Technical Report
TR-2010-001

Wie repräsentativ sind die Messdaten eines Honeynet?*

Felix C. Freiling

Universität Mannheim
Fakultät für Mathematik und Informatik

Technischer Bericht TR-2010-001

10. Februar 2010

Zusammenfassung Zur Früherkennung von kritischen Netzphänomenen wurden in der Vergangenheit viele Arten von verteilten Sensornetze im Internet etabliert und erforscht. Wir betrachten das Phänomen "Verteilung von bösartiger Software im Netz", das punktuell etwa mit dem InMAS-Sensorsystem gemessen werden kann. Unklar war jedoch immer die Frage, wie repräsentativ die Daten sind, die durch ein solches Sensornetz gesammelt werden. In diesem Dokument wird ein methodisches Rahmenwerk beschrieben, mit dem Maßzahlen der Repräsentativität an Messungen von Malware-Sensornetzen geheftet werden können. Als methodischer Ansatz wurden Techniken der empirischen Sozialforschung verwendet. Als Ergebnis ist festzuhalten, dass ein Sensornetz mit mindestens 100 zufällig über den Netzbereich verteilten Sensoren notwendig erscheint, um überhaupt belastbare Aussagen über die Repräsentativität von Sensornetz-Messungen machen zu können.

1 Einführung

Aufgrund der Größe und Komplexität des Internet bilden Daten, die durch verteilte Sensor-Systeme wie CarmentiS [21], IAS [12], AMSEL [4], Nepenthes [6], Monk-it [31,30] oder InMAS [16,15] gesammelt werden, jeweils immer eine zeitlich und örtlich begrenzte Momentaufnahme des Aufkommens bösartiger Software. Derartige Sensornetze erlauben beispielsweise Aussagen der folgenden Art:

- 50% aller Sensoren haben bis jetzt Malware X registriert.
- Innerhalb der letzten Stunde haben 4 Sensoren Malware X registriert.

Zur Einschätzung der aktuellen Gefährdungslage sind aber Aussagen über den globalen Zustand des Internet viel interessanter. Insbesondere sind hier Kennzahlen relevant, die Aussagen über die Repräsentativität der gewonnenen Daten erlauben. Konkret sollen zwei Arten von Aussagen getroffen werden können:

- Absolute Zahlen: Wieviel Malware vom Typ X ist zur Zeit im Umlauf?
- Häufigkeitsverteilungen: Welcher Prozentsatz aller Rechner im deutschen Internet sind zur Zeit durch Malware X infiziert?

* Teile der Ergebnisse entstanden im Rahmen eines Projekts mit dem Bundesamt für Sicherheit in der Informationstechnik (BSI).

In diesem Bericht wird ein methodischer Ansatz vorgestellt, um lokale Messergebnisse von Sensorsystemen mit Methoden der empirischen Sozialforschung zu verallgemeinern.

In Abschnitt 2 gehen wir auf verwandte Arbeiten ein. Abschnitt 3 erläutert Werkzeuge der empirischen Sozialforschung, die im folgenden Abschnitt 4 auf das Problem der Repräsentativität von Sensornetz-Messungen angewendet werden. Abschnitt 5 fasst diesen Bericht kurz zusammen und gibt Empfehlungen zu weiteren Forschungsarbeiten und Experimenten.

2 Verwandte Arbeiten

In diesem Abschnitt gehen wir auf verwandte und relevante Arbeiten ein.

2.1 Epidemiologische Arbeiten

Die Analogien zwischen Computerviren und menschlichen Viren wurden bereits früh untersucht. Die mathematischen Modelle der Ausbreitung von Epidemien werden typischerweise durch eine Differentialgleichung beschrieben. In der einfachsten Form hat diese Gleichung die folgende Form:

$$\frac{dJ(t)}{dt} = \beta J(t)[N - J(t)] \tag{1}$$

Hierbei ist $J(t)$ die Zahl infizierter Individuen zum Zeitpunkt t, N ist die Gesamtgröße der betrachteten Population und β ist die Infektionsrate.

Eine erste Studie zur Anwendbarkeit von epidemiologischen Modellen auf die Verbreitung von Malware ist der Artikel “Directed graph epidemiological models for computer viruses” von Kephart und White [28]. Die relevante Forschung ist zusammengefasst im Buch von Daley [11].

2.2 Einzelanalysen bösartiger Software

Zur Analyse der Verbreitung von Malware gehört auch immer eine Analyse der Malware selbst. Es folgt eine kleine Auswahl von Artikeln, die bekannte und weniger bekannte Malware-Exemplare im Detail beschreiben.

- Der Internet-Wurm (1988, auch bekannt unter dem Namen *Morris Worm*) [37,14,36]: Der Internet-Wurm befiel 1988 einen Großteil der Rechner des damaligen Internet. Er nutzte bekannte Schwachstellen unter anderem in der Implementierung des Finger-Daemons und der Konfiguration der Software Sendmail aus und gilt als “Vater” aller modernen Würmer.
- ILOVEYOU E-Mail-Wurm (2000, auch bekannt unter dem Namen *loveletter*)) [7]: ILOVEYOU war der erste einer Folge von E-Mail-Makro-Würmer, die zur Verbreitung die Neugier und Ahnungslosigkeit der Computerbenutzer benötigen. Der Wurm bestand aus einer E-Mail mit einem Visual-Basic-Programm im Anhang. Beim Öffnen des Anhangs replizierte sich der Wurm auf dem Rechner des Benutzers und verschickte sich an eine Reihe von Addressaten aus dem Outlook-Addressbuch.

– CodeRed (2001) [40]: Der Wurm *CodeRed* nutzte eine Schwachstelle im Windows IIS unter Windows 2000 und existierte in zwei verschiedenen Varianten. Variante 1 besaß einen Programmierfehler im Programmteil, der Zufallszahlen erzeugte, und verbreitete sich nur schwach. Die zweite Variante war diejenige, die weltweit Probleme verursachte. Der Wurm generierte 100 parallele Threads, die parallel zufällig ausgewählte IP-Adressen im Internet über die IIS-Schwachstelle angriffen. Die Angriffsversuche wurden weltweit im Detail protokolliert, so dass relativ gute Vergleichsdaten für eine Modellierung zur Verfügung standen (siehe unten).
– Slammer (2003) [34]: Der Slammer-Wurm wird auch häufig als *SQL Slammer* bezeichnet, weil er eine Schwachstelle im SQL-Server 2000 von Microsoft ausnutzte. Slammer selbst bestand aus einem einzigen UDP-Paket und konnte sich so in einem Schritt verbreiten. Er verbreitete sich dadurch deutlich schneller als beispielsweise CodeRed. Slammer verwendete wie CodeRed *random scanning*, suchte seine Opfer also zufällig aus.
– Witty (2004) [35]: Genau wie Slammer verwendete der Witty-Wurm *random scanning* zur Verbreitung und war sehr klein. Eine Besonderheit von Witty ist, dass er sich ausschließlich im Speicher befand und keine Spuren auf der Festplatte hinterliess. Bemerkenswert ist zudem, dass er eine Schwachstelle in Produkten einer Computer*sicherheits*firma ausnutzte.
– Storm (2008) [25]: Storm ist ein Bot und Bestandteil des Storm Botnetzes, das erstmals 2008 in der Literatur beschrieben wurde. Storm ist ein Beispiel für Malware, die sich nicht autonom verbreitet. Die bisher untersuchten Verbreitungswege von Storm sind E-Mail (wie ILOVEYOU) oder bösartige Webseiten, die Schwachstellen im Webbrowser ausnutzen (*client-side exploits*). Storm ist einer der ersten Bot, der zur Koordination des Botnetzes Peer-to-Peer-Techniken einsetzt.

