A Study of Malware in Peer-to-Peer Networks / Malware Prevalence in the KaZaA File-Sharing Network

Gregor Kopf
(kopf@informatik.hu-berlin.de)

Seminar „Internet Measurement“,
Technische Universität Berlin

SS 2007 (Version vom 2. Juli 2007)

Zusammenfassung

Dieses Papier ist eine Zusammenfassung von [KAM 06] und [SJB 06].
Beschrieben wird der momentane Stand der Verbreitung von Malware in verschiedenen Peer-to-Peer-Netzwerken. Dabei werden die Häufigkeiten von Malware, Suchanfragen, die besonders viel Malware als Antworten generieren, sowie die Verteilung der Malware im Netz und die genaue Funktion der heruntergeladenen Malware untersucht.
Die Messergebnisse besagen, dass in vielen Netzen mit einer Wahrscheinlichkeit von über 30% die heruntergeladenen Binärdateien mit Malware verseucht sind.
Die genaue Verteilung der Malware wird analysiert, sowie deren Funktion näher beleuchtet. Es stellt sich heraus, dass ein Grossteil der von Malware infizierten Rechner unter anderem zum Spamversand genutzt wird.
Daraufhin werden das Schadensausmaß, sowie mögliche Gegenmaßnahmen diskutiert. Mögliche Gegenmaßnahmen bestehen in einer Heuristik zur Erkennung von Malware anhand der Dateigröße oder der Ausstattung von Peer-to-Peer-Clients mit Virenscannern. Weiterhin schlage ich ein verteiltes System zur Bewertung von Dateiinhalten (bzgl. ihrer Infektion mit Malware) vor.

1 Einleitung

Wie viele andere Strukturen, werden auch Peer-to-Peer-Netze von Missbrauch durch Malware nicht verschont. Neben Würmern, die die Infrastruktur des Netzes nutzen, um sich selbstständig zu verbreiten, stellen auch Viren eine Gefahr dar:
Wurden früher vermehrt nicht ausführbare Dateien wie MP3 getauscht, so haben sich diese Netze zu Tauschbörsen für Dateien aller Art entwickelt. Durch die Möglichkeit, ausführbare Dateien zum Download anzubieten, bietet sich ein Angriffsvektor für Viren, die im Gegensatz zu Würmern auf Wirtsdateien angewiesen sind. Insbesondere sind Viren zu beobachten, die speziell für den Einsatz in Peer-to-Peer-Netzen ausgelegt sind.
Durch die Verbreitung solcher Viren können viele Gefahren entstehen: Vom Aufbau von Botnetzen über Spamversand bis zur Zerstörung von wichtigen Dateien ist alles denkbar.

Dieses Papier ist eine Zusammenfassung der Papiere [KAM 06] und [SJB 06] und beschäftigt sich mit der Ausbreitung solcher Malware in den gängigen Netzen KaZaA (FastTrack), Gnutella und OpenFT. In Abschnitt 2 werden einige technische Grundlagen

erläutert, darauf folgt in den Abschnitten 3 und 4 eine Beschreibung der zur Messung geschriebenen oder angepassten Software und deren Einsatz. Abschnitt 5 gibt einen Überblick über die gesammelten Messwerte. Darauf folgt in Abschnitt 6 eine Untersuchung der entstehenden Gefahren, sowie eine Diskussion möglicher Gegenmaßnahmen.

2 Technische Grundlagen

Um die zur Messung verwendeten Methoden besser zu verstehen, ist es hilfreich, einen Überblick über einige Grundlagen von Peer-to-Peer-Netzen zu haben. In fast allen Netzen ist es notwendig, eine gewünschte Datei zuerst im Netz zu suchen, bevor sie heruntergeladen werden kann. Ohne die Suchfunktion eines Peer-to-Peer-Netzes zu verwenden, ist es also nicht ohne weiteres möglich, überhaupt Dateien herunterzuladen und zu analysieren.
Da die Suchfunktion der meisten Clients aber nicht auf ein automatisches Durchsuchen des Netzes nach Malware ausgelegt ist, müssen diese zur Analyse der Virenverbreitung angepasst werden. Dazu ist die genaue Funktionsweise der Suchfunktion hier von vorrangigem Interesse:

Alle drei hier untersuchten Netze weisen eine Supernode-Architektur auf. Dabei wird eine Suchanfrage nicht durch das gesamte Netz an alle Knoten weitergeleitet, sondern spezielle Knoten –sogenannte Supernodes– sind für die Organisation der Suche verantwortlich:
Jeder Knoten im Netz verbindet sich bei seiner Initialisierung mit einem Supernode und überträgt dabei eine Liste seiner angebotenen Dateien.
Somit sind zum Durchsuchen des Netzes nur noch die Supernodes notwendig, da sie über die vollständige Information bzgl. der angebotenen Dateien verfügen.
Wird jetzt eine Suche gestartet, so wird sie vom Anfragesteller zu seinem Supernode übertragen. Dieser Supernode durchsucht seine Liste von angebotenen Dateien und übermittelt dem Client die ggf. gefundenen Dateien. Dann leitet er die Anfrage weiter an andere Supernodes, die ebenso verfahren. Somit wird eine Suchanfrage im Regelfall an alle erreichbaren Supernodes im Netz weitergeleitet. Die Suchergebnisse werden dann wieder zurück an den ursprünglichen Supernode und von diesem an seinen Client übermittelt.
Das führt einerseits zu einem geringen Durchmesser des Netzes, andererseits spart es Bandbreite, da nur die Supernodes für die Beantwortung von Suchanfragen zuständig sind.

3 Messwerkzeuge

Da die meisten Peer-to-Peer-Clients keine Möglichkeit bieten, das Netz voll automatisch auf eine zur Analyse der Virenverteilung geeigneten Art zu durchsuchen, wurde für jedes der Protokolle spezielle Software geschrieben oder zumindest bestehende erweitert.

3.1 KaZaA

Da der offizielle KaZaA-Client nicht quelloffen ist wurde eine spezielle Software names Krawler entwickelt, um das Netz auf geeignete Weise durchsuchen zu können. Krawler besteht aus zwei Teilen: dem *dispatcher*, der eine Liste von bekannten Supernodes verwaltet und dem *fetcher*, der mit dem Dispatcher und den anderen Supernodes im Netzwerk kommuniziert. Krawler beginnt mit einer fest einprogrammierten Liste von

200 bekannten Supernodes im FastTrack Netz und sendet Suchanfragen an eben diese. Nach einer Überprüfung, ob die bekannten Supernodes antworten (durch Aufbau einer TCP-Verbindung), versucht das Programm, eine *supernode refresh list* von jedem der Supernodes zu erhalten, um an bis zu 200 weitere Adressen von aktiven Supernodes zu kommen. Leider macht das Papier [SJB 06] über die Downloadstrategie und den verwendeten Virenscanner keine detaillierten Aussagen.

3.2 Gnutella und OpenFT

Zur Analyse des Malwareaufkommens im Gnutella-Netz wurde der beliebte Client Limewire um einige Funktionen erweitert: Das veränderte Limewire hat die Fähigkeit, alle Suchanfragen, die durch den eigenen Knoten weitergeleitet wurden, zu speichern. Somit kann eine später gesendete Suchantwort mit der entsprechenden Anfrage in Verbindung gebracht werden. Der veränderte Client verbindet sich als Ultrapeer (Supernode) mit dem Netz und lädt jede Datei, die den Download-Kriterien entspricht –also eine von Viren befallbare ausführbare Datei sein könnte–, automatisch herunter. Die eingehenden Dateien werden dann mit dem ClamAV-Virenscanner auf Malware gescannt. Die Veränderungen am OpenFT-Client sind völlig analog, so dass hier nicht näher darauf eingegangen werden muss.

4 Durchgeführte Messungen

Im folgenden wird erläutert, wie die erstellten Messwerkzeuge benutzt wurden, um Daten über die Ausbreitung von Viren in Peer-to-Peer-Netzen zu gewinnen.

4.1 KaZaA

Das KaZaA-Netz wurde mit Hilfe von Krawler (siehe 3.1) nach den folgenden Begriffen durchsucht (dies waren die zum Zeitpunkt der Messung die zehn beliebtesten Programme auf www.download.com):

```
Ad, Spyware, LimeWire, ICQ, Registry, SpyBot, WinZip, Morpheus, All, iMesh,
IrfanView, WinRar, DivX, BitComet, RealPlayer, PC, Adobe, Trillian, Camfrog,
SmartFTP, Nero, MSN, Quick, Knight
```

Das in 3.1 beschriebene Programm Krawler lief auf drei Maschinen (2.1GHz Dualcore/ 1GB Ram, 2.1GHz/1.5GB Ram, 1.42GHz/1GB Ram). Auf diese Weise war es möglich, im Durchschnitt mehr als 60000 Dateien pro Stunde zu analysieren.
Die Analyse lief je drei Tage im Februar und März 2006.

4.2 Gnutella und OpenFT

Im Gegensatz zu KaZaA wurden Gnutella und OpenFT nur passiv überwacht. Das heisst, es wurden keine Suchanfragen an das Netz übermittelt. Vielmehr wurde der aufkommende Datenverkehr mit den in Absatz 3 beschriebenen Werkzeugen geloggt und nur bestimmte Dateien heruntergeladen: Ausführbare Dateien, Archive und Microsoft-Office-Dokumente (letztere wegen ihrer Anfälligkeit für Makroviren).

Der Messzeitraum bei Limewire betrug 45 Tage, bei OpenFT waren es 37 Tage.

5 Messergebnisse

Durch den oben beschriebenen Einsatz der Messwerkzeuge konnten mehrere Messungen vorgenommen werden, die im folgenden näher beschrieben werden.

5.1 KaZaA

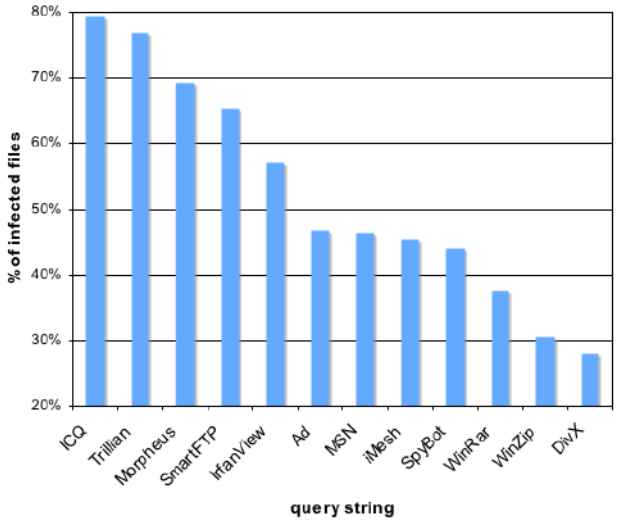
Im Fall von KaZaA wurden ja drei Tage im Februar und Mai 2006 gemessen. Die entstandenen Messwerte „feb-06" und „may-06" werden in Tabelle 1 in Bezug auf Anzahl der Antworten, Anzahl der antwortenden Supernodes und Anzahl der einzelnen Clients, die die betreffende Datei anbieten dargestellt. Dabei besteht die Möglichkeit von doppelten Antworten, da Krawler an mehrere Supernodes je gleiche Suchanfragen stellt.

Tabelle 1: Gesamtübersicht KaZaA

	feb-06	may-06
date	February 23, 2006	May 4, 2006
responses	654254	532610
supernodes	10267	15522
client hosts	19919	28601

5.1.1 Malware-Verteilung

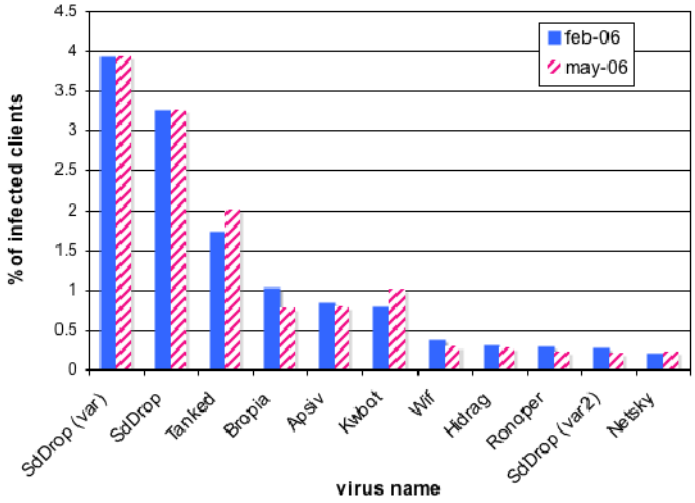
Abbildung 1: Prozentuale Verteilung infizierter Dateien



Wie Abbildung 1 zeigt, gibt es bestimmte Suchbegriffe, auf die besonders viele infizierte Dateien passen. Insgesamt waren im Februar 2006 22.9% der angeforderten Dateien infiziert; im Mai 2006 lag die Quote bei 15.2%.

In Abbildung 2 findet sich eine Häufigkeitsverteilung von Malware im FastTrack-Netz (natürlich nur unter den Dateien, die den gestellten Suchkriterien entsprachen). Dabei ist es interessant, zu beobachten, dass es keine größeren Veränderungen im Bezug

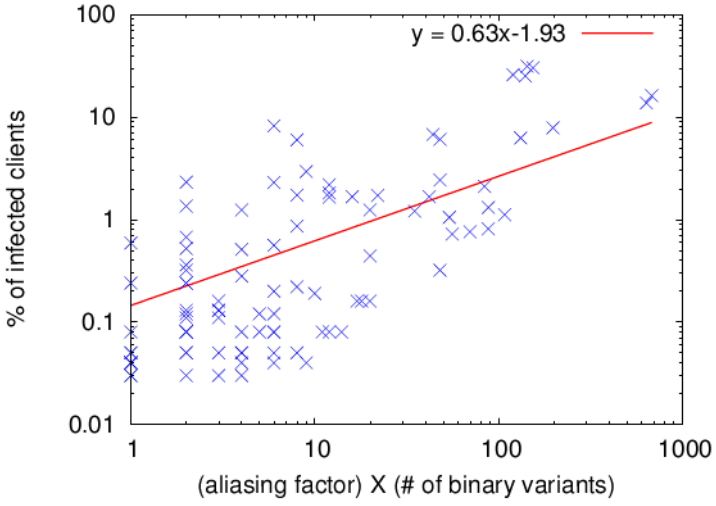
Abbildung 2: Vergleich der Infektionen im Februar und März 2006



auf die Verteilung der Viren zu geben scheint: Sowohl im Februar, als auch im Mai 2006 ist SdDrop der häufigste Virus, gefolgt von Tanked usw. Ausserdem ist zu beobachten, dass die drei am häufigsten anzutreffenden Viren speziell für den Einsatz in Peer-to-Peer-Netzen entworfen wurden.

Abbildung 3 zeigt, dass es eine Korrelation zwischen der Anzahl der möglichen Binärvarianten eines Virus und seiner Ausbreitung im Netz gibt: Viren, die ihren Code ändern bzw. in ihrer Größe variieren, verteilen sich statistisch gesehen besser im Netz, als solche, die dies nicht tun.

Abbildung 3: Korrelation zwischen möglichen Binärvarianten und Verbreitung der Viren



5.1.2 Genaue Art der Malware

Tabelle 2 gibt eine Übersicht über die genaue Art der beobachteten Malware. Es ist zu erkennen, dass eine große Zahl von Schädlingen eine Backdoor enthält. Solche Backdoors können unter anderem dem Aufbau von Botnetzen dienen. Auch sehr verbreitet ist Malware, die dem Versand von Emailspam dient. Versand von Spam über Instant-Messenger, DDoS und Datendiebstahl sind eher selten anzutreffende Funktionen. Das mag darin begründet sein, dass nach Installation einer Backdoor auf einem System solcherlei Schadprogramme in der Regel ohne weiteres nachinstalliert werden können. Da die Aktivitäten vieler Schädlinge schwer nachzuweisen sind (beispielsweise die Nutzung einer Backdoor oder der Diebstahl von Daten), war hier vorrangig die Malwarevon Interesse, die dem Spamversand diente. Zur Überprüfung, ob die infizierten Hosts am Versand von Spam beteiligt waren, wurden sie mit Hilfe der folgenden Black-Lists von

`bl.spamcop.net, cbl.abuseat.org, dnsbl.sorbs.net, list.dsbl.org, opm.blitzed.org, sbl.spamhaus.org`

überprüft. Tabelle 3 zeigt, dass in beiden Messungen über 70% der betreffenden Hosts am Versand von Spam beteiligt waren.

Tabelle 2: Art der Malware in KaZaA

Attack	Virus list
Backdoor	Sndc, Tanked, Kwbot, Bagle, Darby, SdBot, SpyBot, Swen, IRCBot, Agent.Gen, Delf, Dropper
Spam (email)	Bagle, Darby, Mapson-A, Ronoper, Swen, NetSky
Spam (messenger)	Bropia, Supova
DDoS	Darby, Kindal, SdBot
Information stealing	Darby, SdBot

Tabelle 3: Spamversand durch KaZaA-Malware

	feb-06	may-06
infected	1,618	2,576
listed in DNSBLs	1,146 (70.83%)	1,825 (70.85%)

5.2 Gnutella und OpenFT

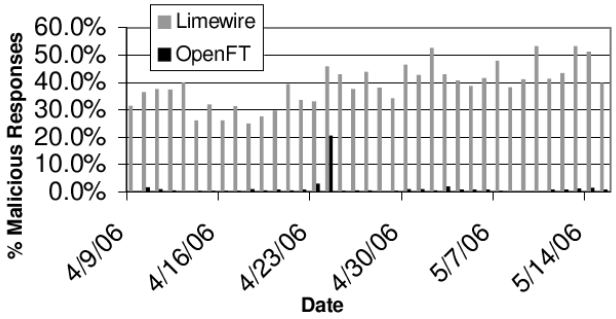
Tabelle 4 gibt eine Gesamtübersicht über die vorgenommenen Messungen im Gnutella und OpenFT-Netz. Man erkennt einen deutlichen Unterschied zwischen den beiden Netzen. