Measuring cybercrime – both its occurrence and cost – is hard. While collecting data is often a time-consuming and arduous task even under the best of circumstances, one must overcome a series of additional impediments in order to measure most cybercrimes effectively.

First, finding suitable data sources can be difficult. Victims may be reluctant to discuss incidents, as they only see downside risk to sharing information with others. At best, disclosing incidents may attract some unflattering press reports, while at worst it can attract lawsuits, particularly for publicly traded companies. In the U.S. for example, the Federal Trade Commission has authority under Section 5 of the Federal Trade Commission Act (15 USC 45) to take action against unfair or deceptive acts and practices that affect commerce. Since 2005, the FTC has occasionally charged companies with acting ‘unfairly’ by failing to adopt reasonable information security practices. The FTC typically relies on public reporting of incidents in order to direct its investigations, which makes companies very reticent to discuss their own failings.

Given a weak incentive to voluntarily disclose incidents, another option is to mandate disclosure through regulation. This has been tried most prominently for breaches of personal information, which must now be reported in most U.S. states and in many countries. Breach-notification laws have been a boon to security economics research. For more information, see Chapter ??.

If the victims do not want to share and absent regulations to compel them, researchers can try to enlist the support of disinterested third parties who might observe evidence of incidents. For example, ISPs already observe every domain name that its customers attempt to visit. Computers query their respective ISPs' DNS servers, in order to convert domain names (e.g., example.com) into the corresponding IP addresses (e.g., 1.2.3.4) that deliver web content. Because many cyber criminals register domain names for purely malicious purposes (e.g., to control computers in a botnet), one can estimate the prevalence of malicious web traffic at an ISP by observing the logs of its DNS servers. So-called passive DNS has been used to estimate the prevalence of malware and the size of botnets [3, 4].

Another popular approach to gathering data on cybercrime is to obtain a copy of the records maintained by criminals. For example, one group of researchers managed to get access to several internal databases run by three different criminal gangs peddling fake antivirus software from 2008–2010 [21].
They found that the criminal groups had kept detailed records of conversion rates from installations to sales, along with the prices paid. The authors estimate that these three groups collectively earned $97 million per year. Similarly, researchers investigating the online sale of illicit pharmaceuticals have analyzed a database of leaked transaction and financial record data from leading pharmaceutical affiliate programs [15].

Despite the greater potential for cooperation with third parties over sharing incident data, many countervailing pressures against sharing remain. For instance, many infrastructure operators are reluctant to share data, claiming concerns for customer privacy. Regulatory concerns also remain, since intermediaries such as ISPs are under increasing pressure to take a more active role in combating online crime.

When security data is not willingly shared, two methods of empirical inquiry remain: direct observation and sampling. Direct observation can take many forms, and is typically tailored to the particular form of wickedness under observation. Several researchers have passively monitored underground markets. For example, researchers have monitored IRC chat rooms for public advertisements of stolen credentials and illicit services such as spamming and machine compromise [11]. These markets have evolved rapidly, reflecting the maturation of the underground economy. In just a few short years, services offered have grown more sophisticated, and have been shown to include ‘pay-per-install’ services that eliminate much of the technical know-how required to perpetrate cybercrime [5].

Some researchers have studied the key aspects of the criminal infrastructure directly. Researchers have studied botnets by reverse-engineering the communications that take place between criminal servers and the infected machines. For instance, researchers took over a portion of the spam-sending Storm botnet [13]. They manipulated the spam emails sent by machines in the botnet so that they linked to cloned websites under their control. They then observed how often people responded to spam, finding for example that just 0.0000081% of spam emails advertising illicit pharmaceuticals attracted customers. Other researchers briefly intercepted the command and control communications of a large botnet designed to steal end-user credentials. In just 10 days, they observed more than 180,000 infected computers participating in the botnet [20].

Another class of direct observation is to simply inspect all known computers for malicious activity. For example, in conjunction with indexing used to update its search engine results, Google has deployed automated crawlers to visit websites masquerading as vulnerable browsers to attract infection by malicious servers [18]. The gathered information is then used to protect users by blocking referrals to the infected websites from Google’s search results. The authors also report summary statistics on the prevalence of malware on web servers, such as that 1.3% of incoming search queries to Google include at least one malicious result. Other researchers have similarly examined millions of tweets on Twitter and found that 8% of tweets have been flagged as malicious [12].

While these summary statistics are interesting, one limitation of comprehensive approaches is that it can be difficult to assess the impact of harm on users.
Just because a website has been hacked or an attacker tweets a malicious URL, it does not necessarily follow that someone will be harmed. A high-profile website that frequently appears in the top search results for popular search terms will cause far more harm than an obscure one that only appears deep in the results for less popular terms.

To draw a closer connection to harm, researchers have collected more limited data that corresponds more closely to user behavior. For example, one study inspected the search results of “trending” terms (i.e., those terms that rapidly rise in popularity) and persistently popular terms to measure how often the links point to compromised web servers [17]. They found that miscreants were more successful infecting trending terms than persistently popular terms, and estimated revenues available from distributing malware rivaled those that could be obtained by running low-quality websites automatically populated with advertisements.

In another example, researchers studied the prevalence of “typosquatting” – the intentional registration of web domains that are close misspellings of popular websites [16]. They generated an exhaustive list of one- and two-letter misspellings of top websites and then checked whether the website had been registered. They then crawled the websites in use and found that the vast majority monetized the traffic through advertising.