2.3 Empirische Studien zur Verbreitung bösartiger Software

Basierend auf empirischen Daten aus den Anti-Malware-Vorkehrungen ihres Betriebssystems veröffentlicht Microsoft halbjährlich eine Studie zu Malware-Aufkommen auf Windows-Rechnern [33]. Neben absoluten Zahlen enthalten diese Berichte auch wertvolle Hinweise zu aktuellen Malware-Trends.

Eine Studie von Göbel et al. [18] untersuchte die Wirksamkeit von Honeynets für Einbruchserkennung im Universitätsnetz der RWTH Aachen, ein Netz mit circa 40.000 Rechnern. In diesem Netz wurde am zentralen Router ein System zur Einbruchserkennung angeschlossen (der sogenannte “Blast-o-mat”), das infizierte Rechner durch charakteristische Kommunikationsmuster (*scanning*) identifizierte und in eine Art Quarantäne verschob. Gleichzeitig wurde in diesem System auf etwa 180 IP-Adressen Nepenthes installiert. Die Autoren berichten, dass mittels dieser 180 Honeypot-Sensoren *jeder* Scanning-Versuch erkannt wurde, der auch durch den Blast-o-mat registriert worden war.

2.4 Modellierung der Verbreitung bösartiger Software

Der Großteil der theoretischen Arbeiten zur Verbreitung bösartiger Software bezieht sich auf die Software, die sich *autonom* verbreitet. Das einschlägige akademische Forum hierfür ist der Workshop on Rapid Malcode (WORM), dessen Tagungsbeiträge

mehrheitlich online abrufbar sind [1,2,3]. Im Folgenden gehen wir auf einige dieser Arbeiten näher ein.

Basierend auf einem formalen Modell von Viren und Trojanern untersuchen Thimbleby et al. [38] grundsätzliche Fragen der Erkennbarkeit von Malware in Computersystemen.

In dem Artikel "Code Red Worm Propagation Modeling and Analysis" [40] untersuchen Zou et al. die Verbreitung des CodeRed-Wurms. Sie nutzen dabei die Eigenschaft von CodeRed, zufällig sein nächstes Opfer auszusuchen, und adaptieren das klassische epidemiologische Modell mit zwei Faktoren: (1) zunehmende administrative Gegenmaßnahmen (Ausbessern der Schwachstelle, Einstellen von Filtern, etc.) sowie (2) Verstopfungsphänomene in einigen Internet-Routern. Das Modell erreicht eine erstaunlich gute Abdeckung der tatsächlichen Verbreitung von CodeRed (siehe Abbildung 1), über die es umfangreiche Datenquellen gab.

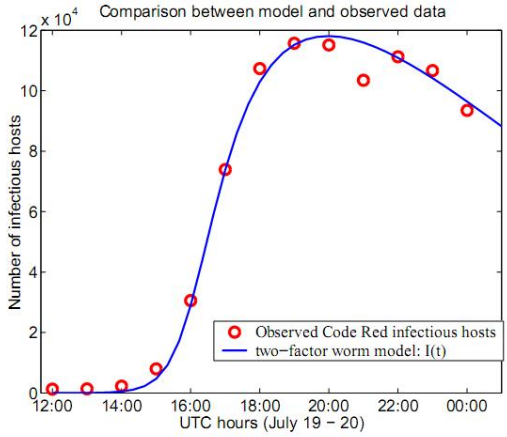
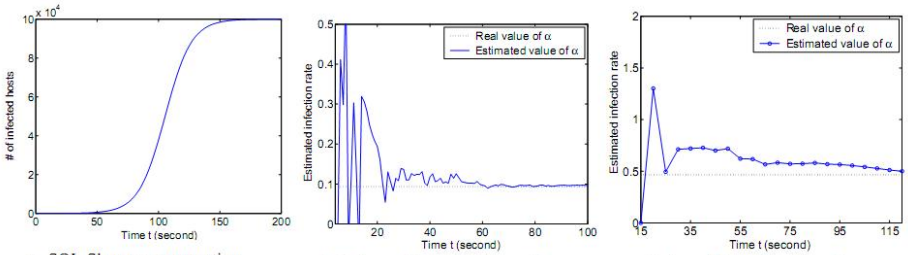


Abbildung 1: Vergleich von modellierter und tatsächlicher Verbreitung des CodeRed-Wurms (Abbildung 9 aus Zou et al. [40]).

Basierend auf Zou et al. [40] stellen Chen et al. [9] in ihrem Beitrag "Modeling the spread of active worms" ein zeitdiskretes Modell für die Verbreitung von CodeRed vor (das Modell von Zou et al. war zeitkontinuierlich). Chen et al. erweitern ihr Modell auf dahingehend, dass das lokale Suchen nach verwundbaren Rechnern (*local scanning*) in Betracht gezogen werden kann. Außerdem wird die Idee einer initialen *hit-list* modelliert, also die Tatsache, dass zu Beginn der Verbreitung mehrere Rechner mit dem Wurm infiziert sind. Auch wenn CodeRed selbst diese Techniken nicht anwendete, so verursachen doch gewisse Netzwerkkonfigurationen (beispielsweise einer Firewall), dass *scanning* und somit die Verbreitung des Wurms nur im lokalen Netz erfolgt.

