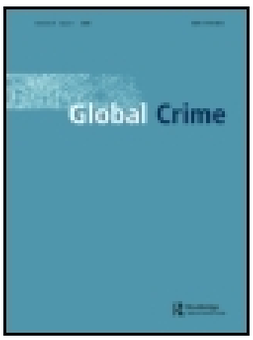

Global Crime

ISSN: 1744-0572 (Print) 1744-0580 (Online) Journal homepage: http://www.tandfonline.com/loi/fglc20

The ecology of trust among hackers

Benoît Dupont, Anne-Marie Côté, Claire Savine & David Décary-Hétu

To cite this article: Benoît Dupont, Anne-Marie Côté, Claire Savine & David Décary-Hétu (2016): The ecology of trust among hackers, Global Crime, DOI: 10.1080/17440572.2016.1157480

To link to this article: http://dx.doi.org/10.1080/17440572.2016.1157480

Published online: 11 Mar 2016.

Global Crime, 2016
http://dx.doi.org/10.1080/17440572.2016.1157480

The ecology of trust among hackers

Benoît Dupont[a]*, Anne-Marie Côté[a], Claire Savine[b] and David Décary-Hétu[a]

[a]*École de criminologie, Université de Montréal, Montréal, Canada;* [b]*Polytech Nantes, Nantes, France*

Malicious hackers profit from the division of labour among highly skilled associates. However, duplicity and betrayal form an intrinsic part of their daily operations. This article examines how a community of hackers uses an automated reputation system to enhance trust among its members. We analyse 449,478 feedbacks collected over 27 months that rate the trustworthiness of 29,985 individuals belonging to the largest computer hacking forum. Only a tiny fraction of the forum membership (2.4%) participates in the vast majority (75%) of 'trust exchanges', limiting its utility. We observe a reporting bias where the propensity to report positive outcomes is 2.81 times greater among beginner hackers than among forum administrators. Reputation systems do not protect against trust decay caused here by the rapid expansion of the community. Finally, a qualitative analysis of 25,000 randomly selected feedbacks indicates that a diverse set of behaviours, skills and attitudes trigger assessments of trustworthiness.

Keywords: online offenders; hackers; trust; reputation; botnets

Online offenders, like any other professional group, have learned to harness information and communication technologies to overcome physical constraints and exploit new economic opportunities. As a result, they are profiting from the automation of fraud and an increased ability to divide labour among highly skilled associates.[1] Although antivirus and security companies, who have a vested interest in overstating the size of the problem, cite extravagant numbers, the most rigorous and conservative academic study conducted to date estimates that the revenues generated by 'genuine' and 'transitional' cybercrime amounted to 16.6 billion US dollars in the early 2010s.[2] Global illicit networks operating in this new technological environment have a lower risk of exposure and arrest – due to the malleable nature of online identities and the fragmentation of law enforcement responses – but they must contend with the considerable challenge of establishing and maintaining trust among co-offenders in online communities that lack the traditional social control and signalling mechanisms found in more traditional criminal settings.[3] This challenge is frequently underestimated in the literature, as researchers, for understandable methodological reasons, tend to focus on the collaborative features of technological platforms such as IRC channels[4] or web discussion forums[5] that help strangers with rare and complementary skills find each other.

However, several recent journalistic accounts of a thriving cybercrime underground have shown that duplicity and betrayal form an intrinsic part of such entrepreneurs' daily operations – and, in the end, are their undoing. In the book *Kingpin*, former convicted hacker turned investigative journalist Kevin Poulsen tells the true story of Max Butler, an

*Corresponding author. Email: benoit.dupont@umontreal.ca

ambitious hacker who engineered the hostile takeover and consolidation of seven rival underground forums specialising in the resale of stolen credit card credentials, eventually depriving 10,000 online fraudsters of their existing accounts and forcing them to trade through his newly formed empire, Carders Market.[6] Butler was arrested two years later after his identity was revealed by one of his co-offenders, during negotiations for a lighter sentence following his own arrest. In the same period, an FBI undercover agent managed to infiltrate a competing carding forum – DarkMarket – and ran it for two years, despite Butler's very public accusations that he was a spy, collecting troves of evidence on every transaction before the forum was shut down and 60 of its most active members arrested.[7] Meanwhile, in Russia, the operators of the two largest illegal online pharmacies – who had previously worked together – launched a destructive confrontation that culminated in the leak of their respective internal databases to journalists and researchers. The complex structure of their criminal networks was exposed, including their heavy reliance on computer hackers, who controlled millions of compromised machines, to promote their goods.[8] More recently, Ross Ulbricht, the founder of Silk Road, the first cryptomarket to allow anonymous drug transactions and the publication of consumer ratings, unknowingly paid an undercover Drug Enforcement Administration officer to have one of his staff members killed. The staff member had not only been arrested and was possibly cooperating with investigators but $350,000 in bitcoins had evaporated from the Silk Road accounts under his control.[9] These examples focus on elite online offenders, whose skills and criminal earnings probably dwarf the more modest achievements of their less talented or committed peers. However, no matter how successful and profit-driven they are, malicious hackers at all levels face the same trust problem.

How to strike the right balance between cooperation and security is a classic dilemma for individuals and groups who operate in hostile social environments.[10] Ever since property crime shifted from a craft to a project form of organisation during the Industrial Revolution, offenders have faced this cooperation/protection puzzle.[11] They must forge, identify and reinforce trusted and productive ties with notoriously unreliable co-offenders while lacking contractual enforcement tools and in a context where failures, mistakes or malfeasance can result in arrests.[12] Gambetta identifies four co-operation motivating mechanisms that are available to offenders: coercion (fear of sanctions), interests (mutual economic benefits), values (belief in the inherent virtue of cooperation) and personal bonds (reciprocal obligations).[13] However, some of these mechanisms become either less effective or operate very differently in online settings.[14] Coercion, for example, which relies on the use of violence when legal recourses are unavailable, is a powerful tool in the physical world but becomes very difficult to wield credibly online, where distance and anonymity represent major barriers to swift physical retaliation.[15] The values and personal bonds that sustain traditional trust networks shield their members from predators – in this case, law enforcement agencies.[16] But the strong ties, boundaries and mutual obligations that help networks thrive offline are harder to sustain online, where breadth and an emphasis on large collections of weak ties are often favoured over depth. Finally, mutual economic benefits are powerful incentives that bring online offenders together, but the absence of credible enforcement mechanisms, combined with the global and potentially unlimited supply of aspiring co-offenders, creates an environment where dishonesty can be more profitable than cooperation, with limited cost for the offending party. Under such constraints, 'trust', which we define as a mechanism for social complexity reduction that relies on expectations about others' future behaviour when making decisions,[17]

becomes much harder to sustain and can be too fragile and contingent to foster a general and profitable climate of cooperation.

A growing number of recent contributions have extended the classical literature on trust among co-offenders to cybercrime settings. They enhanced our understanding of the role trust plays in online offending, where co-offenders have limited pre-existing social ties,[18] by examining the various mechanisms implemented by underground forum administrators to increase self-regulation and prevent deceit among participants. Our research however is the first one to examine an entire online offender community and to measure at the aggregate level how reputation, which we use as a precursor of trust,[19] is distributed among its members and fluctuates over time. In order to do so, we focus on the use by online offenders of a reputation system first implemented by e-commerce websites such as eBay to enhance trust and foster collaborative ties with their peers, and the challenges they face in the process. We are particularly interested in using the feedback data generated by this reputation system to infer the quantity and quality of trust found in the largest community of general and malicious hackers currently operating online. To clarify the various sources of trust in this particular community, we combine quantitative and qualitative methodologies. This innovative approach provides unique insights into the complex mix of transactional, behavioural and cultural factors that establish someone's trustworthiness.

While online interactions are largely anonymous, they have the advantage of being more persistent and easier to collect than those that occur offline. We leveraged this feature to create a database of almost 450,000 events that rated the outcomes of interactions between users of the largest hacking forum on the internet over 27 months. In the first section, we discuss briefly the literature on various strategies used by cybercrime forum administrators to assess the trustworthiness of their members and then outline the theory of reputation systems, initially developed by large e-commerce websites such as eBay. In a second section we describe the data collected and the main features of the forum from which they were extracted, as well as some basic properties of the communities under study. In the third section we present the results of our quantitative and qualitative analyses. Finally, the fourth section examines the theoretical and practical implications of our research and provides our conclusions about the benefits and limits of the large datasets that are becoming ubiquitous in the era of Big Data social science.

The online offender's dilemma: trusting strangers who trade in deception

In underground markets, 'rippers' are dishonest participants 'who do not provide the goods or services for which they've been paid',[20] who provide worthless goods – available for free elsewhere for example,[21] or who sell products equipped with 'backdoors' that allow thieves to steal their peers' loot.[22] Rippers abound in openly accessible underground markets, where they prey on the masses of gullible wannabe offenders lured by the promise of easy money – a promise often unintentionally amplified by security companies.[23] Rippers are the scourge of illicit markets, generating a level of uncertainty that makes participants more reluctant to trade, thereby decreasing the volume of transactions as well as the attraction of a particular marketplace in a highly competitive environment. As a result, underground forums use four main reputation mechanisms to allow transactions to proceed in environments where trust is scarce.[24] First, forum administrators can become reputation managers, often for a fee, and are then authorised to award a special verified status to a selected group of participants after reviewing the quality of their offerings and assessing their reliability.[25] A public variation of verified

status is prowess demonstration, where a hacker's trustworthiness is assessed by his ability to demonstrate essential technical abilities to the community.[26] Second, administrators can offer escrow services to participants, acting as intermediaries and releasing money to sellers once buyers are satisfied with their purchase.[27] Third, forum administrators can provide semi-formal conflict resolution mechanisms, whereby a complaint can be lodged by an unsatisfied customer, evidence (usually chat logs and private messages) presented by both sides is publicly assessed, and a settlement is made that binds all parties, who run the risk of being banned from the forum if they don't comply – although they can obviously return under a new identity.[28] Finally, administrators can delegate reputation management to users by implementing a self-regulatory mechanism that relies on community feedback, similar to systems found on eBay, Amazon, TripAdvisor or Yelp.

The problem of online trust is not unique to cyberoffenders and was addressed in the mid-1990s by e-commerce marketplaces that needed to guarantee the integrity of transactions taking place on a global scale between strangers who had very limited information on which to evaluate each other's reliability.[29] In order to provide such information, online marketplaces designed reputation systems that enable them to replicate traditional word-of-mouth mechanisms on a much larger scale. By letting participants provide public feedback about each transaction, they create a transparent aggregate record of fulfilled or unfulfilled promises on which future decisions can be based. To be successful, these reputation systems must provide information of sufficient quality to allow buyers and sellers to properly evaluate potential vendors or customers, encourage participants to fulfil their obligations and discourage untrustworthy vendors or customers.[30] Online commerce platforms have benefited significantly from these reputation and review systems, although they still have to deal with structural problems, such as low incentives to provide ratings, a strong bias toward positive ratings, difficulty in filtering out unfair or fake ratings, the ease with which untrustworthy online identities can be discarded and replaced by new, untainted profiles and variations in quality over time.[31] One can assume that these problems are magnified when the system is used by a community of strangers who trade in deception.

Using a game-theoretic model, Mell argues that reputation systems should play an instrumental role in online criminal markets and allow them to reach equilibrium by solving the information asymmetry conundrum,[32] a major source of abuse and inefficiency according to Herley and Florêncio.[33] Lusthaus also identifies reputation systems as one of a broader set of strategies designed to foster collaboration and helping to overcome the challenges of anonymity and chronic distrust inherent to underground markets.[34] However, Mell does not test his hypothesis with real world data,[35] and Lusthaus relies on qualitative interviews that provide sufficient material to develop a typology of practices and technologies facilitating the emergence and maintenance of trust,[36] but does not explain how entire communities leverage these tools and how effective each one of these tools is. Most studies conducted on the subject also focus on much smaller groups of hackers,[37] rely on qualitative methodologies that make the measurement of trust impossible,[38] do not examine the very different motivations that reflect the multiple dimensions of trust[39] or focus on the individual rather than the collective distribution of online trust.[40] Moreover, to our knowledge, no criminological research on the role trust plays in online offending has considered the temporal dimension. Hence, the present article makes an original contribution to the understanding of the ecology of trust sustaining communities of online offenders by measuring how positive and negative reputation ratings punctuate social interactions between community members. To quantify the levels and fluctuations of trust in underground forums, we look at four main questions:

how and to what extent do online offenders use systems implemented by underground forum administrators to limit the risks of dealing with unknown peers? How effective are these systems in this particular environment, whose contingencies and constraints differ greatly from those encountered in commerce-oriented platforms? What kinds of information seems most valued in determining the trustworthiness of community members? Finally, can these reputation systems foster enough trust between online offenders to make malicious hacker and online fraudster communities sustainable?[41] We find that not all categories of online offenders are equally effective at leveraging reputation and trust: specialised botnet hackers appear to be much more capable of establishing and maintaining trusted relationships that result in positive outcomes than their generalist counterparts. As well, reporting in reputation systems seems to be heavily biased according to the position of the rater in the system: newcomers refrain from negative assessments while forum staff members and administrators post a majority of negative reviews. The problems created by this reporting bias are compounded by the instability of reputation systems: as the number of online offenders using the system increases, the quality of outcomes and the general levels of trust seem to plummet. Finally, we identify an aspect of the culture of malicious hacker communities that is intrinsically resistant to any form of social control and subverts reputation systems, disrupting the ecology of trust that some forum organisers are trying to establish.

Data and methodology

The data used in this research was extracted from a discussion forum dedicated to hacking, which at the time of writing was the largest online community for the asynchronous discussion of matters related to computer hacking. Launched in May 2007, as of May 2015 it catered to 2,723,344 registered users and had aggregated more than 46 million posts. Sub-forums address a broad range of issues at all levels, from beginner to very advanced, and explore the technical intricacies and vulnerabilities of various computer systems (from Microsoft Windows and Apple OS to Linux and smartphone platforms), languages (Visual Basic, C/C++, Python, Java, etc.) or communication tools (Skype, IRC, ICQ), as well as providing access to tutorials and hacking programs that can be freely exchanged between members. A large number of discussions on this forum are also non-technical and sections dedicated to cultural topics (such as music, films, TV shows, anime art), sports, current events or personal life issues such as health and education are very active. While the forum officially enforces strict rules to prevent 'black hat' hacking, commonly defined as illegal behaviours that result in breaches to the integrity of third-party computer systems, [42] it would be difficult for some sub-forum participants, such as the ones on the 'botnets' section from which most of our data originates, to engage in their preferred form of hacking while remaining within the confines of the law.

To understand why participants in botnets discussions can reliably be described as online offenders, a label that would be problematic for most other forum members, it is necessary to explain briefly what botnets are, what they do and how they provide the infrastructure for most of today's sophisticated online crimes. A botnet can be described as a group of infected computers, or 'bots', which are under the influence of malicious code controlled by an individual known as a 'botmaster'.[43] Botmasters use interfaces known as Command and Control (C&C) centres to send instructions to their botnets and retrieve data captured by the machines under their control. Although the largest botnets, such as BredoLab, Mariposa, Conficker, TDL4 or Cutwail, have been able to co-opt and

exploit millions of machines, the Shadowserver Foundation, a computer security NGO, tracks between 1600 and 2600 functional botnets at any given time, most of them much smaller and involving only several hundred or several thousand infected machines. Unlike traditional viruses, botnets do not entail any observable performance reduction for infected machines. Because of their stealth, versatility and scalability, they provide the technological infrastructure for online crimes and are considered to be one of the main tools facilitating a wide range of online crimes.[44] Once in operation, botnets can be used to launch Distributed Denial of Service attacks that overload and disable targeted websites,[45] distribute spam emails on a massive scale,[46] commit bank fraud and click fraud[47] or offer illegal proxy services.[48] Botnet marketplaces and forums can thus be seen as the nexuses of the online criminal ecosystem, in the sense that they bring together individuals who specialise in malware programming, deployment, obfuscation and exploitation. While this forum's accessibility and large beginner membership mean it is not the most attractive meeting place for sophisticated offenders – who prefer by-invitation and vetted forums – it remains the largest community of its kind and a fertile hub for recruitment, training, collaboration and commerce for aspiring and established hackers.[49] This particular forum's most frequently discussed topics involved Distributed Denial of Service attacks (19%), SQL injection (19%), shell code (16%), spam (14%), cross-site scripting (12%), brute-force attacks (11%) or HTML injection (9%), which represent different ways of compromising the integrity of a website or an internet server.[50]

Like many similar online forums or marketplaces, this forum relies on a reputation system that attempts to rate the trustworthiness of members through a large-scale replication of word-of-mouth processes.[51] Once members have demonstrated a basic understanding of the forum's rules and values, they can rate other members, either positively or negatively. Ratings, weighted according to an individual's position in the forum hierarchy, are either 1 ('l33t' members), 3 (more experienced 'ub3r' members), 5 (staff members who help run the forum) or 10 points (the exclusive group of three administrators). New users, known as '3pic' members, are not permitted to participate in the reputation system. Individual feedbacks accumulate over time to form the global reputation score of a member. The highest ranked members can amass thousands of points, while the lowest ranked can sink to negative ratings – the lead botnet hacker in our sample had accumulated 2330 reputation points, while the most untrustworthy botnet hacker had −708 points. Abusive reputation practices, such as trading in reputation scores, colluding with other members to artificially lower or inflate a member's reputation score and threatening or bribing others to obtain positive reputation scores are banned in principle. Administrators deal with infractions at their discretion and can apply a range of penalties, starting with removal of suspicious feedback and culminating with closing the offending account. An intermediate sanction is the 'Rep Fuck', where all of a member's positive reputation is permanently erased.

The data considered here covers a period of 27 months, from October 2009 to December 2011. Although the data may seem dated, and one could legitimately wonder whether the findings discussed below are still relevant, it is important to remember that we are not trying to dissect the latest botnet technology or money making scheme, but to identify patterns in the social structure of online trust. The forum from which the data was obtained remains five years later the largest hacking community accessible on the Internet, operated by the same committed administrator using the same status structure. Hacking techniques have not changed that much at the moderate level of expertise encountered in this community, botnets remain a major source of online harms, and from the regular monitoring conducted over the past few years, the social dynamics described in the

following paragraphs retain their currency. Technically, the data was collected using a customised web extraction application commonly called a scraper, which automatically captured and parsed the data contained in the profile pages of forum members who received reputation scores during the reference period. Our software gathered information on 29,985 hackers who had been evaluated by 9,177 of their peers through 449,478 discrete rating events. Although these numbers might seem large, they represent only a fraction of the forum's registered users – toward the end of our collection efforts, it listed roughly 250,000 members.[52] This gap between the number of registered users and the much smaller core of active participants suggests that, like other online communities, the vast majority of forum participants are 'lurkers', who, although they regularly read forum contents, do not post or post only minimally. Two studies of popular underground forums have established that lurkers represent on an average between 41.1% and 52.3% of registered users.[53] While lurking has sometimes been considered a free-riding behaviour detrimental to the quality of interactions in online communities, others have pointed out that this silent form of participation, possible because of the low costs of internet transactions, serves as a useful default behaviour in large online groups, keeping discussions from descending into a cacophony of low-value contributions.[54]

Data collected for each individual who had an active reputation report included nickname, sum of reputation points accrued and a list of each rating received during the reference period with a unique identifier for the member who provided the feedback, the number of positive or negative points allocated, the short one-sentence comment explaining the reason behind the rating and the event's timestamp (the date on which the feedback was provided). Hacker nicknames were replaced with a unique number ID to anonymise the data, which was then stored in an Excel file containing 449,478 lines (one for each feedback) and 4,045,302 cells. Most of the statistics presented in this article are descriptive, and once the data had been collected from thousands of this forum's user profiles and consolidated in a single file containing every reputation point provided over a period of 27 months, it became possible to pull the information needed to measure the distribution of positive and negative feedbacks between members, to chart the evolution of this ratio over time and to compare it with monthly activity levels, and to examine how members with different status levels allocated their feedbacks to potential co-offenders. To avoid amalgamating white hat (benevolent) and black hat (malicious) hackers into a single incoherent group, the data collected on this forum was dichotomised based on activities discussed on the botnet sub-forum. Two classes of hackers were created: general hackers, whose offending status is uncertain and should not be assumed based simply on their membership in this particular forum, and botnet hackers, whose involvement in illegal activities is much easier to establish. Any individual who had contributed to discussions about botnets – and could not therefore pursue his interest without breaking a criminal law – was considered to be a botnet hacker. The remaining individuals were considered to be part of the general hacker population, whose intentions and actions may have ranged from the most innocuous to resolutely malicious behaviours. Although we were most interested in the botnet category, comparisons with the general group help identify how trust and reputation patterns can be influenced by the specific resources and constraints associated with illicit action. Table 1 summarises the distribution of individuals among the two groups and their reputation ratings.

Beyond the mere adding up of positive or negative scores, it was essential to understand what led members to vouch for their peers' reputations, since members were not constrained by pre-defined criteria and might have awarded rankings for a broad range of very subjective reasons. Therefore, a qualitative analysis of a random sample of 25,000

Table 1. Descriptive statistics for the two sub-groups of hackers using the reputation system during the reference period.

	General hackers	Botnet hackers
Number of individuals who received a rating	20,768	9,127
Number of reputation ratings recorded	163,788	285,690
Average number of ratings received by individual	7.89	31.30
Number of individuals who provided a rating[a]	8,305	8,824
Average number of ratings provided by individual (for specific group)	19.72	32.37

a. The two categories are not mutually exclusive, as 7952 individuals rated both general and botnet hackers, while an additional 353 members rated general hackers exclusively and 872 members rated only botnet hackers.

feedback comments from the botnet hacker group was undertaken. We chose to focus on this group in order to maximise our limited resources, as coding such a large quantity of comments was time consuming. Each comment was coded to determine whether it reflected a positive or negative judgment on one of the following six dimensions: 1) business relationship; 2) general assessment of someone's contribution to the community at large; 3) specific assessment of someone's interaction with the feedback provider; 4) quality of technical skills as evaluated by the rater; 5) humorous, sarcastic or absurd comment on a member's actions or skills; 6) unreadable or meaningless comments, such as random strings of characters or written in foreign languages.

Results

The ratio of positive to negative feedback for the two communities suggests that, overall, the discussions and interactions on this forum result in overwhelmingly positive outcomes, as seen in Table 2. Positive feedbacks account for 86.3% of interaction outcomes for the botnet community, and 78.1% for the general community. Although members of the latter group seem to be less satisfied, at first glance the community still seems to be highly functional. These numbers might seem high for a group of people who barely

Table 2. Distribution of reputation across general and botnet hacker communities.

		Section					
		General hacker community		Botnet hacker community		Whole forum	
		Count	%	Count	%	Count	%
Reputation ratings	Positive	127,936	78.1	246,473	86.3	374,409	83.3
	Null	12,543	7.7	14,053	4.9	26,596	5.9
	Negative	23,309	14.2	25,164	8.8	48,473	10.8
	Total	163,788	100.0	285,690	100.0	449,478	100.0
Overall reputation scores	Positive	13,933	67.1	7,106	77.9	21,039	70.4
	Null	1,752	8.4	436	4.8	2,188	7.3
	Negative	5,083	24.5	1,585	17.4	6,668	22.3
	Total	20,768	100.0	9,127	100.0	29,985	100.0

Reputation ratings refer to the number of individual feedbacks given members of both communities, while reputation scores describe the performance of individual members.

know each other and who by definition trade in deception, but they are consistent with results from research on other reputation systems, in both legal and illicit settings. For example, the eBay rating system, which has been the subject of numerous empirical studies, has levels of positive feedback that are over 99%.[55] Similarly, members of Silk Road, a drug cryptomarket, gave the highest feedback score (5/5) to 96.5% of their interactions with sellers.[56] Such high levels of positive assessment provide a strong argument in favour of feedback systems, which appear to be effective at creating strong incentives for the self-regulation of behaviour and the subsequent emergence of trust in online communities. The significant difference in positive scores between the general hacker and the botnet hacker communities could thus be explained by the nature of the interactions in those two groups: while discussions in the general hacker community cover a wide range of topics, exchanging knowledge and expressing interests and opinions, including disagreements, the botnet hacker community is much more concerned with trade in botnet tools and the collaborative resolution of specific technical problems. The communities could therefore be inherently different because they appeal to different types of hackers, whose maturity, skills and motivations produce distinct trust ecologies. Alternatively, the variation in the distribution of positive and negative feedbacks could be an artefact of the increased utility of a reputation-rating tool in a market-oriented community whose members have more to lose in case of defection or malfeasance – members of the botnet group have stronger incentives to learn about potential customers and suppliers. We lean toward the second hypothesis, based on data presented in Table 4 showing that botnet hackers use the feedback mechanism twice as much as their generalist peers, but cannot definitively answer the question.

At the individual level, these aggregated ratings produce a global score that summarises past behaviours and provides a quick assessment of a member's trustworthiness. As mentioned above, toward the end of our sampling period, the most trusted botnet hacker had accumulated an overall reputation score of 2330 points, while the most erratic member had a negative score of −708 points. The pattern where the botnet hacking community seems to value reputation more than its general counterpart, because more feedbacks of a more positive nature are provided, still holds.

The unequal distribution of reputation

The ratings given general and botnet hackers and the aggregate scores give us some idea about the overall trustworthiness of the community's members. However, it is crucial to realise that, despite the large number of feedbacks analysed here, the vast majority of comments and ratings were directed toward a small core of very active members. When the distribution of ratings to individual members is measured, a common pattern reflecting a power law or Pareto distribution emerges. A power-law distribution is usually defined as a situation where 'the probability of measuring a particular value of some quantity varies inversely as a power of that value'.[57] This very common occurrence, which is encountered as often in the natural world as in the social sciences (the unequal accumulation of wealth among a tiny sliver of the world's population is a typical example), reveals configurations where a small minority of cases is responsible for a large majority of events.

