MAX PLANCK INSTITUTE FOR FOREIGN AND INTERNATIONAL CRIMINAL LAW

MODELLING AND SIMULATION FOR CYBERCRIME POLICY ANALYSIS[1]

SOUMYO D. MOITRA[2]

Contents:

[1] Much of this work was done when the author was an Alexander von Humboldt Fellow at the Max-Planck-Institute for Criminal Law, Freiburg, Germany from July to December 2004. The author would like to thank the Humboldt Foundation for the Fellowship and also Prof. Hans-Joerg Albrecht, Director of the Max-Planck-Institute, for his encouragement and support.

[2] Professor of Operations Management, Indian Institute of Management Calcutta, Diamond Harbour Road, Joka, Kolkata 700 104. India.

Introduction

The Internet is now regarded as one of the most significant developments in information technology in recent times. Its phenomenal growth, diversity of uses and its applications for information dissemination have been widely reported in the popular press, in journal articles and in books. Thus the terms cyberspace and e-commerce have become ubiquitous. Not only has the number of people who get "on-line" increased dramatically, the use of the Internet by businesses has also grown significantly in importance. For instance, the number of people online in 2003 has been estimated to be between 450 million to 620 million. The growth rate has been particularly phenomenal, rising from 369 million users in 2000 to a projected 940 million users in 2004 (GlobalReach.com, eMarketeer.com). Internet penetration in Europe is now estimated to be over 40%, up from 18% in 2000 (Netwatch.com). It is projected 190 million Internet users in Europe in 2004, up from about 94 million in 2001 (emarketeer.com), while other estimates suggest over 200 million users in the EU in 2004 (Internetworldstats.com). Along with this growth, many risks associated with the Internet, have emerged. In addition to various socio-cultural issues that have arisen, cyberspace has become a new arena for criminal activities. This is because the Internet is a very open network and it is very easy to access computer systems on it. Thus almost anybody can get on it, including malicious hackers. Consequently, any system connected to the Internet (even indirectly) is potentially vulnerable to intrusions and attacks. These intrusions and illegal activities that we shall call "cybercrimes" constitute a significant threat to the functioning of the Internet and a major inhibitor of e-commerce.

In fact, most indications point to substantial illegal/criminal activity over the Internet including many *new kinds* of crime and deviance, such as virus attacks that destroy computer files and systems, and distributed denial-of-service (DDoS) attacks that can paralyze communication links to an organisation's information system. These Internet crimes and security breaches of information systems have now been well documented in reports on cybercrime[3] and it is clear that they present a threat that is serious

[3] BJS 2002 *Cybercrime against Business.* Bureau of Justice Statistics, Department of Justice, US; CSI 2004 *The 2004 CSI/FBI Computer Crime and Security Survey*. Computer Security Institute; DTI 2004 *Information Security Breaches Survey:*

enough to warrant further study and the development of preventive measures.

The threat is potentially serious in many ways since increasingly societies are becoming information-based and any threat to safe communication endangers the functioning of society. According to one scholar, "As never before, industrial societies are dependent upon complex data processing and telecommunications systems. Damage to, or interference with, any of these systems can lead to catastrophic consequences."[4] The Commission of the European Communities also notes that "information and communication infrastructures have their own vulnerabilities and offer new opportunities for criminal conduct" and that "these offences constitute a threat to industry investment, and assets, and to safety and confidence in the information society." Finally national defence systems are also dependent on computer networks and hence threats to the networks are a matter of national security to most countries.

In the US, the first country to have cyberlaws[5], cyberspace security is an important issue as witnessed by the number of agencies and organizations devoting resources to it. Thus the President's Critical Infrastructure Protection Board has formulated "The National Strategy to Secure Cyberspace." The new Department of Homeland Security has network security as one of its major concerns. Various other organizations such as the Department of Justice (including the FBI) and the Department of Commerce are working to enhance network security. In Europe, the Council of Europe has developed a Convention on Cybercrime which the European Union countries have generally adopted[6]. The EU commission has also emphasised the importance of ensuring security on the Internet through a number of docu-

---

Technical Report. Department of Trade and Industry, UK; NHTCU 2004 *HI-TECH CRIME: The Impact on UK Business*. National Hi-Tech Crime Unit, UK, among others.

[4] Grabosky, P. 2000. Computer Crime: Challenge to Law Enforcement. *Law Enforcement Review*, N100, 31- 38.

[5] Hollinger, R. C. 1997. *Crime, Deviance and the Computer*. Dartmouth Publishing, Aldershot, U.K.; Brenner, S.W. 2004. U.S. Cybercrime Law: Defining Offences, *Information Systems Frontiers* 6:2, 115-132.

[6] *Convention on Cybercrime (Council of Europe, Budapest, 2001).*

ments[7]. In fact most countries of the world either have some form of cyberlaws already enacted or are in the process of developing them. In addition, many countries that already have some cyberlaws are extending them or updating them to evolving conditions. There are also many national and international private or semi-private organizations that are concerned with maintaining order on the Internet[8].

1. Policy Analysis Issues

Thus the Internet has attracted the attention of policy makers worldwide and a significant portion of the current literature on cybercrime has been concerned with legal issues where the focus is on the laws related to computer-related crime and on specific cases related to cybercrimes and hackers, rather than on empirical analysis[9]. Legal, policy and governance issues

---

[7] The EU has established a *Forum on Cybercrime* to ensure the "safer use of the Internet." The Commission of the European Communities also notes that "information and communication infrastructures have their own vulnerabilities and offer new opportunities for criminal conduct" and that "these offences constitute a threat to industry investment, and assets, and to safety and confidence in the information society" (Commission of the European Communities 2001).

[8] There are now Computer Emergency Response Teams (CERTS) in many countries around the world. They are designated to receive reports of network intrusions and are also active in preventing network attacks. The original CERT is CERT/CC at the Software Engineering Institute at Carnegie-Mellon University. In Europe, most countries have a CERT and they generally work together to develop a CERT infrastructure within Europe called "EuroCERT." The Deutsches Forschungs Netz has been helping to coordinate this and further information can be obtained from info@dfn-cert.de.

[9] Akdeniz, Y., Walker, C.P. and Wall, D.S. 2000, *The Internet, Law and Society*, Longmans, London; Clifford, R.D. (Ed.) 2001 *Cybercrime: The Investigation, Prosecution and Defence of a Computer-related Crime*. Carolina Academic Press, Durham, NC; Edwards, L. and Waelde, C. (Eds.) 1997. *Law and the Internet*, Hart, Oxford; Rosenoer, J. 1996. CyberLaw: The Law of the Internet. Springer, New York; Westby, J. C. 2003. *International Guide to Combating Cybercrime*, American Bar Association, Chicago.

have also been discussed by a number of authors[10]. The legal discussions naturally bring up many issues and questions that are not fully resolved yet but which could be investigated through empirical analysis. One issue is the efficacy of any legal institutions that users of the Internet can develop on their own[11]. For example, it has been argued that system operators and users are developing a set of robust rules, or at least "netiquette," for maintaining order on the Internet. However, the extent to which this strategy is successful is not known empirically and is surely open to debate. Appropriate policy analysis can examine scenarios that incorporate such controls and assess their effectiveness and limitations.

For traditional criminals, incapacitation is often considered to be one of the goals of sentencing. In the realm of cyber activities, there can actually be two kinds of incapacitation: one is the imprisonment of convicted cybercriminals (including denying them access to the Internet) and the second is "banishment" from cyberspace, where ISPs and other institutions prevent convicted (or known) cybercriminals from using the Internet[12]. While this sounds attractive in theory, it may not be possible to implement such a strategy in practice given the various ways one can access the Internet these days[13]. Also, it has been suggested that a multi-level, formal and informal control mechanism has developed that is keeping cybercrime in check[14]. Finally, the matter of the group behaviour of cybercriminals has also been raised in the literature since there have been indications and anecdotal evi-

---

[10] Kahin, B. and Keller, J.H. 1997. *Coordinating the Internet.* The MIT Press, Cambridge, MA; Kahin, B. and Nesson, C. 1998. *Borders in Cyberspace*. The MIT Press, Cambridge, MA.

[11] See Johnson, D.R. and Post, D. 1996. Law and Borders – The Rise of Law in Cyberspace, *Stanford Law Review*, 48, 1367-1402. They ague that the Internet requires new rules much as Lex Mercatoria was required for the new age of mercantilism, and that online users and service providers can develop a self-governance system. They cite the administration of the domain name system and the power of system operators to control access to the Internet.

[12] Johnson and post, op. cit. pp 1388-9.

[13] For example, anyone can go to cybercafes, public institutions like libraries or colleges, use a friend's computer or subscribe to Internet services with an alias, etc.

[14] Grabosky, P. 2001. Computer Crime: A Criminological Overview. *Forum on Crime and Society*. 1,1, 35-53; Wall, D.S. (Ed.) 2003. *Cyberspace Crime*. Ashgate/Dartmouth, Aldershot, UK.

dence about such behaviour, but we have very little concrete information on its modalities. This is an area where dependable data will probably never be available, and we may need to combine the little information we have with plausible models to understand how group interactions affect the generation of cybercrime.

Thus there are a number of assertions in the cyber-law and cyber-policy that can be empirically tested although very little has been done so far. Such research is urgently needed to support both policy-making and the development of a comprehensive theory of cybercrime but unfortunately the literature does not suggest any specific empirical analysis. Nor is there a sufficiently developed body of theory about cybercrime and this is largely a result of the absence of rigorous empirical studies of cybercrime. We would argue that empirical work is a necessary prerequisite to developing viable theory. In the sciences, it is theory that generally follows observations rather than the reverse. Almost all theories in science, for example from Newton's Gravitational Theory to Planck's Quantum Theory, have been formulated to explain empirical observations[15].