In Reaktion auf CodeRed und Slammer beschäftigen sich Zou et al. [39] in ihrem Artikel "Monitoring and early warning for internet worms" mit Abwehrstrategien für Würmer. Die Idee der Autoren ist es, Erkennungstechniken aus dem Bereich der Signal-

verarbeitung auf Sensordaten anzuwenden, um frühzeitig Trends in Angriffsaktivitäten zu erkennen. Die Erkennungstechnik (ein Kalman-Filter) wird so eingestellt, dass er auf das Muster einer beginnenden epidemiologischen Ausbreitung mit einem "Ausschlag" reagiert. Die Messungen mit simulierten Daten zeigen, dass der Filter sich schon nach kurzer Zeit auf einen konstanten Wert stabilisiert (siehe Abbildung 2).



a. SQL Slammer propagation

b. Kalman filter estimation of α (time unit: 1 second)

c. Kalman filter estimation of α (time unit: 5 seconds)

Abbildung 2: Erkennung eines simulierten Ausbruchs des Slammer-Wurms mittels Kalman-Filter (Abbildung 10 aus Zou et al. [39]).

In einer umfassenden Studie namens "On the performance of Internet worm scanning strategies" untersuchen Zou et al. [42] verschiedene Verbreitungsstrategien von Würmern. Untersucht wurden unter anderem die folgenden Strategien:

- zufälliges Suchen (*random scan*), bei dem jeder Wurm eine zufällige IP-Adresse attackiert;
- sequentielles Suchen (*sequential scan*), bei dem ein Wurm einen IP-Adressbereich sequentiell attackiert;
- kooperatives Suchen, bei dem Würmer kommunizieren, um jede IP-Adresse möglichst nur ein Mal anzugreifen;
- Teile-und-Herrsche-Suchen (*divide and conquer scan*), bei dem erfolgreiche Würmer ihren Kindern immer eingeschränktere Teile des IP-Adressbereichs mitgeben (längere fixierte Präfixe), um das doppelte Angriffsarbeit zu vermeiden;
- lokales Suchen (*local preference scan*), bei dem das lokale Netz bevorzugt angegriffen wird.

Die optimale Angriffsstrategie für einen Wurm ist offensichtlich, in denjenigen Netzbereichen nach Opfern zu suchen, in denen diese mit großer Häufigkeit vorkommen. Zufällige Suche ist also nur bei gleichmäßiger Verteilung von verwundbaren Rechnern zu empfehlen. In diesem Fall sind jedoch auch sequentielles und Teile-und-Herrsche-Suchen genauso gut. Es stellte sich heraus, dass bei nicht-uniformer Verteilung von verwundbaren Rechnern lokales Suchen die Verbreitung des Wurms beschleunigte, insbesondere bei zunehmender Größe des lokalen Netzes.

In dem Artikel “Modeling botnet propagation using time zones” [10] beschreiben Dagon et al. ein Modell zur Erklärung des “pulsierenden” Verhaltens der Größe von Botnetzen. Die Autoren schlagen zunächst vor, das klassische epidemiologische Modell um das Konzept von Tages- und Nachtzeit zu ergänzen. Anschließend fügen sie noch das Konzept einer Zeitzone in das Modell ein. So kann beispielsweise modelliert werden, dass innerhalb einer Zeitzone tagsüber mehr Rechner aktiv sind als nachts, dass aber gleichermaßen zu jeder Zeit irgendwo auf der Welt Tag ist. Die Verbreitung bösartiger Software folgt dann sozusagen dem Verlauf der Sonne um die Erde. Die Modelle erlauben es, Vorhersagen zu treffen über die Verbreitung eines bestimmten Bots, wenn bekannt ist, zu welcher Zeit und wo er das erste Mal freigesetzt wurde.

Die Verbreitung von Bots erfolgt üblicherweise gezielt durch Befehle des Botnet-Betreibers. Obwohl es durch Infiltrationstechniken für Botnetze [17] relativ viele Protokolle derartiger Befehlssequenzen gibt, ist uns eine systematische Studie oder gar eine Modellierung dieser Ausbreitungswege nicht bekannt.

Die Ausbreitung von bösartiger Software, die sich nicht autonom verbreitet, ist bisher wenig untersucht. Beispiele für diese Art von Malware sind E-Mail-Würmer wie ILOVEYOU [7] oder Bots wie Storm [25], die sich durch Benutzerinteraktionen verbreiten. Bei der Modellierung der Ausbreitung von E-Mail-Würmern wurden im wesentlichen dieselben epidemiologischen Ansätze verwendet, wie bei sich autonom verbreitender Malware [41]. Es wurden jedoch *Ruhephasen* (*phases of dormancy*) in das Modell mit eingebracht, um die unbestimmte Verweildauer des Wurms im E-Mail-Postfach des Benutzers zu modellieren.

Noch weitgehend unerforscht sind Ausbreitungsmodelle für Malware, die Schwachstellen in Benutzerapplikationen wie Browsern ausnutzt. Erforscht wurden lediglich Techniken zur automatischen Sammlung und Analyse derartiger Malware [26]. Die Modellierung der Ausbreitung ist deutlich komplexer als die anderer Malware, weil die Ausbreitungsgeschwindigkeit von zusätzlichen Faktoren abhängt. Der Storm-Wurm verbreitete sich etwa über eine Kombination aus Spam-Versand und Benutzerinteraktion. Insbesondere der Faktor Mensch ist notorisch schwierig zu modellieren. Uns sind keine Arbeiten bekannt, die die Verbreitung von derartiger Malware modelliert.

3 Werkzeuge der empirischen Sozialforschung

Dieser Abschnitt stellt Werkzeuge der empirischen Sozialforschung vor, die im Weiteren verwendet werden sollen. Dieser Abschnitt stützt sich im Wesentlichen auf die Lehrbücher von Diekmann [13], Bortz [8], Lehn und Wegmann [29] sowie Hammann [24].

3.1 Repräsentativität in der empirischen Forschung

Grundsätzlich handelt es sich bei der Messung in einem Sensornetz immer um eine Art Stichprobe für die Messung eines umfassenderen, größeren Phänomens. Diese Fragestellung betrifft die *Repräsentativität* von Messungen in der empirischen Forschung. Ziel der empirischen Forschung ist, statistische Aussagen über eine (in der Regel große)

Menge von Objekten zu schaffen. Diese Menge wird auch als *Grundgesamtheit* bezeichnet. Eine *Stichprobe* ist hierbei eine Teilmenge der Grundgesamtheit. Die Messung einer Stichprobe ist repräsentativ, wenn sie Aussagen über die Grundgesamtheit erlaubt. In der empirischen Sozialforschung ist das Problem bekannt im Kontext repräsentativer Umfragen. In dieser Forschungsrichtung gibt es eine etablierte methodische Basis für Repräsentativitätsaussagen. Aus diesem Grund werden wir das vorgeschlagene Vorgehen an diesen Erkenntnissen ausrichten.