In our research, as shown in Table 3, 5% of general and botnet hackers are the beneficiaries, respectively, of 47% and 37% of feedbacks, while at the other end of the distribution, 50% of our sample accounted for only 8% and 5% of allocated feedbacks. Even if the core 5% in the botnet community attracts a smaller share of feedbacks than its general counterpart, the concentration of reputation ratings among the 20% most

Table 3. Concentration of feedbacks among general and botnet hackers.

General hackers		Botnet hackers	
Top rated individuals (%)	Feedbacks received (%)	Top rated individuals (%)	Feedbacks received (%)
1	20	1	13
5	47	5	37
20	76	20	75
50	92	50	95
Remaining 50	8	Remaining 50	5

frequently assessed hackers is identical in both groups. In other words, out of a global membership of 250,000, an elite group of about 6,000 individuals (2.4%) were sufficiently active through mentoring, code sharing, selling tools and services, collaborating on projects, etc., to elicit 75% of all reputation scores awarded. In contrast, studies on the eBay feedback system report rating frequencies (the number of times a vendor or customer was rated regarding a transaction) oscillating between 33% and 78% of all transactions, implying a much higher level of engagement.[58] These numbers need to be kept in mind when assessing how effective online reputation systems are in fostering cooperation in illicit settings.

Reputation, scale and time: how trust decays

So far, we have treated the data only as a static source of information about the outcomes of a large number of hacker interactions as reflected in feedbacks provided through an online reputation system. But, since every feedback in our database is time-stamped, we can also chart how ratings vary over time and if such a reputation system is able to scale up easily and maintain high levels of trust between members, despite the community's size.

Figure 1 shows that the monthly volume of feedbacks provided grew regularly over 27 months for which data is available. Initially the two communities seem to be thriving and able to engage more and more people. However, although the vast majority of feedbacks remain positive, both groups show the same pattern of trust decay, with an increase in the number of ratings awarded by a growing number of members strongly correlated with a slow but inexorable decline in the percentage of positive comments. Over the reference period, the general hacker community's percentage of positive feedbacks drops from 97.2% to 75.8% (−21.4%), while the botnet community does only slightly better with a drop of 18.7% (from 99.1% to 80.4%). The instability in the distribution of reputation among these two hacker groups over time reflects the growing pains of online hacking communities, something botnet hackers also seem exposed to, despite their superior performance in other facets of the 'trust game'.

Reputation, hierarchy and specialisation: reporting biases

An online community that is experiencing rapid growth relies to a large extent on a formal hierarchy of administrators and moderators to ensure that rules are enforced and disputes settled.[59] As stated previously, at the time data was collected this forum operated under a

Figure 1. Evolution of feedbacks provided and the ratio of positive feedbacks.
(a) General hackers: R (−0.86**). (b) Botnet hackers: R (−0.73**). **$P < 0.01$

five-tiered pyramidal structure: at the bottom were newcomers, 3pic (epic) members, who were not allowed to participate in the reputation system and therefore don't appear in our statistics. L33t (shorthand for elite) members were slightly more experienced hackers who could assign positive or negative reputation scores of one. As shown in Table 4, they represented a bit more than 60% of both communities, but contributed less than one-third of feedbacks. One step up were the ub3r members, who reached this status only by invitation from other ub3r or above members and whose rating power was capped at three points (positive or negative) per feedback, with a limit of five feedbacks per day. Ub3r

Table 4. Feedbacks provided based on status within the forum.

Status	L33t	Ub3r	Staff member	Administrator
General hackers				
% of total number of individuals	61.51	38.21	0.24	0.05
% of total number of feedbacks provided	31.14	68.10	0.60	0.15
Average number of feedbacks provided	9.97	35.09	48.90	63.25
% of positive feedbacks	81.32	76.83	67.59	28.85
% of negative feedbacks	0.00	20.50	30.47	67.98
% of neutral feedbacks	18.68	2.67	1.94	3.16
Botnet hackers				
% of total number of individuals	63.15	36.56	0.25	0.05
% of total number of feedbacks provided	27.26	71.76	0.79	0.18
Average number of feedbacks provided	13.93	63.32	102.86	131.00
% of positive feedbacks	86.41	86.41	80.20	33.78
% of negative feedbacks	0.00	11.96	17.10	61.64
% of neutral feedbacks	13.59	1.63	2.70	4.58

members accounted for more than one-third of membership but posted around 70% of feedbacks in both groups. Finally, a small elite group of staff members and administrators could assign positive and negative scores of up to 5 and 10 points, respectively. Although they represented a tiny fraction of the membership, and less than one percent of feedbacks, their ratings distribution is noteworthy. The ratio of positive to negative evaluations follows a diverging trend in which negative comments play a much more significant role as the posters' authority increases, especially for administrators, for whom the general pattern is practically reversed: only one-third of their ratings are positive (28.85% for general hackers and 33.78% for botnet hackers), while two-thirds are a negative assessment of community members (respectively 67.98% and 61.64% of general and botnet hackers).

The mere existence of a reputation system does not guarantee spontaneous trust – far from it. As the numbers in Table 4 attest, forum administrators must exert a considerable amount of social control and coercion to maintain the community's cohesion and weed out members who abuse their peers' credulity. This shift in the nature of feedbacks provided as status increases also dispels the idea of a horizontal and egalitarian community, reflecting instead members' evolving priorities and expectations. At the bottom of the hierarchy, l33t members, eager to climb the status ladder that confers certain user privileges, favour positive ratings in the rational expectation that reciprocal feedbacks will boost their overall score. Such an instrumental strategy is then supplanted by the more critical approach of a core group of higher-status altruistic members who place the community's global performance and harmony above their own scores. By using the rating system primarily to highlight other members' untrustworthiness and unreliability, staff members and administrators demonstrate the limits of the rating system in fostering widespread self-regulation as well as the persistent need to impose order from above.

A comparison of these two hacker communities also uncovers very different dynamics: botnet interactions seem to elicit almost twice as many feedbacks as exchanges conducted on generalist sub-forums, a ratio that holds across the four status levels and might explain why positive feedbacks are relatively more resilient for the botnet group. The spike in distrust is only half as severe for the botnet community at the ub3r and staff member levels (11.96% and 17.10% vs. 20.50% and 30.47% in the general community),

and returns to the pattern observed in the general hacker community at the administrator level. This notable difference might indicate that the propensity to rate interactions more frequently and to share assessments more readily allows members of the botnet community to better identify and deter untrustworthy peers, up to a point.

The meaning of trust

No research on licit or illicit reputation systems has to our knowledge made use of qualitative analysis to complement the more classical statistical analysis, which typically tries to understand how feedback points are assigned based on the frequency of activities, such as sales or posts, nature of activity, status level, experience or social network of participants.[60] In order to understand why positive or negative feedbacks were assigned, we chose a different approach and relied instead on the one-sentence comment provided with each rating, which provide context and motivation for each assessment. Making sense of and coding each comment is a time-consuming activity that cannot be outsourced or even crowdsourced. A dense technical jargon, the lingo of hackers, known as 'l33t speak' because numbers are substituted for certain letters, makes it difficult for the uninitiated to understand the meaning of short sentences without additional context. We randomly extracted from our database a more manageable but still sizeable sample of 25,000 comments from the botnet community, which we manually coded. We quickly identified six major types of comments that represent the justification for a positive or negative evaluation.

The first category refers to the outcomes of business relationships, where tools or services were traded, and reflects the instrumental dimension of these exchanges: was the quality of the product as advertised and was the transaction conducted to the satisfaction of both parties? Under the second category of comments, participants rate their peers on the general contribution they make to the community, such as providing free tutorials or demonstrating a willingness to share knowledge with new members. Members who post low-quality content, repeatedly ignore the rules, try to infect or steal from their peers, or exhibit an arrogant or abusive attitude are frowned upon. In the third category of comments, we find assessments that result from more specific encounters between two members but exclude business exchanges. Very often, these comments relate to the helpfulness, maturity and ego of hackers when they engage in one-to-one discussions. This category contains the largest share of reciprocal feedbacks. A fourth category addresses the perceived technical expertise of community members: some are lauded for their assistance in setting up botnets or generously providing various pieces of malicious code, while others are chastised for lying about their technical skills or trying to hack into systems. Beyond business, social and technical feedbacks, a fifth category of comments provided a different perspective on members' interactions, highlighting the central role played by humour and sarcasm in this community, including a significant number of homophobic and misogynist references. Finally, one last category contains unreadable strings of characters or URLs referring to webpages or images that we were unable to access and therefore interpret. Table 5 provides verbatim examples for each of the six categories extracted from our database.

Analysing the distribution of comments across the six categories provides a more nuanced perspective on the dynamics motivating the assignment of positive, neutral and negative feedbacks in this hacker community, as can be seen in Table 6. Contrary to what was expected, business interactions did not play the most frequent role in establishing or undermining the reputation of botnet hackers. Social and mentoring skills were more

Table 5. Examples of comments for each of the six categories of feedbacks (for both positive and negative assessments, which are indicated in brackets).