What is clear is that the concept of cybercrime has indeed been established in the legal context (even though there are disagreements over its exact nature and definition). While its existence had been recognised, what society's policy should be towards it is very much an open issue and the response to cybercrime has been widely discussed, including the literature already cited. However, in the absence of reliable data, it is obviously difficult to come up with effective solutions given limited resources. There are still misconceptions as well as major gaps in our knowledge and we need more accurate analysis of whatever data is available to arrive at a balanced view of cybercrime[16]. On one hand, the media has sensationalized cybercrime, including making unsubstantiated references to "cyber-war", "cyber-terror", etc. Commercial security firms have also tended to exaggerate the number of attacks and the damage done. Thus the sense of danger has been

---

[15] The cyber-security literature is also a potential source of ideas on cybercrime that could be empirically pursued, but that is not the focus of this paper. Some of the issues arising from cyber-security have been addressed in Moitra, S.D. 2003 *Analysis and Modelling of Cybercrime: Prospects and Potential*. Research in Brief/18, Max-Planck-Institute for Criminal Law, Freiburg.

[16] Wall, D.S. 1999. *Cybercrimes: New Wine, No Bottles?* In Davis, P., Frances, P. and Jupp, V. (Eds.) *Invisible Crimes*. St. Martin's Press, New York.

unduly magnified in the public consciousness. On the other hand there are indeed a significant number of network attacks and cybercrimes that are committed by various kinds of criminals, so this does call for control measures on the part of society.

Analysis for cybercrime policy is particularly relevant since policy-makers and society in general have had only a relatively short historical experience of cybercrime. Thus we have no long record of legal precedence to guide us. To the extent that cyberlaws have the control of cybercrime as part of their goal, it is self-evident that we need to study the impact of these various laws and their enforcement policies on the incidence of cybercrime. To be effective, criminal justice policies need to take into account the actual prevalence of cybercrime, the patterns of cybercrime commission by type and trends in cybercrime. Therefore we need to know them more accurately. Then we could develop more objective predictions about how the incidence of the various cybercrimes will change with changing policies. Otherwise there is the danger that laws will be based solely on popular feelings, transient sentiments fanned by exaggerated and uncritical media reports (including misinterpretation of cybercrime surveys), public overreaction, emotional responses and knee-jerk reactions on the part of legislators and law enforcement officials, since lawmakers tend to be more responsive to popular opinions rather than evidence[17]. Laws that are not properly analysed before being enacted may not achieve their intended goals of controlling cybercrime and could even be counterproductive. Policy needs to be realistic for it to be effective else they may have quite unintended consequences[18]. For example, the resources required for appropriate law-enforcement should be considered a priori. Otherwise the laws may not be implemented as had been imagined. In general, the allocation of resources to various segments of the criminal justice system (police, prosecution, courts, parole boards, etc.) requires careful study. As another example,

---

[17] This point has been made in a very different context in Dustmann, C. and Glitz, A. 2005. Immigration, Jobs and Wages: Theory, Evidence and Opinion. Centre for Economic Policy Research and Centre for Research and Analysis for Migration.

[18] Some examples are discussed in Brantingham, P.L. and Brantingham, P.J. 2004. Computer Simulation as a tool for Environmental Criminologists, *Security Journal*, 17, 1, 21-30. The well-known California law "three strikes and you're out" is another example where the consequences were not quite as anticipated, although prior research had predicted very mixed results but had been largely ignored.

laws and their enforcement should not displace less serious crimes with more serious ones. Finally, an excessive preoccupation with exotic cybercrimes that are theoretically possible, but which hardly exist, may be inefficient by failing to control actual crimes that are harming society.

The legal literature per se offers little guidance on actual implementation, yet it is the quality of implementation that is the key to the effective control of cybercrime. To improve implementation we need empirical studies and analyses of the impacts of alternative policies. Unless we have a clear assessment of their impacts, we cannot promulgate effective, efficient and equitable policies. They are additionally important because they would further our basic scientific knowledge about cybercrime and they would help us to separate perception from reality. Such analyses should result in policies that are more attuned with their own stated goals. To this end, this paper develops a modelling approach for the analysis of cyber-policy. We develop an approach to investigate the prevalence of cybercrime and cybercriminals. We also discuss a simulation approach to help answer some of the important and urgent policy questions such as how long Internet data should be stored or what impact harsher sentences for cybercrimes will have on their incidence. Although a considerable body of cyber law has already been developed, there will very likely be further developments in terms of modifications and the introduction of new cyber laws. Indeed, most policy-making bodies have suggested that much work is still left to be done in the area of cyber policy[19]. Thus it may be hoped that the approaches developed here will be useful in developing future policies. Such an analysis will not only highlight additional data that we need to collect in the future, but also promote a more informed debate on this subject among the public, stakeholders and policy-makers[20].

---

[19] All discussions in Europe and elsewhere have noted the need for new provisions to deal with new developments in cybercrime. In its document to the European Parliament titled *Creating a Safer Information Society*, the European Commission discusses a broad range of computer-related offences and calls for "a large debate at the EU level between all stakeholders on the issue of the security of information infrastructures and combating computer-related crime."

[20] *For an extended discussion of the important questions that need further empirical study, see Moitra, S.D. 2005. Developing Policies for Cybercrime: Some Empirical Issues. To appear in* European Journal of Criminology, Criminal Law and Criminal Justice.

In this paper we primarily consider the issue of estimating the prevalence of cybercrime by crime type, and the estimation of the number of active cybercriminals as a secondary issue (since almost no applicable data on offenders are available). Then we describe a simulation model to address the generation of cybercrime that could be used to explore a variety of policy questions under different scenarios. We do not consider the question of exactly what should be included as 'cybercrime' here[21] and we also do not consider the details of the actions that constitute cybercrime such as any preliminary hacking that is undertaken or the tools used to commit the cybercrime[22]. While these questions are undoubtedly of fundamental importance, they are beyond the scope of this present analysis[23]. Here we focus on the methods and difficulties of modelling the incidence and impact of cybercrime in order to illuminate and highlight some critical policy-related questions.

2. Modelling Cybercrime

A. The main elements of cybercrime models

In modelling cybercrimes, it is clearly important to disaggregate them by type, since they are often quite distinct from each other[24]. This approach

---

[21] We shall essentially rely on the literature and the crime types already considered in the surveys on cybercrime. For a discussion on taxonomies for cybercrime, see *Moitra, S. D. 2004. Internet Crime: Towards an Assessment of its Nature and Impact. International Journal of Comparative and Applied Criminal Justice, 28(2) 105-123.*

[22] This is the matter of the method of operation of the cybercriminal. They are the tools or means of committing the crime. While they can sometimes be related, (some crimes requires specific tools), our interest here is in the crimes.

[23] We also do not discuss the international dimension of cybercrime which is an important aspect. In fact, this has been emphasized by most commentators on cybercrime. The interested reader may see Brenner, S.W. and Schwerha,J.J. 2004. Introduction – Cybercrime: A Note on International Issues. *Information Systems Frontiers* 6:2, 111-114; Castells, M. 1996. *The Rise of the Network Society*. Blackwell, London; Flanagan, A. 2005. The Law and Computer Crime: Reading the Script of Reform. *International Journal of Law and Information Technology*, 13, 1, 98-117; Grabosky, P., Smith, R.G. and Demsey, G. 2001. *Electronic Theft: Unlawful Acquisition in Cyberspace*. Cambridge University Press; Westby, 2003 op. cit.

[24] Cybercrimes are also different from crimes that simply involve computers. Strictly speaking, cybercrimes involve computers both as instruments and victims of the crime *and* the Internet as the key element.

additionally allows us to be flexible about which type to include in a particular analysis. For policy research, it is also important to focus on the key cybercrimes, rather than all possible types. What might constitute the key types is clearly still debatable, but increasingly the more prevalent ones such as *virus attacks* and the more serious ones such as *remote information theft* are being identified as significant for policy analysis. Thus, the current literature and the major surveys on cybercrime can help us identify the more significant ones.

The key concepts we shall need are related to characterizing the behaviour of cybercriminals (that is, the generation of cybercrimes) and the extent of victimization. Thus we need to consider the active cybercriminal or "attacker" population {A} at any time. We also need to consider the individual crime commission rate or attack rate (λ), as well as its distribution across the offender population, and this must be analysed by crime type. The group behaviour of these offenders (who we may also call malicious hackers[25]) is also of considerable interest even though we know very little empirically. Finally their decision process regarding what kind of crime to commit and which target to choose is also very relevant, and we need to investigate it in the future.

The nature and extent of victimization is another obviously important area and we shall discuss the items of information we would like to have on this later. What we might note here is that the segmentation of the victim population is very important but may be quite complex. It may be that traditional segmentation of firms by size and sector, or the segmentation of individuals by age, education, income, etc. will not be appropriate, and that different variables will be needed to make meaningful segmentations, (such as type of computer system, level of network security, Internet activity, etc.), that would explain cybercrime experience. In any case, we shall need to have data on the rate of victimization by cybercrime type by segment and the sizes of the segments.