In der empirischen Forschung ist für die Repräsentativität die Auswahl der Stichprobe entscheidend. Statistisch ist die *Zufallsstichprobe* am besten untersucht. Die Zufallsstichprobe fußt auf einem Urnenmodell. Hierbei werden für eine Stichprobe der Größe n genau n Elemente der Grundgesamtheit gleichverteilt zufällig ausgewählt. Die Messung auf einer Zufallsstichprobe ist immer repräsentativ. Die Aussagegenauigkeit der Messung für die Grundgesamtheit steigt mit der Größe derr Zufallsstichprobe.

Da bei der Befragungen von Personen eine echt zufällige Auswahl sehr schwierig ist, gibt es in der empirischen Sozialforschung auch den Begriff der *Quotenstichprobe*. Hierbei werden die Elemente der Grundgesamtheit in Gruppen eingeteilt. Innerhalb der Gruppen wird eine Zufallsstichprobe getätigt. Die Größe der einzelnen Zufallsstichproben ist proportional zum Anteil der Gruppe an der Grundgesamtheit. Die Quotenstichprobe wird dazu benutzt, um bestimmte Strukturen der Grundgesamtheit in der Stichprobe nachzuahmen, wie beispielsweise die Verteilung des Lebensalters in der Bevölkerung.

3.2 Genauigkeit von Repräsentativitätsaussagen

Die Prüfung der Genauigkeit von Aussagen in der empirischen Forschung erfolgt meist durch Vergleich mit aus anderen Quellen bekannten Werten. Idealerweise erfolgt die Prüfung von Messwerten im Vergleich mit einer vollständigen Messung auf der Grundgesamtheit. Sind derartige Daten nicht verfügbar, versucht man durch entsprechende Stichprobengrößen den systematischen Fehler zu reduzieren. Üblich sind statistische Tests (Signifikanztests) auf den Ergebnissen einer Umfrage. Diese erlauben es, relativ schematisch die Größe einer Zufallsstichprobe zu bestimmen, damit bestimmte Qualitätsparameter eingehalten werden. Auf die weiteren Möglichkeiten, den Messfehler zu reduzieren, soll hier nicht eingegangen werden und wird auf andere Quellen verwiesen [23,22]

Die Qualitätsparameter einer statistischen Aussage sind:

1. die *Fehlerwahrscheinlichkeit* α, also die Wahrscheinlichkeit, dass die gemachte Aussage nicht zutrifft, sowie
2. die *Fehlerspanne* e, also die Genauigkeit der gemachten Aussage.

In der empirischen Forschung hat sich etabliert, sowohl für α also auch für e einen Wert von maximal 0,05 (5%) zu akzeptieren. In der Umfrageforschung wird der Parameter e vorrangig minimiert (es ist also wichtiger, nahe am wahren Wert zu sein, als ihn richtig zu treffen).

Beispiel: Messungen an einer Reihe von (zufällig verteilten) Honeypots ergeben, dass innerhalb der letzten Stunde 70% der Sensoren von einer Malware vom Typ X be-

fallen wurden. Eine Fehlerwahrscheinlichkeit von 5% bedeutet, dass 95 von 100 Messungen, die man mit einem zufällig verteilten Honeynet machen würde, den Infektionsprozentsatz richtig bestimmen. Eine Fehlerspanne von 5% bedeutet, dass der richtige Infektionsprozentsatz zwischen 65 und 75% liegt.

3.3 Größe einer Stichprobe

Generell gilt: je größer die Stichprobe, desto besser die Genauigkeit der gemessenen Werte. Im Extremfall erzielt man exakte Genauigkeit bei einer vollständigen Messung auf der Grundgesamtheit. Aus praktischen Gründen ist es jedoch sinnvoll, eine möglichst kleine Stichprobe zu wählen, die trotzdem die geforderte Mindestgenauigkeit garantiert.

Wenn über die gemessenen Häufigkeitsverteilung nichts bekannt ist, kann man die Stichprobengröße n über folgende einfache Formel berechnen:

$$n = 0.25(\tilde{\alpha}/e)^2 \tag{2}$$

Hierbei ist e wir oben genannt die Fehlerspanne (in Prozent) und $\tilde{\alpha}$ ein aus der Fehlerwahrscheinlichkeit α abgeleiteter Wert (siehe Tabelle 1). Für eine Fehlerspanne von 5% und 1% sind die so berechneten Stichprobengrößen in Tabelle 2 angegeben. Hierbei wird vereinfachend angenommen, dass die Häufigkeitsverteilung einer Normalverteilung entspricht.

α	$\tilde{\alpha}$
0,05	1,96
0,01	2,58

Tabelle 1: Werte für $\tilde{\alpha}$ für 5% und 1% Fehlerwahrscheinlichkeit α.

e	$\alpha = 0,05$	$\alpha = 0,01$
0,05	384	664
0,01	9604	16589

Tabelle 2: Stichprobengrößen, wenn über den zu messenden Sachverhalt keine Zusatzinformationen bekannt sind.

In der empirischen Sozialforschung kann man Kenntnisse über den zu messenden Sachverhalt ausnutzen, um die Stichprobengröße zum Teil deutlich zu verringern. Die Formel 2 geht davon aus, dass der zu messende Sachverhalt eine maximale Varianz hat. Eine große Varianz wirkt sich immer negativ auf die Stichprobengröße aus, weil es relativ vieler Messungen bedarf, um diese Varianz auch in der Stichprobe abzubilden.

Man kann auch die Kenntnis über die Häufigkeitsverteilung selbst verwenden, um die Stichprobengröße zu verkleinern. Meist ist diese aber unbekannt (man möchte sie ja

gerade messen, bzw. approximieren). Manchmal hat man jedoch aus anderen Quellen Informationen, die es erlauben, die zu messende Häufigkeitsverteilung zu schätzen.

Sei p die Häufigkeit des zu messenden Merkmals und $q = 1-p$ die Gegenhäufigkeit. Dann kann Formel 2 verfeinert werden zu:

$$n = (\tilde{\alpha}/e)^2 \cdot p \cdot q \tag{3}$$

Im schlimmsten Fall ist $p = 0,5$. In diesem Fall reduziert sich Formel 3 zu Formel 2. Je größer p ist, desto kleiner wird die errechnete Stichprobengröße.

Beispiel: Es sei aus anderer Quelle bekannt, dass etwa 30% aller Rechner in der Grundgesamtheit durch eine Malware X aktuell befallen sind. Mit $p = 0,3$ (und somit $q = 0,7$) ergibt sich die Formel:

$$n = 0,21(\tilde{\alpha}/e)^2$$

Für eine Fehlerwahrscheinlichkeit α von 5% und eine Fehlerspanne e von 5% benötigt man jetzt statt einer Stichprobengröße von 384 lediglich 323 Elemente.

Statt (aus anderer Quelle) bekannter Werte kann man auch bekannte Merkmalsverteilungen in der Berechnung der Stichprobengröße verwenden, von denen man annimmt, dass sie Einfluß haben auf den zu messenden Sachverhalt haben (also nicht die Merkmalsverteilung selbst sind). In Umfragen ist die Alters- oder Geschlechtsstruktur der Befragten häufig ein solches Merkmal, da die "wahre" Verteilung dieses Merkmals durch die Meldebehörden gemessen und zur Verfügung gestellt wird.

Quotenstichproben ermöglichen es zusätzlich, die Stichprobengrösße insgesamt zu reduzieren. Technisch verringert man durch die Quotierung die Varianz des zu messenden Merkmals. Man kann sogar Merkmale kombinieren und mehrstufige Quotenstichproben durchführen. In Umfragen wird häufig das Alter mit dem Wohnort kombiniert, also Merkmale, über die man recht genaue Verteilungen kennt, die somit also eine geringe Varianz besitzen.

4 Problemstellung

In diesem Abschnitt wird die Ausgangsfragestellung konkretisiert. Allgemein ist das Ziel, die Messung eines Phänomens A anhand eines Phänomens B unter Verwendung der Instrumente der empirischen Sozialforschung. Die zu messenden Phänomene betreffen im wesentlich zwei Arten von Aussagen:

- Absolute Zahlen, beispielsweise: Wieviel Malware eines bestimmten Typs ist zur Zeit im deutschen Internet im Umlauf?
- Häufigkeitsverteilungen, beispielsweise: Welcher Prozentsatz aller Rechner im deutschen Internet sind zur Zeit durch Malware eines bestimmten Typs infiziert?

Methodisch sollen die Messexperimente wie repräsentative Umfragen behandelt werden. Das Phänomen B ist beispielsweise die Anzahl von Honeypot-Infektionen durch eine bestimmte Malware, die durch die Malware-Sensorik aufgezeichnet wird. Diese Messungen sollen dann für alle Rechner in einem bestimmten Netzbereich hochgerechnet werden. Dies erlaubt es dann, Aussagen über die Anzahl der Infektionen eines Netzbereichs durch diese Malware zu machen (Phänomen A).

Bei der Übertragung der Methoden der empirischen Sozialforschung sind einige grundlegende Fragen zu klären, die wir in diesem Abschnitt diskutieren.

4.1 Was ist die Grundgesamtheit?

Die Grundgesamtheit ist die Menge aller Objekte, über die eine Aussage getroffen werden soll. Grob gesprochen, sind dies alle "Computer" im deutschen Internet. Diese Menge ist schwer bestimmbar, da es keinerlei Meldepflicht für am Netz angeschlossene Computer gibt. Es müssen zudem folgende Fragen geklärt werden:

- Sollen Laptops gezählt werden, die möglicherweise nur temporär am (deutschen) Netz sind?
- Sollen Mobiltelefone und PDAs gezählt werden?
- Wie sieht es mit anderer Netzwerkhardware aus (Router, Switches)?
- Wie soll man mit virtuellen Maschinen umgehen? Insbesondere in Rechenzentren werden statt "echten" Computern lediglich virtualisierte Server eingesetzt.
- Sollen die Honeypot-Sensoren selbst mitgezählt werden?

Möglich ist auch eine Definition der Grundgesamtheit, die auf Netzwerkeigenschaften beruht. Relativ gut abrenzbar sind beispielsweise diejenigen Teile der physischen Netzinfrastruktur, die innerhalb der Bundesrepublik Deutschland befinden. Aber auch diese Abgrenzung ist im Hinblick auf des WWW problematisch, da URLs mit Endung `.de` nicht notwendigerweise auf Rechner aufgelöst werden, die geographisch in Deutschland stehen.

Sinnvoll erscheint auch die Möglichkeit, diejenigen IP-Adressen zu zählen, die durch in Deutschland gemeldete Firmen oder Privatpersonen angemeldet sind. Die Registrierungsinformationen von RIPE bzw. DENIC können hierzu ausgewertet werden.

Bei dieser Zählung ist allerdings problematisch, da die Zuordnung von IP-Adresse zu Computer nicht eineindeutig ist, da beispielsweise in dynamischen DHCP-Netzbereichen die Zuordnung von IP-Adresse zu Rechner ständig wechselt und es auch IP-Adressbereiche gibt, die zwar in Deutschland registriert, denen aber gar keine Rechner zugeordnet sind.

Sinnvoll erscheint auch die Zählung von registrierten lokalen Netzen mit gewissen Eigenschaften. Die Menge aller registrierten Firmennetze mit mindestens 100 Rechnern beispielsweise erscheint eine relativ klar abgrenzbare Grundgesamtheit zu sein.

4.2 Welches Merkmal wird gemessen?

Wie eingangs genannt kommen zwei Arten von Aussagen in Frage: absoluten Zahlen und Häufigkeitsverteilungen. Bei entsprechender Repräsentativität einer Stichprobe geben gemessene Häufigkeiten auch die "wahren" Häufigkeiten wieder. Absolute Zahlen müssen aus der Stichprobe auf die Grundgesamtheit hochgerechnet werden.

Im Kontext der einschlägigen Sensor-Infrastrukturen [6,15,4] geht es vor allem um Infektionen mit (einer bestimmten) bösartiger Software. Hierzu müssen folgende Punkte beachtet werden:

- Die Tatsache einer Infektion muss eindeutig definierbar sein. Dies geht beispielsweise durch die Beobachtung einer Infektion durch einen Antivirusscanner.

– Wenn man Zahlen bezüglich einer bestimmten bösartigen Software errechnen möchte, muss diese Software genau abgrenzbar sein. Aufgrund der hohen Zahl von Varianten von Bots oder Tojanern erscheint dies relativ schwierig. Oft ist zudem auf den ersten Blick gar nicht klar, ob zwei Exemplare von bösartiger Software zur gleichen Klasse gehören oder nicht.
– Da Rechner üblicherweise nicht dauerhaft infiziert bleiben, muss auch der zeitliche Aspekt der Infektion in die Fragestellung mit einfliessen. Geht es beispielsweise um die Anzahl aller Rechner, die *jemals* mit einer bestimmten Malware infiziert waren? Oder geht es um Rechner, die innerhalb der letzten Stunde, des letzten Tages oder der letzten Woche mit einer bestimmten Malware infiziert waren?