Categories	Comments
Business relationship	We exchanged money smooth and quick transaction thanks! [+]
	Trusted seller: Thanks for the 500 + 100 extra shells [hacked machines];) [+]
	Don't trade, scammer, scammed me for 80$. [–]
	User scammed members on forum before, and sells public shit, beware when dealing with him, specially newbies. – do not trust under any circumstances! [–]
General contribution to the community	Great member around here. Very helpful and much kind. He really deserves the Rat [remote access trojan, a piece of malicious software that allows remote control of a machine] award as he asks for it. [+]
	Great user, great tutorial he made and thanks to him I'll be making $$$ out of this! [+]
	Wannabe, and as far as I am concerned you seem like a LQ [low quality] member. I already don't like you. [–]
	Trashing my sales thread, his malware version is backdoored, so also trying to infect members. Very childish member, I wouldn't trust him at all… [–]
Specific behaviour directed at feedback provider	Yeeee! Cool dude to talk to. Enjoys small talk:] [+]
	A grown-up, friendly and great member. [+]
	Actually answered my query with good info. [+]
	Lied to me. Blatantly. [–]
	Immature kid whos mad because i called out some bullshit in his 'terms of service' that would allow him to keep your money if he couldnt complete the job. [–]
Quality of technical skills	He made a bunch of crypters [a program to escape antivirus detection] = D [+]
	Thanks for the exploit [a unique way to take advantage of a vulnerability]. [+]
	Hacked my friends site, not cool. [–]
	Says he coded an exploit, he coded my ass. [–]
Sarcasm or no context	Clarinet all the way my friend. [+]
	Thinks im sexy, he has also maintained good grammar all along. [+]
	For the luls [fun derived at another's expense]. [–]
	Your grammar sucks balls my friend. [–]
Unreadable	infwawliabbbbble x. [+]
	♥♥♥ ✿◕ ‿ ◕✿ ♥♥♥ [+]
	Madmadmadmadmadmadmad. [–]
	<2<2<2<2<2<2. [–]

Technical terms are defined in brackets; nature of comment is indicated with + and – signs; spelling errors have not been edited but some terms have been changed without altering the meaning to protect posters' privacy.

highly valued, while their negative expression, such as being disrespectful, disregarding the tacit values of the community, or contributing low quality contents, accounted for more than 43% of negative feedbacks when general and individualised comments are combined. This finding confirms Holt and Lampke's observation that, in stolen data markets, the general attitude of sellers, not just the quality of their products, has a large effect on how they are seen by potential customers.[61] Given such a technically oriented group of offenders, one would expect technical skills to make a major contribution to

Table 6. Nature of comments used to support feedback ratings in the botnet community.

Categories	Positive feedbacks (%)	Neutral feedbacks (%)	Negative feedbacks (%)
Business relationship	9.38	13.00	11.72
General contribution to the community	13.41	28.00	22.95
Specific behaviour directed at feedback provider	24.58	42.89	20.57
Quality of technical skills	20.22	2.84	3.61
Sarcasm or no context	29.65	10.77	36.74
Unreadable	2.75	2.50	4.41
Total	100.00	100.00	100.00

reputations, but they explain only a little more than 20% of positive feedbacks and a mere 4% of negative ratings. Instead, and counter-intuitively for a community that appears to be so dependent on a reputation system to foster trust, sarcastic and humorous comments account for the largest share of feedbacks, 29.65% and 36.74% respectively of positive and negative ratings. They played a more modest role (10.77%) only in neutral feedbacks. Why would participants in an online forum dedicated to illegal activities apparently choose to undermine the effectiveness of the only trust-building tool at their disposal by filling their feedbacks with apparently useless comments?

Two alternative hypotheses can be offered at this stage. On the first, this liberal use of humour, sarcasm and invective is an expressive feature of the hacking community (which, in this subset, was predominantly male and juvenile), where reliable assessments about someone's trustworthiness are wrapped in multiple layers of inside jokes that make little sense to the outside world. Using a reputation-management tool developed for more mainstream uses may then create problems, as the large proportion of decontextualised comments deprives the community of valuable information on users' behaviour. The second hypothesis explains the intensive use of sarcasm differently, ascribing it a more transgressive function, in line with Coleman's depiction of tricksterism in the case of Anonymous.[62] Here, the 'lulz', which Coleman defines as 'laughter at someone else's expense',[63] is not merely a practice designed to make reputation systems more palatable to hackers but instead becomes a subversive tool whose main objective is to undermine such systems' utility by saturating them with meaningless, sarcastic or absurd comments. When things become too serious, instrumental or efficient, the lulz is summoned to restore the balance of a hacking culture that thrives on a healthy dose of chaos and mischief. In other words, the sacrosanct status of subversive behaviours in hacker communities trumps any effort to make interactions between members of these communities more orderly and predictable. These two hypotheses have significant implications for researchers, whose inferences about trust among online offenders will need to be more clearly supported in the future.

Discussion and conclusion

In her seminal review of the multiple functions, motivations and dimensions of trust, Barbara Misztal reminds us of 'the unimpressive record on the part of the social sciences in grasping its essence [...] without a great deal of effort being devoted to its conceptualization' or to its measurement.[64] Very modestly, we hope to have contributed in this article to an expanded understanding of the role trust and reputation play in online

offending and to have made more salient the considerable challenges faced by aspiring cybercriminals when searching for suitable co-offenders. We believe that a lot more empirical and conceptual work needs to be done to fully understand the social mechanisms at play, but the results presented in this article provide preliminary insights into three dimensions of online offenders' social interactions. First, they provide an overview of the distribution of trust in the world's largest known hacker community at general and individual levels. As expected, the data suggests that not all forum participants are considered to be equally trustworthy according to quantitative and qualitative assessments by their peers, but the scale of the imbalance was unexpected for a community so attached to the egalitarian principle. Despite the huge number of registered members and the availability of an easy-to-use and transparent reputation system, familiar to most internet users, only a tiny fraction of the forum membership (2.4%) participates in the vast majority (75%) of 'trust exchanges'. These more active members anchor the whole community, creating an environment in which almost 80% of recorded interactions result in satisfying outcomes, but they seem unable to engage the legions of lurkers, whose passivity dilutes the forum's usefulness. More specialised sub-communities, such as botnet hackers, seem able to nurture higher levels of trust, mainly because they also prompt higher levels of participation through more systematic ratings of peer trustworthiness. In that sense, from a cooperation perspective, specialist communities seem to enjoy a comparative advantage over generalist offender groups. One limit of this research is that it focuses on a large forum that is accessible to all. It therefore attracts mainly beginners, known as newbies or 'noobs', or hackers with mid-level skills, who might assess and allocate trust differently than their more experienced, profit-driven and successful peers. This latter group limits their trade to the more rarefied world of invitation-only forums, from which law enforcement investigators, security analysts, journalists and unqualified participants are excluded. One can see that many of the trust challenges identified in the literature and in this article would be less for members of the more restricted group. It thus seems urgent to expand the research on trust in online offending settings to include the more exclusive forums where high-level transactions are conducted.

Although it is difficult to determine the effectiveness of the reputation system implemented by this particular forum based solely on publicly available data, we can still draw three main conclusions from the results presented above. First, availability alone is not sufficient to encourage widespread adoption and routine use of such a system, even in an environment where the probabilities of malfeasance, mistakes or failures are high.[65] The reliability of a reputation system rests on the quality and quantity of information it makes available, but if a significant proportion of the community refuses or fails to share feedbacks about past transactions, its utility is severely curtailed and its relevance decreases accordingly. Second, the imbalance created by the limited use of the reputation system amplifies reporting biases[66] and the hierarchical positioning of raters significantly influences their assessments. The propensity to report positive outcomes is 2.81 times greater among generalist beginners (L33t hackers) than among forum administrators. As well, beginners did not report a single negative outcome, in contrast with administrators, where negative feedbacks account for 67.98% of their responses. The situation is very similar for botnet hackers. This major distortion might be attributed to fear of retaliation or ostracism among low- and mid-level members, or perhaps to the wilful blindness of 'newbies' still enthralled by the hacker mystique. In any case, the structural reporting biases identified on this forum suggest that reputation systems may be overrated as trust-building mechanisms that can be used by illicit communities. Finally, reputation systems do not seem able to protect against trust decay,[67] caused in this case by the rapid

expansion of the community and a sharp increase in the number of submitted ratings. In other words, and by contrast with what has been observed for online commerce platforms, reputation systems do not appear to make illicit communities more scalable and stable: the greater the number of feedbacks, the lower the number of positive outcomes reported and, by extension, the more fragile the trust between members.

