Concentrating Correctly on Cybercrime Concentration

Richard Clayton [1] Tyler Moore [2] Nicolas Christin [3]

Abstract

We review the cybercrime literature to draw attention to the many occasions on which authors have identified concentrations within criminal activity. We note that the existence of concentrations often leads authors to suggest that concentrations, or 'choke points' are amenable to effective intervention. We then discuss the reasons that concentrations are observed – it is often the result of the criminals being economically efficient, but there are other possible explanations. We then set out a methodology for establishing whether a specific concentration might be the opportunity for a successful intervention. We also argue that the mere possibility of a successful intervention on a specific concentration point does not necessarily mean that incentives of the various stakeholders will be sufficiently well aligned for that intervention to occur.

1 Introduction

Academic studies of cybercrime often draw attention to concentrations of wickedness. For instance, it may be reported that a disproportionate amount of spam comes from a few ISPs; that a large number of phishing websites are using the same registrars; and so forth. In such cases, the authors may well propose an intervention to tackle the criminality which leverages this concentration.

In this paper we explore how and why concentration occurs in online crime and set out a series of tests to apply to determine whether or not this concentration will be relevant when considering how best to deal with the criminality.

We find that concentration is often, but not always, rooted in economic causes. Concentration for instance may emerge due to barriers to entry, network effects and economies of scale. However, we also show that there exist other reasons that may lead measurements of online criminal activity to exhibit concentration. In particular, convenience and inertia of some of the actors may also produce concentrations, but also, and more problematically, measurement biases that are not properly accounted for could mislead researchers into discovering artificial concentrations which are not a reflection of reality.

[1] Computer Laboratory, University of Cambridge, JJ Thomson Ave., Cambridge, CB3 0FD, UK. richard.clayton@cl.cam.ac.uk

[2] Department of Computer Science and Engineering, Southern Methodist University, Dallas, TX 75275, USA. tylerm@smu.edu

[3] Electrical and Computer Engineering & CyLab, Carnegie Mellon University, Pittsburgh, PA 15213, USA. nicolasc@andrew.cmu.edu

Building on these findings, we then discuss a methodology to discern whether an observed concentration can lead to an effective intervention. We also ponder why, even when a concentration point is actionable, it may not necessarily be in the best interest of the various stakeholders to act upon it.

2 Concentration and Intervention in the Literature

We start by discussing some classic studies of concentrations of online criminal activity and the associated proposals for intervention. We emphasize that the literature review presented here is by no means exhaustive: there have now been over ten years of studies on the subject, and comprehensive coverage would fill a paper by itself. However, we argue that the research described below is a representative cross-section of the empirical work on online criminality, consideration of which will allow us to properly substantiate our thesis.

2.1 Concentrations by cybercrime category

During the course of cybercrime measurement studies, researchers have frequently uncovered evidence of concentration in the infrastructure used by criminals. We now review some of the key works and their findings.

Spam-sending botnets Van Eeten et al. studied the prevalence of infected computers sending email spam in the customer base of the 200 largest ISPs globally [30]. They found that 10 ISPs accounted for 30% of all spam, and that 50 accounted for over half the total. They found that the worst offenders were consistent over several years. Based on these findings, they argued that if the ISPs with the most infected machines made a conscious effort to remediate computers, then the harm could be greatly reduced. In follow-up work, the authors provided benchmarked measurements of infection rates to all of the major Dutch ISPs [29]. Two ISPs had a much greater fraction of infected customers than the rest. Upon learning of their relatively poor performance, employees at both ISPs convinced their management to devote more resources to remediation, and within a few months the ISPs met or exceeded the performance of others.

Spamvertised goods In their 2011 "Click Trajectories" paper, Levchenko et al. consider the value chain for advertising pharmaceuticals using email spam [15]. They identified a very large number of URLs, grouped them together and made a number of test purchases. Their explicit aim was "to identify any 'bottlenecks' in the spam value chain: opportunities for disrupting monetization at a stage where the fewest alternatives are available to spammers (and ideally for which switching cost is high as well)." They identified such a bottleneck in the payment tier of the value chain where they found that 95% of their payments were handled by 3 banks and there were only 13 banks used by the criminals in total. Levchenko et al. observed that there are high switching costs here (in both time and money terms) and suggest that payment generally might well be the most promising area in which to intervene.

Typosquatting Typosquatting is the registration of domain names that are similar to popular domains in the hope of being able to make money by showing adverts to visitors who have failed to key in the popular domain name correctly. Moore and Edelman [20] found that typosquatting is highly concentrated with just five individual identifiers (used to record who would be paid) accounting for 63% of the sites displaying Google adverts. Furthermore, some large domain name servers were hosting typosquatting domains as much as four times as often as the web as a whole. They concluded that the most viable intervention would be to crack down on the advertising platforms.