4.3 Was ist eine Zufallsstichprobe?

In der empirischen Sozialforschung werden die Probanden zufällig ausgewählt und befragt. In der Übertragung der Analogie müsste man vor der Messung Elemente der Grundgesamtheit zufällig auswählen und testen. Dies ist methodisch nicht möglich, weil man nicht "im nachhinein" einen normalen Rechner durch einen Malware-Sensor ersetzen kann. Per Definition erfüllt ein Honeypot-Sensor ja auch keinen sinnvolen Zweck im Netzwerk. Insofern kann man auch normale Produktivrechner nicht als Honeypots ansehen.

Interessant ist die Überlegung, normale Rechner für die Zeiten, in denen sie nicht im Produktivbetrieb sind, als Honeypots zu instrumentieren. Beispielsweise könnte beim Start des Bildschirmschoners und bei entsprechend geringer Last der Rechner in einen Honeypot-Modus wechseln. Diese Idee ist jedoch noch nicht einmal im Ansatz umgesetzt.

Generell erscheint es sinnvoll, die Malware-Sensoren möglichst *gleichmäßig* über die Grundgesamtheit zu verteilen. Zu diesem Schluss kommt auch die Literatur. So schreiben Zou et al. [39]:

> "... distributing monitors as uniformly as possible, we can achieve an effective sampling of worm activities. Since worms might choose different destinatioin addresses by using different preferences, [...] we need to use multiple address blocks with different sizes and characteristics to provide proper coverage."

Eine gleichmäßige und großflächige Verteilung der Sensoren erscheint auch sinnvoll, um die nicht-uniforme Ausbreitung von bösartiger Software gut zu registrieren:

> "The address space covered by a monitoring system should be as distributed as possible to accurately monitor the propagation of a non-uniform scan worm, especially a sequential scan worm such as Blaster." [42]

Aus methodischer Sicht ist es etwas problematisch, wenn die Honeypot-Sensoren statisch verteilt sind. Übertragen auf Befragungen bedeutet das, dass immer dieselben Personen befragt werden würden. Vergleichbar ist dieser Ansatz mit dem Verfahren zur Ermittlung von Einschaltquoten, bei der auch die Fernsehgeräte einer festen Menge von Haushalten mit Sensorik ausgestattet ist. In Deutschland besteht dieses sogenannte

Fernsehpanel aus knapp 6000 Haushalten, die über die gesamte Bundesrepublik verteilt zufällig ausgewählt wurden [5].

Problematisch ist auch die Möglichkeit von Angreifern, Honeypot-Sensoren systematisch zu umgehen. Wenn im Zuge von Aufklärungsaktivitäten etwa die IP-Adressen von Sensoren öffentlich bekannt werden würden, könnte bösartiger Software mittels einer *blacklist* diese Sensoren systematisch meiden. Dies ist ein weiteres Argument, warum Sensoren regelmäßig neu plaziert werden sollten.

Das Konzept einer Zufallsstichprobe im Kontext der hier vorgeschlagenen Methode ist also höchst problematisch.

4.4 Welche Stichprobengröße muss gewählt werden?

Es gibt keinen Grund, bei der Ermittlung der Stichprobengröße von den Methoden der empirischen Sozialforschung abzuweichen. Mit den in Abschnitt 3 beschriebenen Formeln ergeben sich methodisch abgesicherte Größen für die Zufallsstichprobe. Insbesondere die Zahlen in Tabelle 2 erscheinen bei zufälliger Verteilung ausreichend.

Jedoch ist zu beachten, dass uns keine Mess-Infrastruktur bekannt ist, deren Zahl von Sensoren auch nur annähernd in die geforderten Mindestzahlen bei gleichzeitiger zufälliger Verteilung erreicht. Die bisher in Projekten erzielten Messungen [27] erfolgten etwa bei einer Verteilung über mehrere Provider mit nicht einmal einem Dutzend Sensoren. Die Messungen mit Nepenthes im Honeynet der RWTH Aachen erfolgen zwar mit mehreren Tausend IP-Adressen. Diese sind jedoch auf ein einzelnes Class B-Netz beschränkt [19].

Durch Quotierung kann man diese Größen der Stichproben noch zusätzlich verringern. Hierzu können Kenntnisse über die Zusammensetzung relevanter Merkmale der Grundgesamtheit verwendet werden. Auch Schätzungen von relevanten Parametern dieser Merkmale aus Modellen können hier wertvoll sein (siehe Abschnitt 2). Jedoch sind in der empirischen Sozialforschung alle Messungen mit weniger als 30 Teilnehmern suspekt.

4.5 Welche Quotierungen sind denkbar?

Relevante Merkmale zur Quotierung der Elemente der Grundgesamtheit müssen einen Einfluss auf die Möglichkeit haben, von einer bestimmten Malware infiziert zu werden. Hierzu zählen:

- Der Hersteller und die Version des Betriebssystems.
- Die Verteilung der Elemente der Grundgesamtheit auf lokale Netze oder autonome Systeme.
- Die geographische Verteilung der Elemente der Grundgesamtheit.

Generell ist festzustellen, dass die Quotierung und damit die Größe der Stichprobe auch von der Art der zu messenden bösartigen Software abhängt. Beispiel: Um eine bösartige Software zu messen, die ausschließlich Windows 2000-Systeme im Raum Hannover angreift (dort aber zufällig), reicht eine relativ kleine Stichprobe aus (zufällig ausgewählte Windows 2000-Rechner im Raum Hannover).

Die Flexibilität bei der Quotierung ist ein weiterer Grund, eine möglichst große und gut verteilte Anzahl von Sensoren zu besitzen. Mittels einer Quotierung kann man dann sogar in begrenztem Umfang derartige geographisch oder systemspezifisch eingeschränkte "Individualphänomene" messen.

5 Zusammenfassung und Empfehlungen

5.1 Zusammenfassung

In diesem Dokument haben wir ein methodisches Rahmenwerk beschrieben, mit dem Maßzahlen der Repräsentativität an Messungen von Malware-Sensornetzen geheftet werden können. Als methodischer Ansatz wurden Techniken der empirischen Sozialforschung verwendet.

Für repräsentative Aussagen über die Verbreitung von Malware müssen folgende Schritte durchgeführt werden:

1. Bestimmung des zu messenden Merkmals.
2. Bestimmung der Stichprobengröße unter Verwendung von Quotierung und bekannten Informationen über (andere relevante) Merkmale.
3. Durchführung der Messung und ggf. Hochrechnen des Ergebnisses.

Punkt 2 ist Grundlage für die Quantifizierung der Verlässlichkeit im Sinne von Fehlerwahrscheinlichkeit und Fehlerspanne (siehe Abschnitt 3).

In die Berechnungen zu Punkt 2 fließen idealerweise folgende Zusatzinformationen ein:

- Informationen über Merkmale wie die typischen Ausbreitungskurven von Malware aus theoretischen Modellen.
- Informationen von anderen Mess-Infrastrukturen (etwa IAS [12] oder Monk-it [31,30] über die Zusammensetzung des Netzwerkverkehrs zur Quotierung.