Finally, for obvious methodological reasons outlined above, very few studies of online trust have considered the explicit motives behind a rater's positive or negative assessment. The qualitative analysis of a large number of sometimes poorly worded feedbacks is time consuming and introduces a level of complexity and ambiguity that makes modelling trust dynamics less compelling than when relying only on raw scores. We believe, however, that this approach is absolutely necessary to fully grasp the multi-layered dimensions of reputation and trust. A diverse set of behaviours, skills and attitudes trigger assessments of trustworthiness and enable or hinder cooperative endeavours. The comments extracted from this forum, which are very different from those found in market-oriented reputation systems, reveal a paradoxical situation in which the most frequently invoked reasons justifying peers' trustworthiness (or lack thereof) are grounded in humour and sarcasm, introducing high levels of uncertainty about the seriousness and reliability of the information provided. Roughly one-third of the comments accompanying numerical scores were questionable, non-sensical or based on seemingly trivial considerations. To a large extent they reflect the distinctive cultural premium hacker communities give to transgressive and subversive behaviours that produce laughter (the lulz). In this context, playfulness and helpfulness seem to trump 'craft(y)ness'[68] and effectiveness, considerably reducing the value of the reported reputations and the subsequent capacity of reputation systems to communicate trustworthiness. These responses also suggest that malicious hacking and online fraud should not be framed exclusively as market crimes driven by economic incentives and neoliberal ethics. Researchers need to acknowledge the distinct (and sometimes contradictory) values embedded in these communities and the unique ecology of trust they shape, which in turn sustains illicit computer-mediated cooperative practices. We are aware that the qualitative analysis provided here remains superficial, and that more insights could be gained from a more systematic discussion of the feedbacks provided to participants with very high and very low reputation scores, in order to better grasp the cumulative nature of trustworthiness and how one manages peer interactions over time to achieve high levels of recognition. A more granular understanding of the negative feedbacks received by the most despised contributors, and of the impact these negative ratings have on participants' careers, would also be welcome. Clearly, more work needs to be done at these opposing ends of the trust-mistrust continuum.

We do not consider how our findings might support law enforcement activities or the disruptive strategies deployed by private security companies.[69] Others have already suggested so-called 'trust attacks' to manipulate reputation systems, either by creating fake trustworthy identities that are then used to initiate fraudulent transactions (the Sybil attack) or by eradicating the high reputation scores of trusted participants through false defamation (the slander attack).[70] We can say, however, that the scale-free distribution of trust observed in this forum implies that it will be very resilient to the random loss of nodes (probably up to 80% in this case). Because such a large number of forum members contribute so little to the trust ecology (and, we assume by extension, to the hacking activities collectively undertaken as a result), their removal from the community through arrests and convictions is unlikely to affect the network's performance. Conversely, these networks are extremely vulnerable to selective attacks that target nodes that play a central role in maintaining the community's trust and connectivity.[71] Police operations based on opportunistic leads are unlikely to cripple

online offender communities that follow a similar pattern, while their chances of success are presumably much higher if they manage to focus on the 20% of individuals who account for 75% of trust exchanges. This finding might also explain why the most successful hackers shroud their activities in the secrecy of invitation-only forums, where strict vetting procedures are enforced before access is granted: open forums certainly provide them with more business and collaborative opportunities, but their transparent reputation systems also expose members who create the highest added value to the scrutiny of law enforcement investigators and intelligence analysts.

As a concluding comment, we would like to briefly reflect on the potential of 'computational' or 'Big Data' social science, as well as its hazards, for the study of online offending.[72] While our sample is rather modest in size compared to projects that use millions, or even billions, of data points, it still provides unusually rich quantitative and qualitative information on the structure and operations of illicit underground communities. The collection, extraction, sorting, coding and analysis of data require basic data-mining skills that are not yet routinely found in social science departments. However, once they become accessible, they open up fascinating new opportunities to answer research questions that were out of reach for our predecessors. The problem is that the seduction exercised by such large datasets and powerful tools can become intoxicating and obfuscate methodological challenges, as well as produce misleading inferences when analyses are not conducted with the required level of care.[73] For example, it would not have been farfetched for us to conclude that this malicious hacker community was relatively successful in harnessing the power of a basic reputation system in order to overcome the classical trust dilemma, even if scalability and stability over time remained of some concern. Indeed, had we not elected to complement our initial quantitative assessment with a qualitative analysis of raters' explicit motives, we would have missed what we believe is a significant insight that allows us to better understand the tension between organisational effectiveness, embodied by mainstream reputation mechanisms, and the cultural values of underground communities that resist the adoption of such utilitarian technologies. In other words, data mining would have been useless – counterproductive even – without a healthy dose of data meaning.

Acknowledgements

The authors would like to thank Peter Grabosky, Tom Holt, Jonathan Lusthaus, Carlo Morselli and Chad Whelan for their valuable feedback. The errors and mistakes remain their own.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was partly funded by the Social Sciences and Humanities Research Council of Canada [Canada Research Chair number 225007].

Notes

1. Castells, *Rise Network Society*; Benkler, *The Wealth of Networks*; Rainie and Wellman, *Networked*; and Leukfeldt et al., "Cybercriminal Networks."
2. Anderson et al., "Measuring the Cost of Cybercrime."
3. Gambetta, *Codes of the Underworld*; Lusthaus, "Trust World Cybercrime."

4. Coleman, *Hacker, Hoaxer, Whistleblower, Spy*; Décary-Hétu et al., "Policing the Hackers."
5. Yip et al., "Trust Among Cybercriminals"; Holt et al., "Examining Risk Reduction Strategies."
6. Poulsen, *Kingpin*, 162.
7. Glenny, *DarkMarket*.
8. Krebs, *Spam Nation*, 139.
9. Bearman and Hanuka, "Rise and Fall of Silk Road."
10. Morselli, *Inside Criminal Networks*, 64.
11. McIntosh, *The Organization of Crime*.
12. Yip et al., "Trust Among Cybercriminals."
13. Gambetta, "Mafia," 168.
14. Lusthaus, "Trust World Cybercrime."
15. Mell, "Reputation," 33; Yip et al., "Trust Among Cybercriminals," 520.
16. Tilly, *Trust and Rule*, 12.
17. Luhmann, *Trust and Power*; Dasgupta, "Trust as a Commodity"; Good, "Individuals"; Misztal, *Trust in Modern Societies*; Smith Ring, "Fragile and Resilient Trust"; Jøsang et al., "A Survey of Trust."
18. Leukfeldt et al., "Cybercriminal Networks."
19. Misztal, *Trust in Modern Societies*, 120.
20. Herley and Florêncio, "Nobody Sells Gold," 35.
21. Holt et al., "Examining Risk Reduction Strategies," 10.
22. Cova et al., "No Free Phish."
23. Franklin et al., "Inquiry Into the Nature"; Herley and Florêncio, "Nobody Sells Gold."
24. Yip et al., "Trust Among Cybercriminals," 520.
25. Franklin et al., "Inquiry Into the Nature," 8; Holt and Lampke, "Exploring Stolen Data Markets," 43; Yip et al., "Trust Among Cybercriminals," 527.
26. Lusthaus, 'Trust World Cybercrime', 87.
27. Holt and Lampke, "Exploring Stolen Data Markets," 42; Yip et al., "Trust Among Cybercriminals," 528; Holt et al., "Examining Risk Reduction Strategies," 11.
28. Lusthaus, "Trust World Cybercrime," 89; Holt et al., "Examining Risk Reduction Strategies," 13.
29. Resnick and Zeckhauser, "Trust Among Strangers."
30. Resnick et al., "Reputation Systems."
31. Jøsang et al., "A Survey of Trust"; Zheng and Jin, "Online Reputation Systems"; Diekmann et al., "Reputation Formation."
32. Mell, "Reputation."
33. Herley and Florêncio, "Nobody Sells Gold."
34. See note 14 above.
35. See note 32 above.
36. Lusthaus, "Trust World Cybercrime."
37. Dupont, "Skills and Trust."
38. Yip et al., "Trust Among Cybercriminals"; Holt and Lampke, "Exploring Stolen Data Markets"; Holt et al., "Examining Risk Reduction Strategies."
39. Franklin et al., "Inquiry Into the Nature."
40. Décary-Hétu and Dupont, "Reputation in a Dark Network"; Monsma et al., "Partners in Cybercrime."
41. Afroz et al., "Honor Among Thieves."
42. Young et al., "Hacking Into Mind Hackers."
43. Abu Rajab et al., "Multifaceted Approach Botnet Phenomenon," 42; Gu et al., "BotMiner," 139.
44. Namestnikov, *The Economics of Botnets*; Anderson et al., "Measuring the Cost of Cybercrime," 7; Décary-Hétu and Dupont, "Reputation in a Dark Network."
45. Goncharov, *Russian Underground 101*.
46. McCoy et al., "PharmaLeaks," 12.
47. Trusteer, *Measuring Effectiveness Antivirus*; Miller et al., "What's Clicking What?"
48. Hu et al., "Auto-detection of Redirection Botnets."
49. Imperva, *Hacker Trend Report #5*; Ablon et al., *Markets for Cybercrime Tools*", 7; Holt et al., "Examining Risk Reduction Strategies," 4.