Our subsequent discussion assumes a "conservation of cybercrimes." That is, we assume that there is one crime for one victim and vice versa. While we believe this would be mostly true, there can be exceptions: we can have one crime with multiple victims and multiple crimes with one vic-

[25] We shall use this term even though we realize that hacking is a subset of cybercriminal behaviour. On the other hand, a very large proportion of cybercrimes in the strict sense do involve some form of hacking so the term is not out of place.

tim. For the first case, we shall need to make an adjustment by considering the distribution of the number of victims per crime. For the second case, we may consider it as a case of multiple victimizations and count each victimization as a distinct one.

B. A Static model of cybercrime prevalence

First we develop a simple, static model for the prevalence of cybercrimes under a number of simplifying assumptions and then we shall consider the possibility of relaxing some of them. The crime rate (for a given crime type) can be derived in two ways since it will be equal to the crimes generated and also equal to the crimes experienced by victims[26].

If $\lambda(a,i,t)$ is the rate at which the a^{th} criminal commits crime type i in time period t, then the crime rate C(i,t) for type i in period t will be given by

$C(i,t) = \Sigma[\lambda(a,i,t)]$ over a (a = 1,2, .. , A) where A is the number of active cybercriminals at t.

If $\nu(v,i,t)$ is the number of victimizations of type i experienced by the v^{th} victim during t, the crime rate may also be written as

$C(i,t) = \Sigma[\nu(v,i,t)]$ over v (v = 1,2, .. , V) where V is the total number of victims during t. (Alternatively, v could represent homogeneous segments and V the total number of segments. Then $\nu(v,i,t)$ would be the total number of crimes of type i experienced by the segment v in time t.)

[26] The terminology used here follows the conventional terminology in the literature on criminal careers. See Piquero, A.R., Farrington, D.P. and Blumstein, A. 2003. The Criminal Career Paradigm, in Tonry, M. (Ed.) Crime and Justice: A Review of Research, Vol. 30. Similar terminology is used in Appendix 2.

Initially, let us assume that cybercriminals are single operators (that is, they commit crimes entirely on their own), that they are specialists (a given individual commits only one type of crime), and that their rates of crime commission (λ's) are constant over time.

Then C(i,t) = A(i)* λ(i) where λ(i) is the average crime rate for type i and A(i) is the number of offenders who commit type i. Similarly, if we take average victimization rates as ν(i) and consider just one time period, we can write

$$\lambda(i)*A(i) = C(i) = \nu(i)*V(i)$$

where V(i) is the number of victims who have experienced crime type i. Thus we have two distinct ways of estimating C(i), which is the quantity of interest to us. Theoretically we can also obtain C(i) directly from police reports[27] (as we do for traditional crimes from Uniform Crime Reports in the US, or from the Annual Criminal Statistics in Germany, for example) but at this point of time they are far from complete with respect to cybercrime. In the future we could compare the estimates from these three methods and perhaps arrive at a more accurate estimate through some reconciliation process.

At present we have virtually no usable estimates for either λ(i) or A(i) and while it is clearly important to attempt to estimate them *in the future*, we shall have to assume some reasonable ranges for λ for now to estimate the A's. On the other hand, we can derive some initial estimates of ν(i) (aggregated over several business segments) based on some surveys. While there are many problems with the data so obtained, and the summaries that are available[28], we shall use some of these data to show how we *might* be able to estimate C(i), *if* we had reliable data. Therefore the estimates should

[27] These reports contain the number of crimes reported to law enforcement authorities and typically have no information on perpetrators or victims.

[28] Moitra, S.D. 2005 The Impact of Cybercrime on Business: An Assessment of Available Data (Submitted to the Journal of Criminal Justice).

not be taken as a reflection of actual cybercrime rates but rather as *hypothetical* numbers to demonstrate a methodology.

The two fundamental problems are either that the survey design is flawed and the variables of interest cannot be estimated correctly or that the key estimates required are not reported. For our present purposes, we need estimates of $\nu(v,i)$ where v could represent an individual victim or a homogeneous victim segment (such as firms in a sector) whose size is known from economic data. However, not only are these not reported anywhere, very often we are not provided with even the averages of the $\nu(i)$'s across the respondents. To estimate prevalence, we have selected two crime types, Denial of Service (DoS) and Viruses, since they are widely reported and we focus on the reports from the US and UK since we have two reports each from these countries[29].

We can summarize the information that can be used as inputs as follows:

- The surveys give some information on the distribution of cybercrimes experienced $f(\nu)$, but only through histograms with very broad ranges.
- We have percentages of respondents experiencing crime type i for a limited set $\{i\}$.
- We can infer averages for $\nu(i,t)$ for period t from some surveys for some types i;
- We *cannot* infer $\nu(v,i,t)$ by segment v since $\nu(v,i,t)$ not given for any v in any report.
- Time t is annual; we have data for two or three years (and more for some sources).
- The surveys cover organizations (mostly businesses) only.
- We still need economic/industry data on the total number of businesses by segment.

Thus it can be appreciated that it is almost impossible to estimate prevalence with any reasonable level of accuracy from the current reported data on cybercrime. However, it may be possible to arrive at approximate ranges from the reported figures since we can estimate $\nu(v,i,t)$ approximately and

[29] Thus we can compare estimates within each country and across countries. The reports are CSI, 2004, op. cit., BJS 2002, op. cit., NHTCU op. cit. and DTI 2004, op cit.

then look up the segment sizes v from economic reports. In Appendix 4 we have attempted to derive such approximate estimates of prevalence for DoS and Viruses for the US and UK. Then, making further assumptions, we have attempted to estimate the number of specialists active in committing DoS and in writing viruses. It is important to emphasize that these are for illustration only, since data are not available at this time at the level of detail that is necessary. There are many other sources of potential biases that we shall identify later in the estimation of prevalence. Thus the estimates themselves can only indicate orders of magnitude at best, and perhaps only a rough indication of even that.

For better estimates, we need to revisit the assumptions and approximations made in the method used to arrive at these figures. It should be reiterated that the key hurdle is the lack of accurate and reliable data. What we really need is the number of each cybercrime type experienced by each respondent in each year, which is $\nu(v,i,t)$, and in fact we could estimate some of them from the original data if they were publicly available. However, the other biases inherent in the surveys would still remain of course.

C. Discussion of the estimation results

The non-representativeness of the respondents along with other sampling errors most likely will result in over-estimates, since it is likely that those who have suffered more from cybercrimes will report it more readily[30]. But we have not made any corrections for this possible bias in these estimates. In other words, we have assumed the respondents in the surveys to be a representative sample from their respective countries. The results are given in Appendix 4.

The average number of crimes per organization by type has been approximated in different ways from the different reports depending on the data provided. While the methods have tended to be conservative, the values are still likely to be upwardly biased for a number of reasons, including possible *over-reporting* on the part of those who *have responded* to the surveys. The fact that we have not been able to get disaggregated victimization rates by type of organization, but only aggregates instead, may have

[30] However, the BJS study which has a more representative set of respondents has produced higher estimates of crime rates (in our example) than the CSI study.

induced some additional bias. Of course there could be a number of other biases in these approximations. That is why we have considered orders of magnitude only, as explained above. A rigorous survey design, systematic fielding of the surveys and appropriate analysis of the results should go a long way to reduce these biases.

In estimating the incidence of virus attacks, two corrections are called for: one is a correction for the *spread rate,* since the same virus can (and usually does) infect hundreds, thousands, perhaps even millions of sites. We have taken 100 and 1,000 as the upper and lower correction factors for this. Another correction factor may be needed since the same site can report multiple attacks which are due to the *same* virus infection within the organization. The NHTCU data on viruses very probably reflects this, since the mean number (255 in 2003) is totally out of line with other survey findings and seems extremely high in absolute terms as well.

As far as virus attacks are concerned we find that in the US there are roughly between 1,000 and 10,000 attacks per year, while in the UK there are roughly between 750 to 70,000 attacks per year. The results for the UK reflect two correction factors, one for the spread rate and the other for multiple reports by the same organization applied to only the NHTCU data. The DTI data suggests 1,000 to 10,000 which are within this range. This of course corresponds to a higher victimization rate in the UK when the sizes of the two economies are considered.

The estimation of the number of hackers (A) will suffer from the same inaccuracies that exist in the estimation of prevalence as they are derived from it. Moreover, the estimate of A depends crucially on the assumed value of the average individual crime rate λ. We have taken this rate to vary from 4 to 10 per year, that is, one crime per quarter per offender to about one per month. This range was based on the consideration of the time required to prepare and launch a virus or DoS attack. If and when we know more about probable values of λ, we can re-estimate A. The numbers of virus writers and DoS attackers have been estimated independently from their respective prevalences. The actual perpetrators could overlap completely, partly or not at all because the versatility (or degree of specialization) of these offenders is not known. Nor was any impact of group behaviour assumed since we have no data on it.

The number of active virus writers who have affected US organizations is seen to be roughly between 100 and 7,500 when the results from the CSI and BJS reports are consolidated. It is important to remember that they represent a worldwide population of virus writers whose viruses happen to

have infected US sites during the year[31]. Similarly, the number of active virus writers who have affected UK organizations is seen to be roughly between 75 and 17,500 when the results from the NHTCU and DTI reports are consolidated. As in the case of the numbers derived from the US data, this represents a worldwide population. If we had accurate figures, we would expect that the numbers should be comparable as they represent a global population. However, even with the imperfections in the data, there is a rough comparability between the numbers, in that the minima and maxima are of the same orders of magnitude respectively, and the US estimates falls within the range from the UK data.