In another striking example of collecting data to measure harm, researchers studied the many facets of the email spam value chain [14]. Most email spam advertises counterfeit pharmaceutical, fashion and software products. The authors monitored several botnets sending out the spam, then followed the advertised URLs to track the websites selling the products and the associated infrastructure. They even made a few hundred purchases, finding that most spam-advertised products are served by a few payment processors.

We can draw several conclusions from the wide range of direct observational studies conducted. There appears to be a natural trade-off between comprehensiveness and precision when measuring online wickedness. Studies often tailor their measurement strategy to the behavior being investigated. One drawback of such focus is that we cannot readily extrapolate findings beyond each case study.

Even when the studies aim to be comprehensive and collect vast amounts of data, they may not be representative of all cybercrime. In particular, the types of criminal activity that can be easily observed by researchers across the globe might be less sophisticated. Consider the IRC chat rooms advertising underground services. These public marketplaces have been infiltrated by law enforcement investigators for several years. Serious criminals might plausibly find more secure means of communication, leaving only amateurs and the truly desperate left in the fora. If so, then it would be unwise to form a view of all cybercriminals based on what can be found there.

One corollary is that crimes that are inherently difficult to measure may go unexamined. The classic example is data exfiltration, where criminals gain unauthorized access to firm computers and copy sensitive data. Firms often do not detect that the event has occurred, since their own copy of the data remains
intact.

Given the detailed domain-specific knowledge required by direct observation along with its limitations, one might prefer to instead survey the victims of cybercrime instead. In theory, surveys allow researchers to measure the impact of cybercrime directly. Some surveys ask people and firms about their experiences with online crime. For example, the long-running CSI survey asks firms questions about their experiences with various attacks, the resulting losses, and their security investments [19]. Meanwhile, other surveys target consumers for their experiences and concerns with cybercrime. One example is the Eurobarometer survey that asks EU citizens questions such as “How often have you experienced or been a victim of the following situation: Identity theft (somebody stealing your personal data and impersonating you, e.g. shopping under your name)” [8].

These surveys highlight several of the difficulties all such security surveys face. First, definitions are loose and left open to interpretation. An “attack” comprises many categories of abuse. Without more specific language, survey respondents will rely on their own often inconsistent definitions. For example, the Eurobarometer definition of identity theft above may or may not be interpreted by its respondents to cover data breaches where personal information is lost but without causing any monetary harm. The fact that reasonable people can arrive at different interpretations can undermine the survey’s accuracy.

This definitional ambiguity problem is more acute for consumer-focused surveys than for surveys of security personnel at firms. In the latter case, questions can be crafted to include terms familiar to those with security knowledge, such as when the CSI survey asks about 21 different forms of attack, ranging from “malware infection” to “theft or unauthorized access to intellectual property”.

Most of the measurement challenges described above for direct observation remain a problem for surveys. Survey respondents can suffer from measurement error in two directions. They might underreport attacks that they do not observe to be attacks, and they can also misclassify benign events as attacks. Furthermore, translating experiences of crime and loss into monetary values is still hard. For example, the CSI survey routinely asks for firms to estimate their financial losses due to security incidents. In the 2011 survey, only 77 out of the 351 respondents (22%) even reported a figure for financial losses. Part of the explanation for such a poor response is that many firms are not able or willing to track and attribute financial losses due to cybercrime.

One potential issue for all surveys is sample bias, when the set of survey respondents does not accurately represent the population being studied. For instance, sample bias is a serious concern for the CSI survey. Their 2011 survey received just a 6.4% response rate. It is reasonable to wonder whether the type of people who do respond are different from those who do not. Indeed, the respondents tend to come from disproportionately large companies who already invest heavily in IT security.

Even after ensuring that the survey sample is truly random, challenges remain when the underlying distribution of what is being measured is skewed. Consider financial losses for cybercrime. The CSI survey could not arrive at a meaningful measure of average loss due to the extreme variation in reported
losses. Of the 77 reports of financial losses, there were two large outliers ($25 million and $20 million) but the average for the remaining 75 was $100,000. Including the two outliers would drive the average up by nearly a factor of 10, but discarding them from consideration completely is not right either since the figures are probably accurate reports of true losses.

A related issue is that for many cybercrimes victims are concentrated in a small portion of the overall population. Many cybercrimes only affect a very small fraction of Internet users. For example, one estimate suggests 0.4% of the Internet population falls for phishing attacks annually [9]. This can have two effects. First, finding victims from a truly random sample of the population may require sampling from a larger pool. But there is also a more subtle effect at play, as pointed out by Florêncio and Herley [10]. Victims of a rare crime may be more likely to respond to surveys, due to the higher salience of the survey topic. When this happens, the victimization rate will be inflated by a factor matching the relative response rate of victims to the population. In other words, if victims are twice as likely to respond to a survey than non-victims, the surveyed incident rate will be twice the true incidence.

Finally, for all forms of empirical research discussed here, one must be wary of perverse incentives to distort the quality of reports. On the one hand, there can be a strong incentive to hype security problems. Antivirus companies and security-services firms have an interest to portray all security threats as severe, in order to justify purchasing their services. One particularly egregious example came from Detica, a defense contractor hired by the UK government to estimate the overall cost of cybercrime in that country [7]. They arrived at an eye-wateringly large figure of £27 Billion lost per annum, but their methodology was later shown to be largely unfounded [2].

On the other hand, there can also be countervailing pressure to underreport the costs of cybercrime. For example, companies fearful of regulation and oversight might try to keep incidents from public view. Similarly, publicly-traded companies might fear that disclosing incidents could harm the stock price, even though past studies have shown only a limited short-term impact in most cases [6, 1].

References