More recently Szurdi et al. [28] have examined a significantly larger sample of domains and found that "most true typo domains cluster at major registrars and are hosted at a few name servers. In particular, 12 name servers and 5 major registrars are responsible for hosting 50% of the true typo domains." They also wrote that "forcing these major registrars to enforce prudent registration practices with respect to typosquatting may be a viable policy option", however they also observed that "registrars and hosting companies do not suffer from typosquatting, thus there is little economic incentive for them expend resources to defend against it."

Websites selling counterfeit luxury goods Wang et al. [31] investigated malicious 'search engine optimization' (SEO), the practice of abusively manipulating search engine results to ensure a stream of visitors to websites that offer counterfeit versions of luxury and lifestyle fashion goods. They identified 52 distinct campaigns targeting 16 brands, and found that apart from the inherent use of search engines the activity was not concentrated in any manner which would suggest some sort of overarching intervention.

Instead, they noted a limited impact of the interventions that were occurring, such as demoting search results, marking websites as malicious or the brands seizing domains from the criminals. They concluded that the criminal activity was "organized as business campaigns, that effective interventions should target their infrastructure at the granularity of these campaigns, and that they [brands and search engines] are being targeted by dozens of campaigns". They also noted that "campaigns have shown great agility in adapting to partial intervention, and in filling in gaps left by the disappearance of other campaigns"

Media piracy In their "Clickonomics" paper Lauinger et al. investigate 'one click hosters' (OCHs) or 'cyberlockers' which are used to host infringing ('pirate') copies of music, films and other copyrighted material [11]. They found over 300 OCH sites and little evidence of activity being concentrated in any way – except in so far as some sites were larger and better resourced.

They found that that the interventions that were occurring were merely moving the problem on, concluding, "convincing or coercing OCHs to implement more advanced anti-piracy measures can drive away pirates. However, instead of ceasing to pirate, uploaders appear to move to less cooperative OCHs. Lower levels of piracy subsist even on the more diligent OCHs." However, they did note that "strategies against economically motivated actors may effectively target certain parts of the ecosystem. Together with increased legal proceedings, we expect them to render many OCHs more diligent, and maybe even proactive, in their own anti-piracy efforts."

High Yield Investment Programs A High Yield Investment Program (HYIP) is an online version of a financial scam in which investors are promised extremely high rates of return on their investments. Payments are made to existing investors from the funds deposited by newcomers, continuing until insufficient funds remain and the scheme collapses. Moore et al. studied HYIPs in 2012, finding a number of areas of concentration: low numbers of digital payment mechanisms were in use; many domains were being registered by a single registrar; and they identified the important role played by a small group of aggregation sites [21]

Neisius and Clayton revisited the topic in 2014, identifying all the flows of money within the HYIP ecosystem and drawing particular attention to the central role played by Gold Coders, a software house whose code was used by three-quarters of the HYIPs and nearly two-thirds of the aggregators [24]. Their view was that Gold Coders "are almost single-handedly removing almost all barriers to entry for a criminal that wishes to join those committing HYIP fraud".

One-click frauds and scareware Christin et al. described "One Click Frauds" [1], a specific type of 'scareware' scam frequent in Japan but seldom seen elsewhere, in which a malicious website operator displays a pop-up window informing the visitor that they entered a legally-binding agreement, and that they need to pay for a registration fee lest they be subject to legal action. There is of course no legal basis to the argument, but given the primarily pornographic nature of the websites, victims often pay fees in the order of USD 500.

By pooling together scams which used the same bank accounts to collect the fraudulent registration fees, and scams that relied on the same customer service phone numbers Christin et al. discovered that the top eight groups were responsible for half of all the scams observed between 2006 and 2009. They ascribed the success of these groups to relatively low barriers to entry and insufficient punishment (i.e., low fines and short or non-existent prison terms) in the cases where the perpetrators were identified and arrested.

In a related study of fake anti-virus distributors, Stone-Gross et al. [27] reported concentrations of affiliate networks engaging in this activity. They closely examined three different affiliate networks, and found that consistently only a small fraction of the affiliates were making the highest profits (e.g., only four out of 140 affiliates made more than USD 500 000 in one of these networks). At the opposite end of the spectrum, a majority of the affiliates they looked at did not make any money at all – for the three networks only 44/140, 98/167, and 541/1 107 affiliates generated any revenue.