Turning to DoS attacks, the US had roughly between 280,000 and 760,000 such attacks per year, while in the UK there were roughly between 33,000 to 200,000 attacks per year. This indicates much higher victimization rates than in the case of viruses. This would seem anomalous, since the average victimization rate is uniformly higher for viruses. The reason for our results is that we have applied a large correction factor for the spread rate in the case of viruses, and this may have been an over-compensation. While the approach of using such correction factors may be necessary, the values chosen here may be too large and hence the estimates are excessively deflated. On the other hand it may well be true that the number of *unique and successful* virus attacks per year is less than the number of DoS attacks annually, (which are necessarily site specific; although multiple DoS attacks can be orchestrated by the same group of malicious hackers). The number of DoS hackers seems to be roughly between 28,000 and 190,000 from the US data, and 3,000 to 50,000 from the UK data. In this case both the upper and lower limits seem rather extreme, and perhaps the number lies between 20,000 and 150,000 approximately. Again, it may be assumed that they represent an international pool of hackers who were active in committing DoS attacks during the given year.

It should be clear that these results have been obtained under very broad assumptions as mentioned already and other assumptions might be just as valid. Similarly, a very straight-forward methodology had been used for the estimations and more rigorous methodologies should be used in the future

[31] This is the estimate for the number that is active in that year. There could of course be many more *potential* virus writers in existence.

when better data are available. Here we have merely reported some approximate findings with the data that happens to be available at the present time. As further information becomes available, these estimates should be adjusted accordingly, and a more accurate picture of cybercrime will emerge.

One interesting issue is that there appears to be a paradox when we contrast the number of reported cybercrimes (which are usually very high) with the number of convictions or sentences under cyberlaws (which are extremely modest to date)[32]. To attempt to explain this partially, we can consider both over-reporting and "under-prosecution." There are actually several reasons why the reported data might be (even highly) inflated. First of all there is the matter of false positives in the process of detection of cyber attacks. It is well known that any detection system (automatic or manual) will have a fairly high false positive rate, that is, many benign signals over the Internet may be falsely interpreted as a sign of hacking or some kind of attempted crime. These may be duly recorded and reported as cybercrimes since it is only much later, if at all, that they are recognised as false positives. It is important to remember that this could be quite a high proportion[33]. Second, there could be imperfect recall, and memories of real incidents could loom larger than they really were, which would upwardly bias reports. Third, and related to this, we should remember that non-victims can only over-report. Thus errors, doubtful cases, recollections of some other network problem or even rumour could result in non-victims reporting some crimes[34]. Fourth, cybercrimes could be imputed for other mis-

---

[32] This shortfall issue is discussed in Wall, D.S. 2005 The Internet as a Conduit for Criminals, in Pattavina, A., *The Criminal Justice System and the Internet*, (77-98), Sage, Thousand Oaks, CA, 2005. Wall actually discusses two types of shortfalls. One is the type we are discussing here. The other is the "shortfall" between the numbers extrapolated from the surveys (what we have done here) and the numbers reported to the police. While this does require a fuller clarification, part of it might lie in the "over-reporting" discussed here and in the general reluctance to report to law enforcement, as noted in the survey reports.

[33] Axelsson, S. 2000. The base-rate fallacy and the difficulty of intrusion detection. *ACM Transactions on Information and Systems Security* 3, 186-205.

[34] Pepper, J.V. and Petrie, C.V. 2003**.** Measurement Problems in Criminal Research**,** National Academies Press, Washington, D.C.

takes (in managing the computer system) or as an excuse for poor performance. Fifth, multiple incidents or experiences from the *same* cyber attack could have been reported as separate incidents when they should really be counted only once. Almost certainly this is the often the case with reports of virus attacks since one infection can be experienced many times in many ways over different time periods. Sixth, there may well be duplication in collating data from multiple individuals (who may separately recall the same incident) or from multiple records of the same incident. Seventh, publicity of a few dramatic Internet crimes combined with dire warnings from security firms may have induced a "cyber-hypochondria" in some cases, and even slight aberrations in Internet transactions or unexpected system responses may be put down to cybercrimes. All of this would suggest that over-reporting could be a significant factor that is biasing the reported data.

However, there could be further factors that are contributing to this "shortfall." There could be (and usually are) many slips between the reporting of a crime and a conviction for it. If we were to trace the intermediate steps we can consider the various probabilities involved in proceeding to each successive step. There has to be an identification of the perpetrator, and the probability of this (say p[f]) is exceptionally small for cybercrime given the well-known "anonymity" of Internet users, especially of malicious hackers who cover their tracks. Next there is some probability of an investigation (say p[v]), since that is by no means automatic (because of a variety of factors, such as the feasibility of collecting evidence). Then there will be some probability of prosecution (p[s]), depending on the evidence, circumstances, etc. Finally, after the case is taken to court, there is the probability of a conviction (p[c]). Thus the probability of conviction given a reported crime is the product of all these probabilities and it is likely to be extremely small as illustrated in the Appendix 3. Combining this with the over-reporting probabilities helps to explain the observed shortfall to a certain extent at the prosecution, conviction and the sentencing levels.

D. Additional issues

There are a number of complicating factors that need to be considered when modelling cybercrime. One is the distribution of the rate of crime commission across the population of active offenders by crime type. It is quite possible that an offender commits different types of crimes, but if we can estimate the rate for each type, we can proceed with the modelling. We

can segment offenders according to their rates if they have homogeneous offending patterns. This assumes that offenders have a stable rate of offending. If they do not, then the rate must be taken as a function of time. Generally, the rate may be modelled as a function of other variables that influence it, such as policy variables (for example, sanctions). A second important variable of interest is the number of active offenders at time t, or A(t). If it is possible to segment the offender population, this will be a vector of the sizes of the segments at each time period t. All the evidence at hand suggests that the offender (or malicious hacker) population is extremely heterogeneous, with different segments having different motivations and skills[35]. In terms of motivations, we can perhaps distinguish two broad segments: those that commit cybercrimes for psychological satisfaction and those who do it for financial gain[36]. However, we know extremely little about this subject at this stage. We also know very little about whether cybercriminals are specialists (that is they commit only one type of crime) or whether they dabble in many different types of crimes (leveraging their hacking skills), or if each tends to stay within a small set of crime types (cluster specialists). In reality there is probably a mix of these types. There is also the issue of a possible correlation between versatility and crime rate: for example, more versatile criminals may have a higher rate of offending. A third issue that is very important is the group behaviour of malicious hackers. There is a suggestion in the literature that it is significant and perhaps there is a division of labour among cybercriminals, but we clearly require more data in order to model this group interaction. A particular effect of interest would be if group participation increases the effective crime rate (complementary effects) or reduces it (substitution effects). A fourth issue that is very important for the understanding of the actual patterns of cybercrime is the decision process involved in selecting a victim target or targets

---

[35] Parker, D. 1998. Fighting Computer Crime: A New Framework for Protecting Information. New York, John Wiley & Sons; Lilley, P. 2002. Hacked, attacked and abused: digital crime exposed. Kogan Page, London.

[36] There may be parallels to vandalism. See Beck, A.J. et al. 2003 Vandalism of vending machines: Factors that attract professionals and amateurs. *Journal of Criminal Justice*, 31, 85-95.

and the choice of crime type. The published literature has very little discussion of this, but informally, there are a number of models of how malicious hackers go about committing cybercrimes. However, in the absence of data, we cannot identify the major modes of operation. We also lack information on the times taken to prepare for and commit a crime, which would depend on the type of crime, the composition of the group if one is involved and the skills of the perpetrator or perpetrators.

Similarly, there are many questions concerning victimization by cybercrime to which we have no reliable answers. For one, we do not know the distribution of detection rates and reporting rates for cybercrime, nor how they vary by victim characteristic. Although some survey data addresses the issue of reporting rates, no estimates of its distribution are available as input for modelling. We must remember that the data we get comes from a "twice-filtered" process: that is, it is conditional on detection and then reporting. However, we also need to keep in mind that there can be errors on both sides. There can be *over-detection* (false positives) as well as under-detection and *over-reporting* as well as under-reporting. Some idea of these error rates is obviously important for making correct estimates. Second, we do not yet have a basis for segmenting victims that is relevant for cybercrime. It is quite possible that traditional demographics (whether for businesses or individuals) do not explain variations in cybercrime experience as noted earlier, but we do not know the relevant ones yet. Much more research and analysis is needed to understand this issue. Similarly, we do not know how to characterize the vulnerabilities of potential victims. Considerable work has been done on system vulnerabilities[37] but we have very little information on how *variation* in vulnerabilities explains *variation* in cybercrime experience. It may be a very complex issue since the probability of victimization may depend not only on the technical aspects of security (system vulnerability) but also on the digital assets of the victim and its general

[37] Doll, M.W., Rai, S. and Granado, J. 2002, *Defending the Digital Frontier: A Security Agenda*, John Wiley and Sons, New York; Garfinkel, S. and Spafford, G. 1997 *Web Security and Commerce*, O'Reilly & Associates, Sebastopol, CA; Ghosh, A.K. 2001. *Security and Privacy for E-Business*, John Wiley and Sons, NY; Krusl, I. V. 1998. Software Vulnerability Analysis, Ph.D. Thesis, Purdue University; Stallings, W. 2002 *Network Security Essentials* Prentice Hall.

attractiveness as a target (which would be different to different criminal segments). Third, we have very little information on which to estimate the distribution of cybercrimes experienced by type and by segment. Unless we know this, we cannot estimate the general prevalence, as has been pointed out above. However, these distributions may be quite complex and very different by segment, and moreover, there can be correlations between the rate of victimization, type of victimization[38] and the probabilities of detection and reporting. These will of course introduce significant biases in the reported data unless corrected. Moreover, the detection probabilities may well be increasing since improved security technology is being developed and deployed by organizations (and individuals) and Internet usage patterns may be changing which would change the visibility of organizations in cyberscape[39]. All these factors contribute to making the problem of modelling cybercrime (its generation or resulting victimization or both) extremely complex and analytically intractable. It is for these reasons that we propose a simulation model for cybercrime.

E. A Simulation Model

A simulation model would be most appropriate for such analyses since simulation is particularly suitable for analysing such a complex stochastic system. Realistically, it is very likely that there will be gaps in our knowledge of cybercrime. The simulation model can help in overcoming the limitations due to lack of data, since we can use ranges of values for parameters over which there are uncertainties and perform sensitivity analysis over the ranges. Simulation can provide insights into the cybercrime process[40]. Such

---

[38] *See Moitra, S.D. and S. L. Konda 2004. An empirical investigation of network attacks on computer systems.* Computers & Security*, 23, 43-51. for a more detailed discussion.*

[39] Therefore apparently increasing trends in reported cybercrimes may not signify a real increase, *even if the reporting base is the same*.

[40] Cohen, F. 1997a. Simulating Cyber Attacks, Defences, and Consequences. Fred Cohen & Associates; Moitra, S. D. and Konda, S. L. 2000. A Simulation Model for Managing Survivability of Networked Information Systems. Technical report CMU/SEI-2000-TR-020. Software Engineering Institute, Pittsburgh PA. Available at ww.cert.org/research/00tr020.pdf; Narasimha, K.B. and Moitra, S.D. 2002. A

a model would allow us to better comprehend the different aspects of cybercrime by allowing us to explore various scenarios and observe the effects of different policies on cybercrime under different conditions. Thus we can consider such questions as whether a given policy will really reduce crime or merely shift the crime rate of one type to another. As another example, it would be possible to model the multi-level control scheme that has been suggested[41] and investigate its scope and effectiveness. Similarly, other issues raised in the literature can be examined analytically and perhaps resolved with the model.

The inputs to the model can be derived in part from the analysis of cybercrime data suggested above and the simulated scenarios could incorporate whatever is known about the characteristics of the malicious hackers and the security measures in computer systems on the Internet. The impacts on the victims of cybercrime can also be studied with the model. The structuring of the model will make explicit the relationships among the different variables affecting cybercrime and we shall be able to see which variables are influential (and which are not) through sensitivity analysis. This will help identify critical data needs for future studies. In addition, the model can help us to judge the reasonableness of the assumptions made about the nature of cybercrime as reflected in the current literature by comparing the results we obtain from simulation with what we can judge from our empirical knowledge of cybercrime.

Simulation models have been used in criminal justice studies many times previously. The contexts have ranged from viewing the criminal justice system as a whole (JUSSIM), to various aspects of the system such as the flow of court cases, the processing of juvenile cases and the impact of enforcement policies on the availability of drugs[42]. The model proposed here

---

Simulation Model for Measuring the Effectiveness of Networked Information Systems. Indian Institute of Management Calcutta Working Paper.

[41] Grabosky, P. 2001 op. cit.; Wall, D.S. (Ed.) 2001 *Crime and the Internet*. Routledge, London.

[42] Brantingham, P.L. and P.J. Brantingham. 2004. op. cit.; Caulkins, J.P., Crawford, G. and Reuter, P. 1993. Simulation of Adaptive Response: A Model of Drug Interdiction. Mathematical and Computer Modelling, 17 (2), 37-52; Stewart, A., Spencer, N., O'Conner, I., Palk, G., Livingston, M. and Allard, T. 2004, Juvenile Justice Simulation Model: A Report to the Australian Research Council Strategic Partnerships with Industry. An earlier model, JUSSIM, was developed by A.Blumstein and colleagues at Carnegie-Mellon University. See Belkin, N.,

will simulate crime rates by crime type given the best estimates we can obtain and explore various impacts of cybercrime such as patterns of victimization. Initially, one should start with the important types, such as denial-of-service, data theft and viruses, and consider more crime types and interactions as the model is developed. The basic idea is to model the dependencies and relationships between crime commission (by type), criminal justice/law enforcement policies and the Internet environment that would include levels of security, user responses, informal controls and so on. The model needs to take into account the fundamental drivers of any crime, namely motivations, opportunities (including absence of guardians) and skills of the criminal in the context of cybercrime[43].

The model will be able to include the diverse impacts on victims: both the systems and users. Among users, individuals and organizations must be considered separately, since they obviously will have very different characteristics. The impacts will have to be modelled as a function of crime (or attack) type, hacker skills and defences incorporated into the systems. The results of the simulation will reveal the victimization patterns and we shall be able to assess the effects and implications of different assumptions. An advantage of simulation is that the effect of variables and structural relationships can be assessed relatively easily and thus we can simulate a large number of scenarios to reflect varying assumptions regarding the possible values of parameters when we are not certain of them. Since the scenarios would include controllable variables as explained below, the impact of a given change in criminal justice policy (even a radically new policy) or a change in Internet conditions such as security measures deployed or user

---

Blumstein, A., Glass, W. and Lettre, M. 1972. JUSSIM: An interactive computer program and its uses in criminal justice planning. Proceedings of Project SEARCH Symposium.

[43] The "routine activity approach", Cohen, L.E. and Felson, M. 1979. 'Social Change and Crime Rates and Trends: A Routine Activity Approach', *American Sociological Review*, 44: 588-608, has been invoked in the context of cybercrime (Adamski 1998 Crimes Related to the Computer network. Threats and Opportunities: A Criminological perspective. Nicholas Copernicus University, Poland; Grabosky 2001 op. cit.). However, in the case of cybercrime it would be more appropriate to investigate the triad of *motivation*, *opportunity* and *skills* with "absence of guardians" contributing to the opportunity available. This is because skills play a very important role in cybercrimes.

behaviour can be assessed with the model. As a consequence the model can be used for forecasting cybercrime as well since it will predict the prevalence of cybercrime under future, postulated conditions. Some work has been done along these lines[44] but it is suggested here that those models need to be extended to include more details of the cybercrime generation process, such as a sub-model for the rate of crime commission by crime type. It is also important to enrich the relationships between crime generation and policy-controlled variables based on current data and theory. Then it will be possible to see more clearly the implications of alternative policies in terms of their impact on the prevalence and patterns of cybercrime. For this, the relevant parameters of the model must be considered as functions of the policy variables. For example, the number of active hackers (A), their individual crime rates (λ), their group behaviour, their choice of crime type and target can all be modelled as functions of the policy variables of interest (for example, sentencing levels, intensity of prosecution, degree of informal controls, etc.). Thus we can track how policy changes influence cybercriminal behaviour and how this change in behaviour in turn results in changes in the incidence of cybercrime.

This methodology permits us to test for various biases that we may suspect as existing in the available data. For example, we might hypothesize a possible bias. We can then adjust the model to counteract the bias and then run it again. The new results will reflect the situation we would have observed if there was no bias. It is quite possible that this could be a more accurate view of reality. As a simple illustration, we might postulate a certain bias due to non-reporting. By simulating the model with adjusted reporting rates we can arrive at new estimates of cybercrime prevalence that could be more realistic. Since we may not know of the degree of actual bias we can try out a range of values for it and judge the results. Since the model will reveal the effects of changing the values of any of the parameters on related variables (based on the assumed relationships) it could be possible to resolve potential inconsistencies in findings related to cybercrime. Further, the model can always be extended to other (possibly newer) crime types, different hacker behaviour (based on new tools) and so on. In any case, the model can always be updated as necessary. The simulation model will complement the knowledge we derive from available data and help us to

[44] See footnote 40.

construct a more accurate picture of crime in cyberspace. This in turn will suggest how cybercrime may be expected to evolve under different law enforcement policies and security measures. Such a model will allow us to forecast trends and move from a reactive approach to cybercrime to a proactive approach. The model would support governments, criminal justice/law enforcement officials, Internet policy makers, and major organisational users of the Internet in making more appropriate decisions regarding all aspects of risk management in cyberspace.

The simulation we are proposing here is important because it will be possible to model the stochastic and dynamic aspects of cybercrime as well as the interdependencies among the variables in the model, such as between attack rates and hacker group size or between versatility and individual crime rates. It is generally difficult or impossible to obtain closed-form, analytic solutions for stochastic systems (especially with complex distributions and correlated variables), so it is often preferable to conduct a simulation rather than find approximations. An outline of the simulation procedure is given in Appendix 1.

One policy goal could be general deterrence of potential cyber attackers through sentencing or fines. However, there are some problems regarding its effectiveness since the probability of a conviction for a cybercrime is very low in the first place, as explained in the Appendix 3. Additionally, since the criminal justice system may regard cybercrime as a white-collar crime, prison sentences may be relatively rare and the deterrence effect of fines may not be that strong. The special deterrence effects of imprisonment or fines are not known. Even if they are quite strong, it would apply to relatively few people (see Appendix 3).

There can be two kinds of incapacitation with respect to cybercrime policy: one is the physical imprisonment of convicted criminals[45] and the other is "banishment" from cyberspace. The latter incapacitation involves preventing a person from using the Internet[46]. In the first case, its effectiveness

[45] Spelman, W. 1994. *Criminal Incapacitation*. Plenum Press, New York, has an extensive discussion on incapacitation for traditional crimes including modelling approaches to estimate its impact.