50. Imperva, *Hacket Trend Report #13*, 4. A total sample of 439,587 discussions was analysed by the Imperva Team, using the forum's advanced search capabilities to measure the frequency of occurrence for the most common keywords associated with hacking techniques.
51. Resnick et al., "Reputation Systems"; Dellarocas, "Digitization Word-of-mouth."
52. Imperva, *Hacker Trend Report #5*.
53. Motoyama et al., "Analysis Underground Forums," 72; Afroz et al., "Honor Among Thieves."
54. Kollock and Smith, "Managing Virtual Commons"; Nonnecke and Preece, "Lurker Demographics."
55. Resnick and Zeckhauser, "Trust Among Strangers," 139; Dellarocas, "Digitization Word-of-mouth," 1411.
56. Christin, "Traveling the Silk Road," 218.
57. Newman, "Power Laws," 323.
58. Dellarocas et al., "Self-interest, Reciprocity, and Participation"; Jian et al., "Prevalence Reciprocation Feedback."
59. Yip et al., "Trust Among Cybercriminals," 531.
60. Décary-Hétu and Dupont, "Reputation in a Dark Network"; Monsma et al., "Partners in Cybercrime."
61. Holt and Lampke, "Exploring Stolen Data Markets," 46.
62. Coleman, *Hacker, Hoaxer, Whistleblower, Spy.*
63. Ibid., 31.
64. Misztal, *Trust in Modern Societies*, 13.
65. Tilly, *Trust and Rule*.
66. Dellarocas and Wood, "The Sound of Silence."
67. Dupont, "Skills and Trust."
68. Steinmetz, "Craft(y)ness."
69. Lerner, "Microsoft the Botnet Hunter."
70. Franklin et al., "Inquiry Into the Nature," 13; Mell, "Reputation," 25.
71. Albert et al., "Error and Attack Tolerance."
72. Boyd and Crawford, "Critical Questions for Big Data"; Kitchin, "Big Data, New Epistemologies."
73. Tufekci, "Big Questions for Social Media Big Data"; Chan and Bennett Moses, "Big Data Challenging Criminology?"

Notes on contributors

Benoît Dupont is a Professor of Criminology at the Université de Montréal, where he also holds the Canada Research Chair in Security and Technology. He is the Scientific Director of the Smart Cybersecurity Network (SERENE-RISC), one of Canada's Networks of Centres of Excellence. SERENE-RISC brings together social scientists, computer scientists as well as government and industry stakeholders to mobilise evidence-based knowledge for a safer and more resilient digital ecosystem. His research interests include the co-evolution of technology and crime, as well as the polycentric governance of security.

Anne-Marie Côté is a PhD Candidate in Criminology at the Université de Montréal. She is the first PhD student at the Cyber Criminology Lab (LC2) and a research assistant at the International Center for Comparative Criminology (CICC). Her interests include cybercrime, new technologies and cybersecurity issues. Her thesis examines how different conceptualisations of cybersecurity issues (by private firms, government and hybrid organisations) impact public policies. She is also involved in other research projects, such as being part of a computer engineering team at École Polytechnique of Montréal and assisting a project based on the use of social media by law enforcement.

Claire Savine is an Information Systems engineering student at Polytech Nantes (France). She specializes in the field of Business Intelligence, which includes database management, statistics and data mining. She is interested in discovering models or correlations in data and in finding the best representation to visualize results. Claire is currently working in Paris in the CRM field where she is developing applications to increase the quality of relationships between companies and their customers.

David Décary-Hétu is an Assistant Professor at the School of Criminology of the Université de Montréal and a regular researcher at the International Centre for Comparative Criminology. His interests include online illicit markets that sell stolen financial information, drugs and stolen intellectual property. His main current project focuses on the structure of online illicit drug markets known as cryptomarkets.

Bibliography

Ablon, L., M. C. Libicki, and A. A. Golay. *Markets for Cybercrime Tools and Stolen Data: Hacker's Bazaar*. Santa Monica, CA: RAND Corporation, 2014.

Abu Rajab, M., 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*, edited by P. Barford, 41–52. Rio de Janeiro: Association for Computing Machinery, 2006.

Afroz, S., V. Garg, D. McCoy, and R. Greenstadt. "Honor Among Thieves: A Common's Analysis of Cybercrime Economies". Paper presented at the *eCrimes Research Summit*, San Francisco, September 16–19, 2013.

Albert, R., H. Jeong, and A.-L. Barabási. "Error and Attack Tolerance of Complex Networks." *Nature* 406, July 27 (2000): 378–382. doi:10.1038/35019019.

Anderson, R., C. Barton, R. Böhme, R. Clayton, M. J. G. Van Eeten, M. Levi, T. Moore, and S. Savage. "Measuring the Cost of Cybercrime". Paper presented at the *11th Workshop on the Economics of Information Security*, Berlin, June 25–26, 2012. http://weis2012.econinfosec.org/papers/Anderson_WEIS2012.pdf.

Bearman, J., and T. Hanuka. 2015. "The Rise and Fall of Silk Road". *Wired Magazine*, May. http://www.wired.com/2015/04/silk-road-1/.

Benkler, Y. *The Wealth of Networks*. New Haven, CT: Yale University Press, 2006.

Boyd, D., and K. Crawford. "Critical Questions for Big Data: Provocations for a Cultural, Technological and Scholarly Phenomenon." *Information, Communication & Society* 15, no. 5 (2012): 662–679. doi:10.1080/1369118X.2012.678878.

Castells, M. *The Rise of the Network Society*. Malden, MA: Blackwell, 1996.

Chan, J., and L. Bennett Moses. "Is Big Data Challenging Criminology?" *Theoretical Criminology* 20, no. 1 (2016): 21–39. doi:10.1177/1362480615586614.

Christin, N. "Traveling the Silk Road: A Measurement Analysis of a large Anonymous Online Marketplace." In *Proceedings of the 22nd international conference on World Wide Web*, edited by R. Baeza-Yates and S. Moon, 213–224. Geneva: Association for Computing Machinery, 2013.

Coleman, G. *Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous*. New York, NY: Verso, 2014.

Cova, M., C. Kruegel, and G. Vigna. "There Is No Free Phish: An Analysis of 'Free' and Live Phishing Kits." In *Proceedings of the 2nd USENIX Workshop on Offensive Technologies*, edited by D. Boneh, T. Garfinkel, and D. Song. San Jose, CA: USENIX Association. Accessed July 28, 2008. https://www.usenix.org/legacy/events/woot08/tech/full_papers/cova/cova.pdf

Dasgupta, P. "Trust as a Commodity." In *Trust: Making and Breaking Cooperative Relations*, edited by D. Gambetta, 48–72. New York, NY: Basil Blackwell, 1988.

Décary-Hétu, D., and B. Dupont. "Reputation in a Dark Network of Online Criminals." *Global Crime* 14, no. 2–3 (2013): 175–196. doi:10.1080/17440572.2013.801015.

Décary-Hétu, D., B. Dupont, and F. Fortin. "Policing the Hackers by Hacking Them: Studying Online Deviants in IRC Chat Rooms." In *Networks and Network Analysis for Defence and Security*, edited by A. Masys, 63–82. New York, NY: Springer, 2014.

Dellarocas, C. "The Digitization of Word-of-Mouth: Promise and Challenges of Online Feedback Mechanisms." *Management Science* 49, no. 10 (2003): 1407–1424. doi:10.1287/mnsc.49.10.1407.17308.

Dellarocas, C., M. Fan, and C. Wood. "Self-Interest, Reciprocity, and Participation in Online Reputation Systems". *MIT Sloan Working Papers*, 2004. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=585402.

Dellarocas, C., and C. Wood. "The Sound of Silence in Online Feedback: Estimating Trading Risks in the Presence of Reporting Bias." *Management Science* 54, no. 3 (2008): 460–476. doi:10.1287/mnsc.1070.0747.

Diekmann, A., B. Jann, W. Przepiorka, and S. Wehrli. "Reputation Formation and the Evolution of Cooperation in Anonymous Online Markets." *American Sociological Review* 79, no. 1 (2014): 65–85. doi:10.1177/0003122413512316.

Dupont, B. "Skills and Trust: A Tour Inside the Hard Drives of Computer Hackers." In *Illicit Networks*, edited by C. Morselli, 195–217. Oxford: Routledge, 2013.

Franklin, J., V. Paxson, A. Perrig, and S. Savage. "An Inquiry Into the Nature and Cause of the Wealth of Internet Miscreants." In *Proceedings of the 14th ACM Conference on Computer and Communications Security*, edited by S. de Capitani di Vimercati and P. Syverson, 375–388. New York: ACM, 2007.

Gambetta, D. "Mafia: The Price of Distrust." In *Trust: Making and Breaking Cooperative Relations*, edited by D. Gambetta, 158–175. New York, NY: Basil Blackwell, 1988.

Gambetta, D. *Codes of the Underworld: How Criminals Communicate*. Princeton, NJ: Princeton University Press, 2009.

Glenny, M. *DarkMarket: How Hackers Became the New Mafia*. New York, NY: Vintage, 2012.

Goncharov, M. *Russian Underground 101*. Cupertino, CA: Trend Micro, 2012. http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-russian-underground-101.pdf.

Good, D. "Individuals, Interpersonal Relations, and Trust." In *Trust: Making and Breaking Cooperative Relations*, edited by D. Gambetta, 31–48. New York, NY: Basil Blackwell, 1988.

Gu, G., R. Perdisci, J. Zhang, and W. Lee. "BotMiner: Clustering Analysis of Network Traffic for Protocol- and Structure-Independent Botnet Detection." In *Proceedings of the 17th Conference on Security Symposium*, edited by P. van Oorschot, 139–154. San Jose, CA: USENIX Association, 2008.

Herley, C., and D. Florêncio. "Nobody Sells Gold for the Price of Silver: Dishonesty, Uncertainty and the Underground Economy." In *Economics of Information Security and Privacy*, edited by T. Moore, D. J. Pym, and C. Ioannidis, 33–53. New York, NY: Springer, 2010.

Holt, T. J., and E. Lampke. "Exploring Stolen Data Markets Online: Products and Market Forces." *Criminal Justice Studies: A Critical Journal of Crime, Law and Society* 23, no. 1 (2010): 33–50. doi:10.1080/14786011003634415.

Holt, T. J., O. Smirnova, Y. T. Chua, and H. Copes. "Examining the Risk Reduction Strategies of Actors in Online Criminal Markets." *Global Crime* 16, no. 2 (2015): 81–103. doi:10.1080/17440572.2015.1013211.

Hu, X., M. Knysz, and K. G. Shin. "Rb-Seeker: Auto-Detection of Redirection Botnets". Paper presented at the *Network & Distributed System Security Symposium*, San Diego, CA, February 8–11, 2009. http://www-personal.umich.edu/~huxin/papers/xin_RBSeeker.pdf.

Imperva. *Monitoring Hacker Forums – Hacker Intelligence Initiative Monthly Trend Report #5*. Redwood City, CA: Imperva, 2011. http://www.imperva.com/docs/HII_Monitoring_Hacker_Forums.pdf.

Imperva. *Monitoring Hacker Forums – Hacker Intelligence Initiative Monthly Trend Report #13*. Redwood City, CA: Imperva, 2012. http://www.imperva.com/docs/HII_Monitoring_Hacker_Forums_2012.pdf.

Jian, L., J. K. MacKie-Mason, and P. Resnick. "I Scratched Yours: The Prevalence of Reciprocation in Feedback Provision on eBay." *The B.E. Journal of Economic Analysis & Policy* 10, no. 1 (2010): 1935–1682. doi:10.2202/1935-1682.2470.

Jøsang, A., R. Ismail, and C. Boyd. "A Survey of Trust and Reputation Systems for Online Service Provision." *Decision Support Systems* 43, no. 2 (2007): 618–644. doi:10.1016/j.dss.2005.05.019.

Kitchin, R. "Big Data, New Epistemologies and Paradigm Shifts." *Big Data & Society* 1, no. 1 (2014): 1–12. doi:10.1177/2053951714528481.

Kollock, P., and M. Smith. "Managing the Virtual Commons: Cooperation and Conflict in Computer Communities." In *Computer Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, edited by S. C. Herring, 109–128. Amsterdam: John Benjamins, 1996.

Krebs, B. *Spam Nation: The Inside Story of Organized Cybercrime – From Global Epidemic to your Front Door*. Naperville, IL: Sourcebooks, 2014.

Lerner, Z. "Microsoft the Botnet Hunter: The Role of Public-Private Partnerships in Mitigating Botnets." *Harvard Journal of Law & Technology* 28, no. 1 (2014): 237–261. http://jolt.law.harvard.edu/articles/pdf/v28/28HarvJLTech237.pdf.

Leukfeldt, E. R., E. R. Kleemans, and W. P. Stol. "Cybercriminal Networks, Social Ties and Online Forums: Social Ties Versus Digital Ties Within Phishing and Malware Networks." *British Journal of Criminology* (2016). doi:10.1093/bjc/azw009.

Luhmann, N. *Trust and Power*. Chichester: Wiley, 1979.

Lusthaus, J. "Trust in the World of Cybercrime." *Global Crime* 13, no. 2 (2012): 71–94. doi:10.1080/17440572.2012.674183.
McCoy, D., 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 *Proceedings of the 21st USENIX Security Symposium*, edited by T. Kohno, 1–16. Bellevue, WA: USENIX Association, 2012.
McIntosh, M. *The Organisation of Crime*. Macmillan: London, 1975.
Mell, A. "Reputation in the Market for Stolen Data". Discussion Paper no. 611. Department of Economics, University of Oxford, 2012.
Miller, B., P. Pearce, C. Grier, C. Kreibich, and V. Paxson. "What's Clicking What? Techniques and Innovations of Today's Clickbots." In *Detection of Intrusions and Malware, and Vulnerability Assessment*, edited by T. J. Holt and H. Bos, 164–183. Berlin: Springer, 2011.
Misztal, B. *Trust in Modern Societies: The Search for the Bases of Social Order*. Cambridge: Polity Press, 1996.
Monsma, E., V. Buskens, M. Soudijn, and P. Nieuwbeerta. "Partners in Cybercrime." In *Advances in Cyber Security: Technology, Operations and Experiences*, edited by D. F. Hsu and D. Marinucci, 146–172. Bronx, NY: Fordham University Press, 2013.
Morselli, C. *Inside Criminal Networks*. New York, NY: Springer, 2009.
Motoyama, M., D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker. "An Analysis of Underground Forums." In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement*, edited by P. Thiran and W. Willinger, 71–79. New York: ACM, 2011.
Namestnikov, Y. *The Economics of Botnets*. Moscow: Kaspersky Lab, 2009. http://securelist.com/large-slider/36257/the-economics-of-botnets/.
Newman, M. E. J. "Power Laws, Pareto Distributions and Zipf's Law." *Contemporary Physics* 46 (2005): 323–351. doi:10.1080/00107510500052444.
Nonnecke, B., and J. Preece. "Lurker Demographics: Counting the Silent." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, edited by T. Turner and G. Szwillus, 73–80. New York, NY: ACM, 2000.
Poulsen, K. *Kingpin: How One Hacker Took Over the Billion-Dollar Cybercrime Underground*. New York, NY: Crown Publishers, 2011.
Rainie, H., and B. Wellman. *Networked: The New Social Operating System*. Cambridge, MA: MIT Press, 2012.
Resnick, P., K. Kuwabara, R. Zeckhauser, and E. Friedman. "Reputation Systems." *Communications of the ACM* 43, no. 12 (2000): 45–48. doi:10.1145/355112.355122.
Resnick, P., and R. Zeckhauser. "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System." In *The Economics of the Internet and E-commerce*, edited by M. R. Baye, 127–157. Amsterdam: Elsevier Science, 2002.
Smith Ring, P. "Fragile and Resilient Trust and their Roles in Economic Exchange." *Business & Society* 35, no. 2 (1996): 148–175. doi:10.1177/000765039603500202.
Steinmetz, K. "Craft(y)ness: An Ethnographic Study of Hacking." *British Journal of Criminology* 55, no. 1 (2015): 125–145. doi:10.1093/bjc/azu061.
Tilly, C. *Trust and Rule*. Cambridge: Cambridge University Press, 2005.
Trusteer. *Measuring the In-The-Wild Effectiveness of Antivirus Against ZeuS*. New York, NY: Trusteer, 2009.
Tufekci, Z. "Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls." In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, edited by M. de Choudhury, B. Hogan, and A. Oh, 505–514. Palo Alto, CA: AAAI Press, 2014.
Yip, M., C. Webber, and N. Shadbolt. "Trust among Cybercriminals? Carding Forums, Uncertainty and Implications for Policing." *Policing & Society* 23, no. 4 (2013): 516–539. doi:10.1080/10439463.2013.780227.
Young, R., L. Zhang, and V. R. Prybutok. "Hacking Into the Minds of Hackers." *Information Systems Management* 24, no. 4 (2007): 281–287. doi:10.1080/10580530701585823.
Zheng, W., and L. Jin. "Online Reputation Systems in Web 2.0 Era." In *Value Creation in E-business Management*, edited by M. L. Neslon, M. J. Shaw, and T. J. Strader, 296–306. Berlin: Springer-Verlag, 2009.