Online anonymous marketplaces Online anonymous marketplaces are a relatively recent development, in which web sites running as Tor hidden services [6] are used in conjunction with the Bitcoin payment system [23] to form electronic commerce marketplaces with little restrictions on the goods that can be sold and bought. Christin [3] provided an empirical characterization of 'Silk Road,' which was the first large-scale such marketplace. Relevant to our study, he showed that the top 100 sellers were responsible for over 60% of all transactions in the first six months of 2012. Poring over the (publicly available) data [2] from Christin's paper, we can perform a new calculation which is of

direct relevance to this present paper: the top 3% of vendors were responsible for over 50% of all transaction volume.

Pharmaceutical inventories Leontiadis et al. [13] showed that approximately half of all unlicensed online pharmacies they surveyed appear to source their inventories from at most nine production facilities. They conjecture this is because production of high-quality counterfeit drugs is difficult, and requires specialized equipment or connection – while on the other end of the supply chain, setting up a website for distributing these pharmaceutical drugs is considerably easier.

Traffic to online pharmacies In the same arc of research, Leontiadis et al. [12, 14] measured that even though the entire network of pharmacies relying on poisoning of search-engine results for advertising is relatively large (2 232 pharmaceutical domains) most of it involves one of 382 "traffic broker" domains, which funnel traffic to these pharmacies from compromised sites. Even more symptomatic of concentration, an overwhelming majority of these traffic brokers appear to have been hosted on machines belonging to one single autonomous system (corresponding to a large cloud-service provider) before being displaced.

2.2 Concentrations by infrastructure component

Even though in the above discussion some of the work evidenced, as a by-product of their measurements, concentrations at the level of the supporting infrastructure, these papers primarily focused on describing separate case studies. On the other hand, researchers have also collected evidence of concentration by directly measuring its presence in various components of the Internet infrastructure. We review these studies now.

Payment system intervention In their 2012 "Priceless" paper – which followed up on the "Click Trajectories" paper mentioned earlier – McCoy et al. [17] observe that other parties have independently concluded that intervening to disrupt payments to criminals would be a useful strategy. They report on an initiative by VISA which came into effect in June 2011 that deemed online selling of pharmaceuticals to be high risk, which meant merchants had to be more substantial companies and acquiring banks could be fined for extending credit card services to unsuitable companies. Also in 2010 a new mechanism was introduced to allow undercover purchases from dubious companies to enter a streamlined complaints system.

By examining the usage of merchant accounts over time and correlating this with the time at which complaints were made the authors show that the complaints are effective, merchant accounts are closed and the criminal enterprises have to find new payment processors. The criminals of course moved on to other banks but they found evidence that where a purchase/complaint strategy was pursued on a consistent basis over time it had a considerable impact on the criminal activity.

Domain name interventions Liu et al. [16] describe their measurements of the impact of two initiatives to prevent criminals from registering domain names for use in email spam. The first initiative was a change in policy by the registry for `.cn`, the country code top level domain (ccTLD) for China. The policy for registering domains was changed overnight from a open system where domains cost approximately one US dollar, to a highly regulated regime under which the identity of those registering domains was checked, there was a limit on the number of domains per person and the price went up by a factor of ten. The impact was extremely dramatic, with abusive registrations of .cn domains effectively ceasing altogether – but the paper shows that the criminals moved to exploit the `.ru` (Russian) ccTLD instead. So the effect was one of displacement, along with a small increase in costs for registering the .ru domains.

The second initiative examined was the actions taken by a single registrar (eNom), albeit one of the biggest, in disabling domains associated with fake pharmacies. Here Liu et al. found some temporary disruption during an initial phase. But the criminals merely used other registrars, proactively transferring domains away from eNom that were as yet unaffected, but that they perceived to be at risk of future action. The paper concludes "concrete effects" had been measured but that "the current ecosystem provides spammers with ample room to adapt. We conclude that local interventions on a registry/registrar level are likely to be ineffective. To have an impact on spammers domain registration these interventions have to be extended to a global scale by ICANN."

Hosting provider shutdowns (McColo) In November 2008, following an investigation by the Washington Post journalist Brian Krebs, the Internet connectivity of the McColo Corp hosting company of San Jose, California was simultaneously disconnected by its two providers of Internet connectivity [9]. Krebs had drawn attention to the high concentration of malicious activity at the site – malware command and control servers, websites hosting fake pharmacies, fake designer goods, fake security products and the sale of child sexual abuse images delivered by email – and the connectivity providers had decided they should act.

The most obvious effect was an immediate reduction in global spam volumes because the criminals had not foreseen the entire site being disconnected. DiBenedetto et al. [5], show that the number of spam sources decreased by almost a half immediately after the shutdown and various industry observers such as Trend Micro [7] measured reductions of 50% to 75% in total volume. However, Clayton observes that the spam that disappeared was mainly of the easy-to-block kind, so that the impact on in-boxes (rather than spam-folders) was perhaps more limited [4].