[46] In extreme cases, hackers have been prevented from using even a self-standing computer; this has arisen from a confusion between computers and the Internet on the part of some law enforcement officials in the past.

is limited because incapacitation can be imposed on very few people, since very few cybercriminals are convicted, and in the second case, the restriction cannot be fully implemented in practice as argued above, since these days anyone can usually find a way to access the Internet. In other words, it is impractical to implement an incapacitation policy by simply notifying local ISPs and some systems operators that a certain person should not get Internet access. However the actual effectiveness has not been analysed or measured. On the other hand it has been considered as a policy option as has been discussed.

There are of course a wide range of preventive measures that can be taken and are being increasingly taken. There is now a computer network security industry that has been spawned by the fear of cybercrimes[47]. This is a user-specific approach – each user puts up some defensive measures. There are also developments in network technology and digital security generally, that can be implemented on a network basis to control some types of crime. Thus, filters to screen out spam or undesirable materials, encryption and other safeguards can be seen regularly on Internet-based services. One can extrapolate these trends and envisage a future technological state where much of cybercrime has been eliminated through software and hardware design. However, it is difficult to forecast when it would happen, or if at all.

Finally we can consider yet another traditional policy goal in criminal justice, and that is rehabilitation. However, it would have to work somewhat differently in the case of cybercriminals (if it did work). There is the notion in the literature that a segment of hackers commit cybercrimes in part because they do not appreciate the harm they might cause (or even that their activities are illegal). Educating such offenders (as well as skilled Internet users at large) about the negative aspects of their actions might be cost-effective. This can be done at any stage, for example after identification or arrest. It has even been suggested that relevant computer science courses have an ethics component. However, in the absence of appropriate empirical studies, it is impossible to say much about the overall effectiveness of education. On the other hand, there might be another rehabilitative path that will be effective, and that is a policy of cooptation. This is the

[47] Also, there are now a very large number of books on network security: interested readers may see among many others the one cited in footnote 37.

process of converting malicious hackers into using their computer skills for legal activities, including becoming computer security managers[48]. However much it may violate our finer sensibilities, and however much firms may deny that they hire converted hackers, it may ultimately be the most practical solution. Given sufficient incentives, it is probable that a large proportion of malicious hackers will indeed turn to legal activities. There is no reason to assume that all or most malicious hackers are irrational. This is also something that can be modelled (under various scenarios) and explored through simulation.

To sum up, there are many urgent policy issues that should be analysed before any definitive policy is formulated, and although we have mentioned some previously, we shall highlight a few of them here. One is the matter of data retention by the ISPs. Law enforcement officials naturally want ISPs to retain data for a long time – for weeks or even months, since it could facilitate cyber forensics. However, the data in question are the records of all Internet communication, and the total amount (even per day) is huge. Therefore, it quickly becomes very expensive to store and maintain all this data for long periods of time. The optimal balance depends on the benefits of having the data to extract evidence of cybercrimes and the costs of retention. The benefits presumably would come from higher probabilities of successful arrest and prosecution of the perpetrators. Thus in turn would presumably translate to lower levels of cybercrime through deterrence and possibly incapacitation. Whatever the right balance is, data cannot be retained for very long in practice (let alone for ever) and thus at some point it will "disappear."[49]

There are two additional factors that would encourage the retention of data for only a limited time. One is the matter of privacy: privacy advocates

---

[48] This has obvious parallels to the labelling of hackers as white-hat and black-hat hackers. White-hat hackers include those who are inherently law-abiding and those who have been converted to performing strictly legal activity. Black-hat hackers are those who have strayed to the "dark side." However, it is not clear that this distinction is always maintained or is a reflection of reality in the first place. It might be more realistic to consider hats of various shades of grey. It is conceivable that a well-designed package of incentives can result in more and more hats of lighter shades.

[49] Wall, 2005 op. cit. talks about the possibility of "disappearance of disappearance," that is, the prospect that records will never disappear, but for a number of reasons that may not really occur in practice.

(and officially both the European Union and the Council of Europe) have urged the deletion of personal data as soon as possible to reduce the possibility of individuals' privacies being invaded and lawmakers have paid heed to this argument. The second factor is the matter of processing all that data to extract useful information. Even with modern data processing technology (data mining, etc.) our ability to find useful information from really huge amounts of data is extremely limited[50]. Thus it is likely that it will soon be seen that keeping all that Internet traffic data is pointless.

A second policy issue is the assessment of the efficacy of various formal and informal controls that are already in operation on the Interne. For example, it has been argued that official law enforcement agencies may not have to intervene very much since order on the Internet is already being maintained by a number of private or semi-private organizations in concert with limited efforts on the part of official agencies[51]. Similarly, it has been suggested that a multi-tiered system of controls monitored by ISPs, user groups, watchdog organizations, corporate users and other non-governmental agencies have successfully kept cybercrime in check[52]. While these optimistic viewpoints are certainly valid to some extent, and scarce public resources should not be spent unnecessarily on redundant control measures, we do not quite know the true efficacy of this non-official or quasi-official system of controls[53]. We can incorporate these control mechanisms in our simulation models and see, under a reasonable range of assumptions, their effectiveness and limits.

The allocation of resources to the various parts of the criminal justice system is clearly a key issue in policy making and implementation. Ideally we would like to allocate available resources for the greatest effectiveness and one measure is clearly reduction in crime. For such optimal allocation, we first need to understand how prevalence of crime changes with changing policies[54]. For example, if the probabilities of investigation or prosecu-

---

[50] There is a natural limitation because totally innocuous data and incriminating data can often look alike, and there will inevitably be considerable errors in identification.

[51] Grabosky, Smith and Dempsey, 2001 op. cit.

[52] Wall, 2001 op. cit.

[53] It is also important that cyber-vigilante-ism is not unduly encouraged by excessive enthusiasm for privately operated controls on the Internet.

[54] The model can be extended to any other measure of effectiveness if the relationship between that measure and policy variables is specified.

tion for cybercrime increases as a result of allocating more resources to these activities, we can estimate the extent to which prevalence will decrease (if at all) by taking into account the impact of investigation or prosecution on the incidence of cybercrime. This kind of analysis should be done by crime type, and such issues can be explored through the simulation model. In this context we might consider a "priority index" such as a weighted crime index that takes both prevalence and seriousness into consideration[55]. Unfortunately, not only do we not have working estimates of prevalence by crime type, we do not have any systematic model for measuring a seriousness index for cybercrime[56].

Laws and law enforcement can often have unforeseen consequences. A set of such possibilities is the different kinds of displacements that could occur in criminal activity[57]. For cybercrime, three kinds of displacements are especially relevant:

- Crime displacement – criminals shifting from one type to another;
- Tactical displacement – criminal turning to different methods of operation;
- Target displacement – criminals targeting different segments of victims.

As we gather more information on cybercrime patterns, we can incorporate the displacement effects into our model and analyse the final effects of various proposed laws and enforcement policies. We can also analyse the displacement effects of preventive security measures and informal control mechanisms on the Internet itself.

---

[55] If only one of these factors, say seriousness, is considered for policies, we might end up ignoring minor offences (such as spam) entirely and devote all our resources to trying to prevent theoretically serious but hypothetical crimes which in practice never or almost never occur. For a discussion of measures of seriousness of offences see Stylianou, S. 2003. Measuring crime seriousness perceptions: What have we learned and what else do we want to know. *Journal of Criminal Justice*, 31, 37-56.

[56] One exception is a study reported in Furnell S. 2002. *Cybercrime: Vandalizing the Information Society*. Addison-Wesley, New York.

[57] See Smith, R.G., Wolanin, N. and Worthington, G. e-crime solutions and crime displacement, Australian Institute of Criminology, T&I no. 243, 2003 for a more detailed discussion on crime displacement.

F. Data requirements for simulating cybercrime

From the discussion of the data and models so far we can summarize the needed data. If we had the listed below, we would have for all practical purposes an "ideal dataset" for the analysis and understanding of cybercrime. We shall discuss the required data in 4 parts:

i] what is available and known; ii] what could be estimated from available data; iii] what requires new surveys and collection methodologies; and iv] what probably will not be available in the near future.

What is available and known:

Essentially all that we know about cybercrime comes from unrepresentative samples. An exception is the BJS survey, but it was a limited, pilot study. Most surveys ask about the general experience of cybercrime, organizational procedures regarding cyber security and the responses of organizations to cybercrime. The respondents are also asked about their losses due to cybercrime but no reliable methodology has been developed yet to help respondents estimate them accurately, and as a result, the reported financial losses may well be exaggerated in many cases.

What could be estimated from available data:

The data collected from the surveys could easily be analysed in greater detail and much more information about cybercrime patterns could be obtained. For example, we could get the distribution of the number of cybercrimes experienced by type of crime from most surveys. We could also get the mean victimization rate for each crime type by industry sector. This would be very useful to make projections about the general prevalence about cybercrime (although the bias due to non-representativeness will remain). Similarly, we could estimate the victimization rate by type of computer security network defence the organizations have. In general, many correlations and contingency table analyses could be performed with the survey data.

What requires new surveys and collection methodologies:

Ideally, we should field new surveys with proper sampling schemes so that we get a representative sample of respondents from the population

we wish to study. In the past, surveys have focussed on organizations, particularly businesses, but we need to sample individuals and households as well. We should also try to locate and survey cybercrimnals to the extent possible, in much the same way that prison populations have been studied, and collect data on self-reported criminal activity. We need to consider a comprehensive set of true cybercrimes. In the past, significant types have been omitted in surveys and crimes that are not really cybercrimes have been included. In some cases, the crime included were hardly even computer crimes, but really were property crimes, such as theft of computer-related objects. The surveys should ask about the precise damages caused, and have a step by step procedure to estimate financial losses more accurately. The times of the attacks by type would be extremely useful for Internet risk analysis. Among the many other improvement possible, one could consider ways to improve the response rate, investigate the low reporting rates to law enforcement, validate the collected data (as the BJS study has done) and to follow up on non-respondents.

