wild,” in *2017 38th IEEE Symposium on Security and Privacy (SP)*. Ieee, 2017, pp. 770–787.

[117] B. Cheswick, “An evening with berferd in which a cracker is lured, endured, and studied,” in *Proc. Winter USENIX Conference, San Francisco*, 1992, pp. 20–24.

[118] G. Gu, P. A. Porras, V. Yegneswaran, M. W. Fong, and W. Lee, “Bothunter: Detecting malware infection through ids-driven dialog correlation.” in *USENIX Security Symposium*, vol. 7, 2007, pp. 1–16.

[119] D. Moore, C. Shannon *et al.*, “Code-red: a case study on the spread and victims of an internet worm,” in *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*. ACM, 2002, pp. 273–284.

[120] A. Dwyer, “The nhs cyber-attack: A look at the complex environmental conditions of wannacry,” *RAD Magazine*, vol. 44, 2018.

[121] P. J. Brantingham, P. L. Brantingham *et al.*, *Environmental criminology*. Sage Publications Beverly Hills, CA, 1981.

[122] L. E. Cohen and M. Felson, “Social change and crime rate trends: A routine activity approach (1979),” in *Classics in Environmental Criminology*. CRC Press, 2016, pp. 203–232.

[123] R. V. Clarke and D. B. Cornish, “Modeling offenders’ decisions: A framework for research and policy,” *Crime and justice*, vol. 6, pp. 147–185, 1985.

[124] R. V. G. Clarke, *Situational crime prevention*. Criminal Justice Press Monsey, NY, 1997.

[125] M. R. Albert, “E-buyer beware: why online auction fraud should be regulated,” *American Business Law Journal*, vol. 39, no. 4, pp. 575–644, 2002.

[126] H. Liu, K. Levchenko, M. Félegyházi, C. Kreibich, G. Maier, G. M. Voelker, and S. Savage, “On

the effects of registrar-level intervention." in *LEET*, 2011.

[127] D. N. Khey and V. A. Sainato, "Examining the correlates and spatial distribution of organizational data breaches in the united states," *Security Journal*, vol. 26, no. 4, pp. 367–382, 2013.

[128] S. Hinduja and B. Kooi, "Curtailing cyber and information security vulnerabilities through situational crime prevention," *Security journal*, vol. 26, no. 4, pp. 383–402, 2013.

[129] D. B. Cornish, "The procedural analysis of offending and its relevance for situational prevention," *Crime prevention studies*, vol. 3, pp. 151–196, 1994.

[130] R. Wortley, A. Sidebottom, N. Tilley, and G. Laycock, *Routledge Handbook of Crime Science*. Routledge, 2018.

[131] D. Huang, K. Thomas, C. Grier, D. Wang, E. Burztein, T. Holt, C. Kruegel, D. McCoy, S. Savage, and G. Vigna, "Framing dependencies introduced by underground commoditization," in *Workshop on the Economics of Information Security*, 2015.

[132] T. Rid and B. Buchanan, "Attributing cyber attacks," *Journal of Strategic Studies*, vol. 38, no. 1-2, pp. 4–37, 2015.