Die direkte Messung der Verbreitung von Malware durch IAS und monk-it ist nicht praktikabel.

Insgesamt erscheinen die Methoden der empirischen Sozialforschung geeignet, um Repräsentativitätsaussagen der InMAS-Messungen zu beziffern.

5.2 Empfehlungen

Die folgenden Empfehlungen können unabhängig voneinander umgesetzt werden.

Etablierung eines großen Sensornetzes. Um realistische Ergebnisse zu erzielen, muss ein großes Sensornetz im untersuchten Netzbereich eingerichtet werden. Dabei müssen die Elemente der Grundgesamtheit im Sinne einer Zufallsauswahl möglichst gleichmäßig abgedeckt werden.

Am sinnvollsten erscheint uns die Definition der Grundgesamtheit nach lokalen Netzen einer bestimmten Größe. Die Quotierung kann anschliessend nach der Art des installierten Betriebssystems erfolgen. Voraussichtlich sind hierdurch handhabbare Stichprobengrößen herleitbar, auch für unterschiedliche zu messende Merkmale.

Für sinnvolle Repräsentativitätsaussagen ist voraussichtlich ein Sensornetz mit mindestens 100 Sensoren notwendig.

Präzisierung von Modellen autonomer Malware. Die in Abschnitt 2 beschriebene Literatur enthält zum Teil schon sehr weit fortgeschrittene Modelle für die Verbreitung von autonomer Malware. Diese Modelle wurden jedoch jeweils in Folge eines konkreten Wurmausbruchs erstellt und an die historischen Daten angepasst. Es ist unklar, ob die Modelle auch für andere Arten von Malware zutreffend sind.

Zur Präzisierung der existierenden Modelle schlagen wir folgenden experimentellen Aufbau vor: Ein virtuelles Netz mit x IP-Adressen wird aufgesetzt. An jeder IP-Adresse lauscht Nepenthes. Wird ein Nepenthes-Sensor infiziert, ersetzen wir den Sensor mit einem infizierten Rechner (in Form einer virtuellen Maschine). Nach einer wählbaren Zeit wird der infizierte Rechner wieder durch einen Nepenthes-Sensor ersetzt (*cleaning*). Initial wird eine bestimmte (echte) Malware auf einer vorgegebenen Anzahl von IP-Adressen ausgesetzt. Anschliessend beobachtet man die Ausbreitung der Malware in Echtzeit.

Mittels eines solchen experimentellen Aufbaus ist es möglich, Messungen über die Verbreitung von Malware mittels "echter" Malware im Labor nachzuvollziehen (eine Art verhaltensbasierte Verbreitungsanalyse). Man kann damit auch eine große Anzahl von Malware-Exemplaren in automatisierter Form untersuchen und die Ergebnisse mit den theoretischen Modellen vergleichen.

Erforschung von Benutzermodellen und Malware, die sich nicht autonom verbreitet. Es gibt einen evidenten Forschungsbedarf zu Verbreitungsmodellen von Malware, die Benutzerapplikationen angreift. Dafür ist die Erarbeitung entsprechender Benutzermodelle notwendig. Ansatzpunkte hierfür sind beispielsweise:

- Die Studie von McCoy et al. [32] über das Verhalten von Benutzern des Tor-Netzwerks.
- Die allgemeine Untersuchung von Benutzerverhalten im WWW, etwa durch Beobachtung von Testpersonen, die Analyse von veröffentlichten Bookmarks oder die Analyse von populären Suchbegriffen (etwa bei Google Zeitgeist [20]).

Mit diesen Informationen kann man eine Software konstruieren, die auf Basis verschiedener Parameter einen menschlichen Surfer simuliert. Infektionen, die durch die Aktivitäten dieses Benutzers hervorgerufen werden, können dann mit sogenannten Anwendungs-Honeypots (*client-side honeypots*) erkannt und analysiert werden.

Literatur

1. Workshop on Rapid Malcode. Internet: `http://www1.cs.columbia.edu/~angelos/worm05/`, November 2005. George Mason University, Fairfax, USA.
2. Workshop on Rapid Malcode. Internet: `http://www.eecs.umich.edu/~farnam/worm2006.html`, November 2006. George Mason University, Fairfax, USA.
3. Workshop on Rapid Malcode. Internet: `http://www.auto.tuwien.ac.at/~chris/worm07.html`, 2007. TU Wien.
4. Martin Apel, Joachim Biskup, Ulrich Flegel, and Michael Meier. Early warning system on a national level — Project AMSEL. In *Proceedings of the First Workshop on Early Warning and Network Intelligence (EWNI)*, Hamburg, Germany, January 2010.

5. Arbeitsgemeinschaft Fernsehforschung. Stichprobe und Anwerbung. Internet: `http://www.agf.de/fsforschung/methoden/stichprobe`, 2008.
6. Paul Baecher, Markus Koetter, Thorsten Holz, Maximillian Dornseif, and Felix C. Freiling. The nepenthes platform: An efficient approach to collect malware. In Diego Zamboni and Christopher Krügel, editors, *Recent Advances in Intrusion Detection, 9th International Symposium, RAID 2006, Hamburg, Germany, September 20-22, 2006, Proceedings*, volume 4219 of *Lecture Notes in Computer Science*, pages 165–184. Springer, 2006.
7. Matt Bishop. Analysis of the ILOVEYOU worm. Internet: `http://nob.cs.ucdavis.edu/classes/ecs155-2005-04/handouts/iloveyou.pdf`, May 2000.
8. J. Bortz. *Statistik für Sozialwissenschaftler*. Springer, 5. auflage edition, 1999.
9. Z. Chen, L. Gao, and K. Kwiat. Modeling the spread of active worms. In *IEEE INFOCOMM*, 2003.
10. David Dagon, Cliff Zou, and Wenke Lee. Modeling botnet propagation using time zones. In *NDSS*. The Internet Society, 2006.
11. D. J. Daley and J. Gani. *Epidemic Modeling: An Introduction*. Cambridge University Press, Cambridge, UK, 1999.
12. Matthias Deml and Norbert Pohlmann. Internet early warning system — overview and architecture. In *Proceedings of the First Workshop on Early Warning and Network Intelligence (EWNI)*, Hamburg, Germany, January 2010.
13. Andreas Diekmann. *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen*. Rowohlt, 18. auflage edition, August 2007.
14. M. W. Eichin and J. A. Rochlis. With microscope and tweezers: An analysis of the internet virus of november 1988. In *Proc. IEEE Symposium on Security and Privacy*, pages 326–343, 1989.
15. Markus Engelberth, Felix C. Freiling, Jan Göbel, Christian Gorecki, Thorsten Holz, Ralf Hund, Philipp Trinius, and Carsten Willemsa. The InMAS Approach. In *Proceedings of the First Workshop on Early Warning and Network Intelligence (EWNI)*, Hamburg, Germany, January 2010.
16. Markus Engelberth, Felix C. Freiling, Jan Göbel, Christian Gorecki, Thorsten Holz, Philipp Trinius, and Carsten Willems. Frühe Warnung durch Beobachten und Verfolgen von bösartiger Software im Deutschen Internet: Das Internet-Malware-Analyse System (InMAS). In *11. Deutscher IT-Sicherheitskongress*, Bonn, Germany, May 2009.
17. Felix C. Freiling, Thorsten Holz, and Georg Wicherski. Botnet tracking: Exploring a root-cause methodology to prevent distributed denial-of-service attacks. In Sabrina De Capitani di Vimercati, Paul F. Syverson, and Dieter Gollmann, editors, *Computer Security - ESORICS 2005, 10th European Symposium on Research in Computer Security, Milan, Italy, September 12-14, 2005, Proceedings*, volume 3679 of *Lecture Notes in Computer Science*, pages 319–335. Springer, 2005.
18. Jan Göbel, Jens Hektor, and Thorsten Holz. Advanced honeypot-based intrusion detection. *;login*, 31(6):17–25, December 2006.
19. Jan Goebel, Thorsten Holz, and Carsten Willems. Measurement and analysis of autonomous spreading malware in a university environment. In Bernhard M. Hämmerli and Robin Sommer, editors, *Detection of Intrusions and Malware, and Vulnerability Assessment, 4th International Conference, DIMVA 2007, Lucerne, Switzerland, July 12-13, 2007, Proceedings*, volume 4579 of *Lecture Notes in Computer Science*, pages 109–128. Springer, 2007.
20. Google Inc. Google Zeitgeist. Internet: `http://www.google.de/press/zeitgeist.html`, 2008.
21. Bernd Grobauer, Jens Ingo Mehlau, and Jürgen Sander. Carmentis: A co-operative approach towards situation awareness and early warning for the internet. In *IT Incidents Management and IT Forensics – IMF 2006, Conference Proceedings*, pages 55–66, October 2006.

22. Robert M. Groves. Research on survey data quality. *Public Opinion Quarterly*, 51(2):156–172, 1987.
23. Robert M. Groves. *Survey Errors and Survey Costs*. Wiley, 1989.
24. P. Hammann and B. Eichson. *Marktforschung*. UTB, 3rd edition, 1994.
25. Thorsten Holz, Felix Freiling, Moritz Steiner, Frederic Dahl, and Ernst Biersack. Measurements and mitigation of peer-to-peer-based botnets: A case study on storm worm. *Proceedings of the First USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET 08)*, 2008.
26. Ali Ikinci, Thorsten Holz, and Felix C. Freiling. Monkey-spider: Detecting malicious websites with low-interaction honeyclients. In Ammar Alkassar and Jörg H. Siekmann, editors, *Sicherheit 2008: Sicherheit, Schutz und Zuverlässigkeit. Konferenzband der 4. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI), 2.-4. April 2008 im Saarbrücker Schloss*, volume 128 of *LNI*, pages 407–421. GI, 2008.
27. Internet Malware Analyse System (InMAS). NE1: Bedrohungserkennung und -analyse Netzwerk (BEAN). December 2007.
28. J. O. Kephart and S. R. White. Directed-graph epidemiological models of computer viruses. In *Proc. IEEE Symposium on Research in Security and Privacy*, pages 343–359, 1991.
29. Jürgen Lehn and Helmut Wegmann. *Einführung in die Statistik*. Vieweg, 5th edition, June 2006.
30. Tobias Limmer and Falko Dressler. Survey of Event Correlation Techniques for Attack Detection in Early Warning Systems. Technical Report 01/08, University of Erlangen, Dept. of Computer Science 7, April 2008.
31. Tobias Limmer and Falko Dressler. Seamless Dynamic Reconfiguration of Flow Meters: Requirements and Solutions. In *16. GI/ITG Fachtagung Kommunikation in Verteilten Systemen (KiVS 2009)*, Kassel, Germany, March 2009. Springer. to appear.
32. D. McCoy, K. Bauer, D. Grunwald, T. Kohno, and D. Sicker. Shining light in dark places: Understaning the Tor network. In *Privacy Enhancing Technologies Symposium*, 2008.
33. Microsoft. Microsoft security intelligence report. Internet: `http://www.microsoft.com/sir`, 2008.
34. David Moore, Vern Paxson, Stefan Savage, Colleen Shannon, Stuart Staniford-Chen, and Nicholas Weaver. Inside the slammer worm. *IEEE Security & Privacy*, 1(4):33–39, 2003.
35. Colleen Shannon and David Moore. The spread of the witty worm. *IEEE Security & Privacy*, 2(4):46–50, 2004.
36. Eugene H. Spafford. The Internet worm: Crisis and aftermath. *Communications of the ACM*, 32(6):678–687, June 1989.
37. Eugene H. Spafford. The Internet worm program: An analysis. *Computer Communications*, 19(1):17–57, Januar 1989.
38. H. Thimbleby, S. Anderson, and P. Cairns. A framework for modelling trojans and computer virus infections. *The Computer Journal*, 41(7):444–458, 1998.
39. Cliff Changchun Zou, Lixin Gao, Weibo Gong, and Don Towsley. Monitoring and early warning for Internet worms. In Vijay Atluri and Peng Liu, editors, *Proceedings of the 10th ACM Conference on Computer and Communication Security (CCS-03)*, pages 190–199, New York, October 27–30 2003. ACM Press.
40. Cliff Changchun Zou, Weibo Gong, and Don Towsley. Code red worm propagation modeling and analysis. In Ravi Sandhu, editor, *Proceedings of the 9th ACM Conference on Computer and Communications Security*, pages 138–147, Washington, DC, USA, November 2002. ACM Press.
41. Cliff Changchun Zou, Donald F. Towsley, and Weibo Gong. Email worms modeling and defense. In Ronald P. Luijten, Luiz A. DaSilva, and Antonius P. J. Engbersen, editors, *Proceedings of the International Conference On Computer Communications and Networks (ICCCN 2004), October 11-13, 2004, Chicago, IL, USA*, pages 409–414. IEEE, 2004.

42. Cliff Changchun Zou, Donald F. Towsley, and Weibo Gong. On the performance of Internet worm scanning strategies. *Perform. Eval*, 63(7):700–723, 2006.