In his recent book [10], Krebs explains that McColo was operated by a Russian group who were entirely aware of the type of activity their customers were engaged in and took steps to ignore or deflect abuse reports. Within a few weeks these customers had rebuilt their spam sending networks, and Krebs says that the only significant loss was some extensive lists of email addresses (the destinations for the email spam) and the criminals are now believed to more cautious about concentrating their systems at a single location.

2.3 The effect of biased measurements

When a study of criminal activity uses data about a subset of that activity (whether knowingly or unknowingly) then there is clearly a risk that the view of the activity will be biased – and, relevant to our discussion, it may throw up apparent concentrations of activity that are an artifact of the measurements rather than of reality. Examples of this effect in the published peer-reviewed literature are of course rare, but phishing (theft of credentials using fake websites) is notoriously difficult to measure without some bias creeping in and so we select two examples from this activity.

Phishing website lifetimes In their first published paper on phishing Moore and Clayton measured how long it took for the fake phishing websites to be 'taken down' [18]. This paper used data from PhishTank [25] and one of its figures was a comparison of take-down times for different banking brands. A year later Moore and Clayton revisited take-down time measurements, but now with several other sources of data about the location of phishing websites [19]. In particular they were now receiving data from two of the companies which provided 'take down' services to the banking industry.

In this second paper Moore and Clayton showed that take-down times were markedly lower for phishing websites that were known to the take-down company providing a service to the brand being phished. If a rival take-down company learned of the website then they took no action if the brand was not their customer – however, they still provided information about the phishing URL for Moore and Clayton's research – thereby starting the clock ticking on a significantly longer lifetime for the phishing website.

In retrospect, much of the variability in take-down times between brands in the earlier work could be ascribed to measurement bias and the way in which some, but not all, of the take-down companies were also looking at the PhishTank data – there was no criminal conspiracy to concentrate on using more robust website hosting arrangements when targeting particular brands.

Phishing toolbar effectiveness Sheng et al. looked into how effective toolbars were in detecting that a phishing webpage was being visited [26]. Their methodology required a source of brand new phishing URLs that they would then regularly visit with each of the toolbars active in turn to assess whether or not the page was identified as a phish – that is that the company involved had learned of URL and correctly categorized it as malicious. They obtained their URLs from the University of Alabama at Birmingham 'Spam Data Mine' which in turn received data from, among others, a spam filtering company handling over a billion emails a day.

The work went well when considering an October 2008 dataset, but when they repeated it on a December 2008 dataset they found that two of the toolbars were suddenly much more effective. They reviewed this and concluded that "the two tools acquired new sources that were similar to our feed". That analysis underlines the entire difficulty of determining whether it is useful to use blacklists to counter criminality. Only the items in the blacklist will be blocked, but measuring how good the blacklist is – whether the criminality is sufficiently concentrated for everything relevant to get into the blacklist – is very difficult to measure objectively.

3 Why Does Concentration Emerge?

We next discuss why concentration emerges. We argue that, although economic rationality plays a preponderant role in the emergence of concentration points, it is not the only reason why concentration occurs.

3.1 Rational economic behavior

It is generally believed that online criminals are rational economic actors – in that they wish to make money as efficiently as possible. This means that we would expect to see concentration when it is economically efficient and thus it is relevant to review economic drivers towards concentration.

Comparative advantage In economies where individuals and firms are free to trade goods and services, the theory of comparative advantage predicts that specialization will emerge. Individuals and firms that can most efficiently produce a good or service will tend to focus on providing that good and selling it to others, buying remaining goods and services from others.

On the Internet, companies specialize in providing different services cheaply, from domain name registration to hosting to advertising. Such specialization is not confined to legitimate purveyors, however. Unscrupulous providers specialize in offering services that are attractive to criminals, for example by basing themselves in countries where they can operate without fear of prosecution. Criminality subsequently concentrates in these areas, such as where domain names are offered for free or at (considerably more expensive) 'bullet-proof' hosting companies that boast that they will ignore abuse complaints.

Network effects Towns grow in the real world because of network effects. Raw materials coming into a port are processed in the surrounding area and more expensive goods are re-exported. The market towns develop financial sectors and become centers of administration. Similarly, the three major US automakers are all headquartered in Detroit, where an extensive industry of parts suppliers emerged.

It has commonly been observed that certain types of criminality are more closely associated with particular parts of the world: fraud with West Africa, auction fraud with Eastern Europe, financial malware with Russian speaking countries. These types of criminality generally require multiple actors specializing in different aspects of the crime and although the individuals involved could be anywhere in the world, issues of trust and language may mean that they are offline friends or acquaintances first, and criminal colleagues later.