What probably will not be available in the near future:

A wish list of the data on cybercrime for analysis would be quite long. Restricting ourselves to estimating the prevalence of cybercrime and what is most relevant for policy analysis, it would have been useful to know about the rates of detection of cybercrime, that is the probabilities of cybercrimes being detected by different victim segments, and also the false positive rates. Additionally if we had a standard system of collecting and reporting known crimes as we have for traditional crimes, we would certainly have a better idea of prevalence. However, at this point it seems almost impossible to set up such a repository for a variety of reasons and it is unlikely to be set up in the future.

As far as the behaviour of malicious hackers is concerned, we know virtually nothing and it is unlikely that we shall ever get sufficient information to make any confident estimates about their behaviour and in particular how that behaviour might be affected by changing policies. Among the data we would like to know ideally are the number of active offenders over time, their individual offending rates by crime type, the distribution of these rates, the nature of group activity among the offenders, their crime type switching patterns, and their methods of crime type and target selection. In the absence of statistically reliable data,

there are many notions about hacker behaviour in the literature and popular press. However, to truly understand cybercriminal behaviour, we need to distinguish between conjecture about that behaviour and valid inductions from it.

3. Summary and Future Work

A. Summary

The enormous growth of the Internet has concurrently spawned new opportunities for crimes and this has resulted in various kinds of social responses. Much has been written on cybercrime and most jurisdictions now have cyberlaws and cyber policies. However, to further develop cyber policies, we need a better understanding of cybercrime itself. To this end, we have considered two empirical issues in this paper: a) the estimation of prevalence of cybercrime and b) the modelling of cybercrime generation as well as victimization by it. We discuss the current problems in the estimation of cybercrime and present some approximate ranges for the prevalence of viruses and DoS and the number of active offenders involved based on available data. In view of the complexity of cybercrime, we propose a simulation model for a more detailed study of cybercrime. We note some advantages of this approach and the data required to run the simulation model.

B. Future Research Needs

The above discussion has highlighted a number of research needs. We reiterate here the key areas that are important for future research.

- More data needs to be systematically collected on offenders, cybercrime victimization and overall prevalence by type. The major problem is in verifying the accuracy of the survey results and how to reduce the biases that will inevitably exist.

- More refined analysis is required with more detailed secondary data on victim segments.

- The analysis should consider cybercrime disaggregated by type to be meaningful.

- It would be very useful to have data on investigations, prosecutions, convictions and sentences arising out of the new cyberlaws.
- To understand the impact of cybercrime we need to extend the proposed model and run simulations with it. In particular, we need to test various hypotheses about hacker behaviours since we know very little about it.
- The simulation model should enable us to estimate the effectiveness of the traditional goals of policies with respect to the control of crime: deterrence, incapacitation, and rehabilitation.
- In addition, other and newer policy options should be analysed.

As we have emphasised several times in the preceding discussion, it is extremely important to first study policy implications from simulation and other analyses before making recommendations regarding new cybercrime policy. Even though this will require patience on the part of all interested stakeholders, it would be worth the wait if more effective, efficient and equitable policies evolve.

Appendices:

1. Procedure for the proposed simulation

In order to simulate the cybercrime process from the generation of the crime by individual malicious hackers to tracking the impact on victim systems, we propose a set of modules that would simulate different parts of the whole process. The three basic modules are I) the generation of cybercrimes, II) target selection and III) victim characterization. A fourth module is described for the summarizing and analysis of the data collected in the course of the simulation. The mathematical formulations are provided in Appendix 2.

I) Generation of cybercrimes

There can be a number of approaches to accomplish this. We may a) select an attacker with a set of characteristics from suitable probability distributions, then simulate his or her activities given an individual crime rate and propensities to commit various crime types; then we repeat the simulation for the other hackers; b) derive a process that represents the generation of crimes by the whole population of hackers but interactions might be complex to model by this method; c) consider the crime types as independent and generate a sequence of crimes over time for each type given the individual (or segment) rates for committing the different crime types.

If we have enough information to segment malicious hackers into homogeneous groups by their characteristics, such as individual crime rates and the type of crimes they commit, we can simulate the activities of each segment independently. For example, hackers may be classified as generalists (potentially committing all possible crimes), strict specialists (committing only one type of crime) or cluster specialists (committing a small set of related crimes). We might also segment them by the activity rate (high rate/low rate or a finer discretization) or consider a continuous distribution of the rates within each segment.

Group activity is another phenomenon that can be incorporated into the model through this module. We can simulate the propensity to coordinate attacks in groups, for example, and we can also vary the attack

rates of the groups as a function of their group size or composition. If we wish, we can include the times to prepare for and conduct an attack.

Finally, and most importantly, we can make both A (size of the attacker population) and λ (the individual crime rate) functions of policy variables. This will enable us to observe how the patterns of cybercrime change with policy, and this is the crucial part of the policy analysis. In reality A and λ will be disaggregated as appropriate.

II) Target selection

This aspect of offender behaviour is also quite complex. Actually we know very little how malicious hackers go about committing their attacks. Do they select a target first and then decide on the crime type conditional on the target? Or do they decide on a crime type to commit first, and then find a matching target conditional on the crime selected? Or are both the selections completely random? If they do not succeed with one target, how persistent will they be with other potential targets? Depending how we wish to model the attack behaviour, we can simulate the target selection. This module will consider those aspects of potential victim sites about which the attackers have information and which influence their choice probabilities. In other words we should include the concept of "proneness" to cybercrime, since there may be some characteristics of potential victims that make them more likely to be victims, perhaps victims of particular types of crimes.

Then we shall have to select (probabilistically) a victim from a target segment or we can select any victim from the whole set of potential victims. In either case we can use the uniform distribution to select the victim of the attack we have generated in I). The target selection module can also include policy variables, so that later the sensitivity of victimization patterns to policy variations can be studied.

III) Victim characterization

Next we need to establish a set of characteristics of this victim beyond its segment membership (if segmentation is used). For this we can use information on the distribution of the various characteristics of interest, such as level of security the victim has, its information assets, etc. With these characteristics, and knowing the type of crime being committed, we can model the impact (that is, damage) done to it. If we have mod-

elled attacker skills, we can use that information as well in simulating the impact. Mostly, these variable will be ones which do not influence target selection either because the malicious hackers do not have this information or are indifferent to it.

IV) Summarizing and analysis of the data collected.

Finally, the data collected during the simulation has to be collected and the required estimates along with their standard errors (or confidence intervals) need to be computed. There are two key measures that are essential: i] the distribution of the number of victimizations by victim type and crime type or $\nu(v,i)$. From this we can compute the average victimization rate by crime type over all segments, average victimization rate by segment type over all crimes and the grand mean of victimization rates; and ii] the total number of crimes committed by type or the prevalence of the different cybercrimes.

These results can then be replicated under any other set of assumptions and the sensitivity of the prevalence and victimization rates to any other parameter of the model can be calculated by varying that parameter. Of course, patterns in the generation of cybercrime can also be analysed according to what we might be interested in. For example, we can track patterns of repeat victimization and study victimization histories analogous to criminal careers[58]. The simulation exercise can yield victimization patterns for a site in a given victim class for any class[59]. If we can also collect data on victimization on sites knowing the class each belongs to, we can validate the simulation model. In general, this simulation approach allows us to study the cybercrime process from many different aspects.

---

[58] Farrell, G. et al. 2000. Career Victims and Victim Careers, *Crime Prevention Studies* Vol 12.

[59] We can further study repeat victimization both from data on individual sites and also from the simulation results which of course will depend on the model assumptions. For more details, see Farrell, G. and Pease, K. 2003. Measuring and Interpreting Repeat Victimization, in Smith, M.J. and Cornish, D.B. (Eds.) *Theory for Practice in Situational Crime Prevention*, Criminal Justice Press, Monsey, NY.

2. Mathematical details and model derivations

A) Characterization of attackers and the generation of cybercrimes

I) Segmentation: The *active* offender population can be taken as A = {A(s)} a set of offender segments and the size of segment s is given by A(s). There can be alternative bases for segmenting this population. i] It could be on the basis of the types of offences they commit: for example, we could have specialists (committing only one type of crime), generalists (committing any crime) or cluster specialists (who would be somewhere in between). The problem may be that we could end up with many sub-segments of them. We have no theory on which to cluster offenders but it may be the most significant dimension for clustering. The current groupings of cybercrimes, such as "content-related" offences, may not correspond to unique offender segments. ii] It could be on the basis of skills and/or group activity since it is likely that these are related and both are likely to influence the crime commission rate. iii] It could be on the basis of crime propensity if there are sharply differing segments in this respect. If not, we can take a probability distribution $f(\lambda)$ over the offender population or within each segment otherwise defined.

II) Crime commission rates: The individual crime rate λ may have a distribution $f(\lambda; i, \underline{Z})$ over the offender population where $\underline{Z}$ is the vector of policy variables. If there are discrete segments in terms of the individual rates, we can take $\lambda(s, i, \underline{Z})$ for each segment s. In the case where the cybercriminals are versatile (generalists or cluster specialists) we can include p(i,j), where p(i,j) is the probability of switching from crime type i for one offence to type j for the next offence. There may be a correlation $\rho(\lambda,\underline{p(i,j)})$ between an individual's crime rate and the person's versatility, represented by $\underline{p(i,j)}$,and we have to consider that as well. $\lambda(i)$ may depend on the offender's skill level as well.

III) Group behaviour: It is important to model group behaviour as well. We need to consider the distribution of group sizes f(g), and it may not be a smooth function of size g. For instance, a pairing of two offenders may be very common, and the distribution will have a spike at that point. The key issue is how group behaviour affects the crime rate of the members: does the effective rate increase as a result of synergistic effects or does it perhaps decrease, as would be the case if they committed crimes at the same rate together as they would have separately. Thus of the λ's of two offend-

ers is 8 per year, acting individually they will commit 16 crimes a year. However, if working together they commit 8 crimes between them then only 8 crime will be committed and their effective rate will be 4 per year. Thus λ can increase or decrease, but we have no data to estimate this effect. The distribution of skills may also influence the crime rates of a group.

IV) Selection of crime type and target: There are several ways to model this, and we describe one method here. We can assume that p(i), the probability of selecting crime type i is dependent on offender segment s to which the offender belongs. Thus we have to consider a set of values [p(i) | s]. We need to get some data or make some judgements regarding these values. The given a crime type i, we can simulate a victim class l, one of whose members will be the target of this attack. We currently have some data on which to estimate Pr{l|i} but they are probably not reliable. However, future surveys should be able to estimate these. What this assumes is that potential victims {V} have been classified into segments or classes $\{V_1, V_2, \text{ to } V_L\}$ which correspond to their cybercrime experiences. Again this should be possible with good survey data. Then we can further simulate the selection of a particular site v in class l with probability $q(v) = 1/V_l$. Given this site, we can simulate particular characteristics for it given the distribution of these characteristics within l, since these may be independent of the characteristics that partition {V} into the L classes. There is the final issue of selecting multiple targets, if that is relevant for this particular crime. We can simulate this number (n) from a distribution f(n), which we have to estimate from data that should be available from the CERTs.

B) Characterization of victimization patterns

We also need to model q(v), the probability that victim v is chosen(if it has not been done above), and which can be represented as q(v; X(v), i, D, $\underline{Z}$), that is, as function of defence level (D) and the characteristics of the victim, X(v), and the other factors $\underline{Z}$ that may be relevant. Then we can consider the sequence of attacks $\underline{\Lambda}$ generated by the attackers in A (which is the summation of λ(a,i) in theory). $\underline{\Lambda}$ is a stochastic point process, and can be modelled relatively easily. This will lead to ν(v,i) which is the number of attacks experienced by victim v of type i. From this simulated data, we can estimate the victimization patterns for every potential victim v. We can further estimate the damage caused if we can develop a probability distribu-

tion of the "end states" (say ψ to ψ') conditional on i (as well as X(v), D, $\underline{Z}$) and even the skills of a (if that is possible). The assessment of this new and possibly degraded state ψ' will yield an estimate of the damage. By simulating a large number of attacks, we can compute the expected damage. This damage could include a variety of dimensions such as financial loss, data compromises and impairment of functionality, that is, factors similar to the issues the surveys have tried to address.

C) Change in aggregate prevalence

For simplicity we shall consider a homogeneous crime, individual crime rate and active criminal population for the mathematical representations by dropping the index "i." Then the crime rate is

$$C = \lambda * A$$

And a change in the crime rate as a function of the changes in A and λ can be represented as

$$\Delta C = \lambda * \Delta A + A * \Delta \lambda$$

This expression will allow us to calculate the impact of a change in policy (say, sentencing) which would result in changes in both λ and A in general.

3. Estimation of the probabilities of conviction and sentencing given a reported crime.

Step in the criminal justice process	Notation	Range of likely values[60]
Identification | Report	p(f)	.01 to .1
Investigation | Identification	p(v)	.1 to .2
Prosecution | Investigation	p(s)	.5
Conviction | Prosecution	p(c)	.6
Conviction | Report	π (c)	.003 to .06
Sentence | Conviction	p(s)	.7 to .9
Sentence | Report	π	.0021 to .054

$\pi(c) = p(c) * p(s) * p(v) * p(f)$

Therefore number convicted $n(c) = \pi(c) * C$, where C is the aggregate prevalence of reported crime. From this rather simple formulation (ignoring crime types, etc) we can see that convictions would be about two orders of magnitude less (about a hundredth) than C, and probably about three orders of magnitude less (about a thousandth) if we corrected for over-reporting. Similarly, $\pi = p(s) * \pi(c)$, and hence the number of people sentenced would also be three or even four orders of magnitude less than the number of reported crimes. In fact, some of the available indications suggest that $\pi(c)$ and π are in fact considerably lower. This would imply that the values assumed in the above table are actually over-estimates.

[60] These are approximate values based on various values proposed in the literature, industry newsletters and press reports. There is actually no reliable data on these. The values should be revised in light of future data.

The expected sentence $E = \pi * S$ where S is the mean sentence length when any sentence is imposed at all upon conviction. [any deterrence literature?] Therefore if we were to estimate the impact of a change in sentencing policy, for example, an increase in the average sentence meted out to convicted cybercriminals, we need to consider the corresponding change in prevalence. If ΔC is the change in the prevalence as a result of a change of ΔS in the average sentence S, then $\Delta C/\Delta S$ is the relative impact on the crime rate and is equal to $\pi * \Delta C/\Delta E$. Now if $\Delta C/\Delta E$ represented the actual deterrence effect on cybercriminals, we can see that the reduction in the crime rate would be quite small even for a significant increase in sentences (as a result of a policy change) and even if there were a deterrence effect, because of the small value of π. Unless all the above probabilities are increased (through a more intense criminal justice intervention policy) so that E is closer to S, we should not expect deterrence to play a major role in the control of cybercrime.

In assessing the cost-effectiveness of various policies we need to consider $\Delta C(i)$ versus the cost of $\Delta(p.)$ (where p. stands for any of the above probabilities), or $\Delta C(i)$ versus the cost of $\Delta(\pi)$. That is, we are examining how resources spent in increasing any of the above probabilities would reduce crime. In particular the *utility* to society of $\Delta C(i)$ is the important quantity and we may use a weighted crime index or some general function to estimate that utility.

4. Estimation of prevalence of virus and DoS attacks and attackers

(Computed numbers may not be exact because of rounding; oom = "order of magnitude")

CSI/FBI data:

Variable	Method of approximation	Approximate range of values
Prevalence of virus attacks (in the US)	Average = average of all attacks * fraction that are viruses = 5 * .32	1.6 to 2.0 per year (Correction for more than one experience per firm)
	Prevalence = average * victim population (= average * .53 * 1.59 mill)	1.27 to 1.59 million per year
	Correction for spread rate 100 to 1,000	1,000 to 10,000 (oom)
Prevalence of virus writers	$A = C/\lambda$	λ = 4 to 10 per year
	Active per year	100 to 2,500
Prevalence of DoS attacks	Average = average of all attacks * fraction that are viruses = 5* .07	.35
	Prevalence = average * victim population (= average * .53 * 1.59 mill)	About 280,000 per year
Prevalence of DoS hackers	$A = C/\lambda$	λ = 4 to 10 per year
	Active per year	28,000 to 70,000

BJS/Census data:

Variable	Method of approximation	Approximate range of values
Prevalence of virus attacks	Average from distribution = 3	3 to 8
	Prevalence = average * victim population = 3 * 64.1% of 1.59 mill.	About 3 million per year (with lower estimate of 3)
	Correction for spread rate 100 to 1,000	3,000 to 30,000 or: 1,000 to 10,000 (oom)
Prevalence of virus writers	$A = C/\lambda$	λ = 4 to 10 per year
	Active per year	300 to 7,500
Prevalence of DoS attacks	Average from distribution = 2	About 2
	Prevalence = average * victim population = 2 * 25.3% of 1.59 mill.	About 760,000 per year
Prevalence of DoS hackers	$A = C/\lambda$	λ = 4 to 10 per year
	Active per year	76,000 to 190,000 or: 50,000 to 200,000 (oom)

NHTCU data:

Variable	Method of approximation	Approximate range of values
Prevalence of virus attacks	Average given as 255 Fraction experiencing = .77 C = 255 * .77 * 39,000	
	Correction for multiple reports times a spread factor = 100x100	750 to 70,000 viruses per year in the UK
Prevalence of virus writers	A = C/ λ	λ = 4 to 10 per year
	Active per year	75 to 17,500
Prevalence of DoS attacks	Average given as 5 Fraction experiencing = .17 C = 5 * .17 * 39,000	About 33,000 per year
Prevalence of DoS hackers	A = C/ λ	λ = 4 to 10 per year
	Active per year	3,000 to 8,000 or: 1,000 to 10,000 (oom)

DTI data:

Variable	Method of approximation	Approximate range of values
Prevalence of virus attacks (in the UK)	Average = 2.82 ~ 3 Fraction = .5 "universe" = 215,000	
	C = 322,500 per year (Spread correction of 100)	1,000 to 10,000
Prevalence of virus writers	A = C/ λ	λ = 4 to 10 per year
	Active per year	300 to 800 or: 100 to 1,000 (oom)
Prevalence of DoS attacks (in the UK)	Average = 10.7 ~ 11 Fraction = .07	
	C = 161,000	100,000 to 200,000
Prevalence of DoS hackers	A = C/ λ	λ = 4 to 10 per year
	Active per year	10,000 to 50,000