Economies of scale Any tour of a medieval town will take in streets named after the trades that operated there. Although there are many reasons for these concentrations a key one is the economies of scale. Dealing with the mess created by the butchers is much

simpler if they are all in one place; the inhabitants of a Chinatown may co-locate for mutual support or because they are unwelcome in other neighborhoods.

Online, scaling is everything and most of the costs of a criminal enterprise are upfront. Once you have written your malware and worked out how to distribute it, then the more people you can infect then the bigger the return.

Winner-take-all dynamics Once one city becomes the largest in the land, its leading role may last for centuries. London's importance as a commercial center meant the King moved there from Winchester. Its economic and political importance made it the center of the UK road network ('all roads lead to Rome'), then of the railway network and then the location of the most important UK airports.

Online, the low levels of 'friction' leads to numerous monopolies in legitimate online businesses – and also in less legitimate operations, such as the way that GoldCoders dominates the market for HYIPs.

Barriers to entry Towns and cities are now, in the developed world, subject to extensive planning regulations so those who want to found new towns or build new housing estates face considerable difficulties.

Online the main barriers to entry into criminal activity are the expertise needed to be efficient and effective. Spam sending and spam blocking have been in an 'arms race' for so long that there is a very steep learning curve for new entrants.

Online criminal activity linked with the trade of physical goods (e.g., pharmaceutical counterfeits or narcotics) typically presents much higher barriers to entry than 'pure' online crime focusing only on digital goods. Indeed, access to, or production of, physical goods is typically considerably more expensive than that of digital assets.

An example of this is in the way Silk Road was launched: its operator grew psychedelic mushrooms himself so he could to sell them on his nascent marketplace. This made the case his website was 'legitimate' and thereby addressed a steep barrier to entry. Once customers were satisfied that the site 'worked', other vendors (and many more customers) joined—leading to the aforementioned network effects. Likewise, the concentration observed in inventories offered in the sale of counterfeit pharmaceuticals suggests that the production (or obtaining) of such goods by miscreants incurs very high start-up costs.

3.2 Non-economic factors

However, not all criminals are economically efficient all of the time, and there are other, more structural, reasons why concentrations can be observed. We identify two such reasons here – which are nothing to do with economic rationality but are instead forms of inertia.

Whack-a-mole Many countermeasures to criminal activity such as phishing or spamming involve the sending of abuse complaints to hosting companies or ISPs. The competent and caring companies will remove malicious content or ensure that compromised

machines are cleaned up. However, the less competent will do less and so over time the badness will mainly be found in these uncaring locations. That is the concentration occurs not because the criminals choose to host at these sites, but rather that they end up there by default.

Copying successful patterns When criminals are seen to be successful then others may seek to imitate that success. This may lead to concentrations because the imitators copy most aspects of the initial operation without necessarily assessing alternatives. If one criminal uses eGold for payments and seems to do well then others may follow. Of course, over time the economic effects we discussed above come into play – for example, there's network effect and monopolistic tendencies because criminals receive eGold as income and wish to use it directly for expenditure.

3.3 Measurement methodology artifacts

Furthermore, studies may identify a concentration whereas there is something else going on altogether.

Measurement bias As discussed earlier, measurements may be biased – the online reporting center IC3 received large numbers of reports of auction fraud – but this was because eBay made a point of drawing IC3 to the attention of their customers when they had been victimized. This skewed the incidence of one type of crime.

If a study uses a biased dataset then large numbers of similar crimes may be missed. Many academic studies of phishing use the publicly available data from PhishTank – this only contains about 40% of all phishing URLs, but 100% of all PayPal phish, which the wide-ranging studies consistently identify as the most attacked brand. Clearly, if composition of the PhishTank dataset is not understood, then the concentration on attacking PayPal will be overestimated.

Another example, discussed by Kanich et al. [8], is that of botnet measurements. Kanich et al. observed in the course of their measurement of the Storm botnet that other researchers were undertaking measurements at the same time, thereby biasing the overall results. While, in that specific study, the bias in measurement does not seem to have affected findings related to potential concentration points, it certainly could have had an impact on any findings given that a number of "bots" were actually measurement hosts that should not have been accounted for in any analysis.

Measuring something else altogether It's not unusual to see maps or tables of the location of malware infected machines – typically the USA or China heads these lists. However, the numbers are seldom normalized to account for the size of the population or, more relevantly, the proportion of that population who own a computer and whether or not they leave it permanently online.

Apparent correlations between two datasets may not indicate a causal linkage, but that the two datasets are both, independently, linked to the same thing.[1]

Small numbers of participants If there are only a handful of criminals operating a particular scam in a particular manner then an apparent concentration of criminality is inevitable. In 2007–2008 the RockPhish gang sent about two-thirds of all phishing URLs and operated in a distinctive manner using fast-flux DNS techniques. This meant that almost anyone who studied phishing would come across them – but their methods were seldom copied by anyone else, so that the apparent importance of fast-flux to phishing success was to some extent merely an artefact of the scale they operated at.

Additionally, as we have already seen, many crimes operate under an affiliate system – so that a large number of individual criminals, with very disparate behavior are advertising a relatively small number of fulfillment locations.

3.4 Reviewing concentration causes in previous studies

In light of the above discussion, we summarize in Table 1 the factors that appear to have presided over the establishment of concentration in some of the previous work we introduced. The table shows, for each type of online criminal activity we surveyed, whether observed concentrations are due to economic factors, or to other reasons. We observe that, in a number of cases, the observed concentrations are partially due to relatively small number of actors overall; in other cases (one-click frauds, pharmacies and HYIP scams), "whack-a-mole" or inertia effects may have caused somewhat artificial concentrations that we expect would not lead to fruitful intervention.

Category	Perpetrators	Payments	Registrar	Hosting	ISP	Phone	Ad Platform
Spam-sending botnets	E				E		
Typosquatting	E C S		E W				S
HYIPs	E S	S					
Luxury goods	S		E	E			
Media piracy							
One-click frauds	E S		S W			S	
Fake A/V	S		W	E			
Unlicensed pharmacies	S	S		S W			
Goods sold on anon. mkts.	E C			S			
Phishing	S		E B	E B			

Table 1: Legend of concentration factors: E: Economic, W: Whack-a-mole, C: Copycats, B: Bias, S: Small number of participants.

[1] Munroe [22] humourously considers the similarity between heat maps of viewers of furry pornography and the Martha Stewart website. In practice, both maps are of centers of population in the USA.

4 A Methodology for Proposing Interventions

When documenting the results of a cybercrime study, it is natural to wish to propose some countermeasures that would reduce the incidence or impact of the cybercrime. When the study has identified some sort of concentration, usually of the use of some resource, then it may well be that intervening to disrupt access to that resource will disrupt the crime. Indeed a whole discipline, 'crime script analysis,' is based on the notion that crimes can be broken down into a series of stages and countermeasures and interventions can be seen as an attempt to disrupt one or more stages of that script.

We propose the following methodology that will ensure the soundness of an intervention proposal that aims to leverage the presence of concentration:

Step One – is the concentration real? As we have discussed above, concentrations can be caused by numerous types of measurement bias. Hence any discussion of a proposed intervention whose effectiveness depends upon a concentration must be preceded by a discussion of the validity of the analysis – why this is not the equivalent of the drunk looking for their keys under the streetlight.

That said – there may be some economies of scale in tackling concentrations of cybercrime whatever the nature of the concentration. For example, there will be value in working with an engaged ISP that is prepared to clean up malware on their customer machines even if the only reason that they head an infection league table is because they are the biggest ISP with the largest number of customers. However, it should not be claimed that this is making an existential difference to the malware's prevalence.

Equally, if it can be determined that there are only five criminals operating a particular scam then arresting all five will make a difference. However, such an intervention will not, of itself, ensure that five more (or indeed fifty more) criminals do not follow in their footsteps. In fact, we have seen instances of this in practice: after the Silk Road was taken down, a number of copycat marketplaces emerged to take its place, ultimately resulting in a richer – and more robust – ecosystem.

Step Two – identify how a viable intervention would work For an intervention that leverages a concentration to be likely to make a difference there must be some sort of structural reason for the concentration to exist.

It is, for example, difficult for criminal websites to receive payments, so if concentrations can be found in their financial arrangements then it may be reasonable to assume that intervening to disrupt these arrangements will be effective.

Likewise, production of nigh quality counterfeit pharmaceutical drugs usually requires either access to a production facility in which some of the stock can be stolen or acquired illicitly; or access to expensive and highly controlled equipment, namely 'pill presses', not to mention the expertise required in producing chemically indistinguishable products.[2]

[2]Alternative distribution models, such as exploiting arbitrage, where somebody purchases pills in a given country and resells them online to customers from a different country, may also occur – but would not cause the same concentration effects.

However, a concentration which only exists for reasons of simple convenience may indicate that an intervention although in many ways viable, could be of short term impact and perhaps counterproductive. For example, if many different criminals are using a registrar solely because of the price they pay for domains, then persuading the registrar to turn down their business may slightly increase the criminals' costs, but is unlikely to put them out of business – and by dispersing the activity among competing registrars may make it harder to track the activity and intervene in future.

Step Three – predict the criminals' response Successful criminals will wish to continue their activity despite an intervention. Thus, for an intervention to have a long term impact it must be difficult or inefficient for the criminals to make a substitution.

Thus, if all of the payments for a service are currently directed through Ruritania the success of an intervention that involves shutting down Ruritanian services will depend upon whether the criminals can immediately move operations to Freedonia, or whether there was something unique about the Ruritanian arrangements.

Step Four – assess the practicality of the intervention Even if a plausible strategy for intervention is identified, it may not turn out to be practical. Observing that Law Enforcement officials would be able to disrupt a crime in a particular way does not mean that they consider the crime sufficiently important to see this as something to spend time on. Identifying a particular intermediary as being uniquely capable of disrupting criminal activity does not mean that the intermediary will be prepared to co-operate.

Even more problematically, internal incentives within Law Enforcement may actually play against intervening on a specific concentration point. In countries where law enforcement personnel are promoted based on the number of arrests made or cases solved, it may be more advantageous to, for instance, take down multiple instances of a given criminal activity (e.g., websites selling pharmaceutical drugs) rather than going after a handful of important intermediaries (e.g., payment processors). In situations like this, it is extremely important that policy makers be educated about the potential incentive misalignment and rectify it.

5 Summary and Conclusions

We have reviewed some of the cybercrime literature to draw attention to the many occasions on which authors have identified concentrations within criminal activity and have seen that this often leads them to suggest that this leads to a possibility of an effective intervention. We do not claim this literature review is exhaustive: rather than a complete description of the entire body of work in the area, we wanted to focus on the few papers that explicitly tried to assess whether concentration points existed – be it among actors, or parts of the infrastructure.

We then discussed the general ways in which concentrations can arise (or appear to arise) because unless the reason for the concentration is understood, then going on from that to suggest an intervention is unlikely to be successful.

We then set out the bare bones of a methodology for proposing a successful intervention in criminal activity. It is essential to understand the relevance of the concentration and understand how the criminals will respond. We also observed that identifying a viable intervention does not mean that the incentives will be sufficiently well aligned for that intervention to occur.

Acknowledgements

The authors are funded by the Department of Homeland Security (DHS) Science and Technology Directorate, Cyber Security Division (DHSS&T/CSD) Broad Agency Announcement 11.02, the Government of Australia and SPAWAR Systems Center Pacific via contract number N66001-13-C-0131. This paper represents the position of the authors and not that of the aforementioned agencies.

References

[1] N. Christin, S. Yanagihara, and K. Kamataki. Dissecting one click frauds. In *Proc. ACM CCS'10*, pages 15–26, Chicago, IL, October 2010.

[2] Nicolas Christin. Traveling the Silk Road: Datasets, 2012. `https://arima.cylab.cmu.edu/sr/`.

[3] Nicolas Christin. Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. In *Proceedings of the 22nd World Wide Web Conference (WWW'13)*, pages 213–224, Rio de Janeiro, Brazil, May 2013.

[4] Richard Clayton. How much did shutting down mccolo help? In *Sixth Conference on Email and Anti-Spam (CEAS 2009)*, 2009.

[5] Steve DiBenedetto, Dan Massey, Christos Papadopoulos, and Patrick J. Walsh. Analyzing the aftermath of the mccolo shutdown. In *Proceedings of the 2009 Ninth Annual International Symposium on Applications and the Internet*, SAINT '09, pages 157–160, Washington, DC, USA, 2009. IEEE Computer Society.

[6] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. In *Proceedings of the 13th USENIX Security Symposium*, August 2004.

[7] J. M. Hipolito. Spam volume plummets as ISPs pull the plug on McColo, 2008. `http://blog.trendmicro.com/trendlabs-security-intelligence/spam-volume-plummets-as-isps-pull-the-plug-on-mccolo/`.

[8] C. Kanich, K. Levchenko, B. Enright, G. Voelker, and S . Savage. The Heisenbot uncertainty problem: challenges in separating bots from chaff. In *Proceedings of USENIX LEET'08*, San Francisco, CA, April 2008.

[9] Brian Krebs. Host of Internet spam groups is cut off. *Washington Post*, 12 Nov, 2008.

[10] Brian Krebs. *Spam Nation: The Inside Story of Organized Cybercrime – from Global Epidemic to Your Front Door*. Sourcebooks, 2014.

[11] Tobias Lauinger, Martin Szydlowski, Kaan Onarlioglu, Gilbert Wondracek, Engin Kirda, and Christopher Kruegel. Clickonomics: Determining the effect of anti-piracy measures for one-click hosting. In *NDSS*, 2013.

[12] N. Leontiadis, T. Moore, and N. Christin. Measuring and analyzing search-redirection attacks in the illicit online prescription drug trade. In *Proceedings of USENIX Security 2011*, San Francisco, CA, August 2011.

[13] N. Leontiadis, T. Moore, and N. Christin. Pick your poison: pricing and inventories at unlicensed online pharmacies. In *ACM Conference on Electronic Commerce*, 2013.

[14] N. Leontiadis, T. Moore, and N. Christin. A nearly four-year longitudinal study of search-engine poisoning. In *Proceedings of ACM CCS 2014*, Scottsdale, AZ, November 2014.

[15] Kirill Levchenko, Andreas Pitsillidis, Neha Chachra, Brandon Enright, Márk Félegyházi, Chris Grier, Tristan Halvorson, Chris Kanich, Christian Kreibich, He Liu, et al. Click trajectories: End-to-end analysis of the spam value chain. In *IEE Symposium on Security and Privacy*, pages 431–446, 2011.

[16] He Liu, Kirill Levchenko, Márk Félegyházi, Christian Kreibich, Gregor Maier, Geoffrey M Voelker, and Stefan Savage. On the effects of registrar-level intervention. *Proc. of 4th USENIX LEET*, 2011.

[17] Damon McCoy, Hitesh Dharmdasani, Christian Kreibich, Geoffrey M Voelker, and Stefan Savage. Priceless: The role of payments in abuse-advertised goods. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, pages 845–856. ACM, 2012.

[18] Tyler Moore and Richard Clayton. Examining the impact of website take-down on phishing. In *Proceedings of the Second APWG eCrime Researcher's Summit*, Pittsburgh, PA, October 2007.

[19] Tyler Moore and Richard Clayton. The consequence of non-cooperation in the fight against phishing. In *Third APWG eCrime Researchers Summit*, Atlanta, GA, October 2008.

[20] Tyler Moore and Benjamin Edelman. Measuring the perpetrators and funders of typosquatting. In Radu Sion, editor, *Financial Cryptography and Data Security*, volume 6052 of *Lecture Notes in Computer Science*, pages 175–191. Springer, 2010.

[21] Tyler Moore, Jie Han, and Richard Clayton. The postmodern ponzi scheme: Empirical analysis of high-yield investment programs. In Angelos D. Keromytis, editor, *Financial Cryptography and Data Security - 16th International Conference, FC 2012, Kralendijk, Bonaire, February 27-March 2, 2012*, volume 7397 of *Lecture Notes in Computer Science*, pages 41–56. Springer, 2012.

[22] Randall Munroe. Heatmap, 2012. http://www.xkcd.com/1138/.

[23] S. Nakamoto. Bitcoin: a peer-to-peer electronic cash system, October 2008. Available from http://bitcoin.org/bitcoin.pdf.

[24] Jens Neisius and Richard Clayton. Orchestrated crime: The high yield investment fraud ecosystem. In *Proceedings of the Ninth APWG eCrime Researcher's Summit*, 2014.

[25] PhishTank. https://www.phishtank.com/.

[26] Steve Sheng, Brad Wardman, Gary Warner, Lorrie Faith Cranor, Jason Hong, and Chengshan Zhang. An empirical analysis of phishing blacklists. In *Proceedings of the Sixth Conference on Email and Antispam (CEAS)*, July 2009.

[27] Brett Stone-Gross, Ryan Abman, Richard A. Kemmerer, Christopher Kruegel, Douglas G. Steigerwald, and Giovanni Vigna. The underground economy of fake antivirus software. In *Proceedings of the 10th Workshop on the Economics of Information Security*, Fairfax, VA, June 2011.

[28] Janos Szurdi, Balazs Kocso, Gabor Cseh, Jonathan Spring, Mark Felegyhazi, and Chris Kanich. The long taile of typosquatting domain names,. In *23rd USENIX Security Symposium (USENIX Security 14). San Diego, CA: USENIX Association*, pages 191–206, 2014.

[29] Michel van Eeten, Hadi Asghari, Johannes M. Bauer, and Shirin Tabatabaie. Internet service providers and botnet mitigation: A fact-finding study on the dutch market. Technical report, Netherlands Ministry of Economic Affairs, Agriculture and Innovation, 2011. http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2011/01/13/internet-service-providers-and-botnet-mitigation.html.

[30] Michel van Eeten, Johannes M. Bauer, Hadi Asghari, Shirin Tabatabaie, and Dave Rand. The role of internet service providers in botnet mitigation: An empirical analysis based on spam data. In *9th Annual Workshop on the Economics of Information Security, WEIS 2010, Harvard University, Cambridge, MA, USA, June 7-8, 2010*, 2010.

[31] David Y. Wang, Matthew Der, Mohammad Karami, Lawrence Saul, Damon McCoy, Stefan Savage, and Geoffrey M. Voelker. Search + seizure: The effectiveness of interventions on seo campaigns. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, pages 359–372, New York, NY, USA, 2014. ACM.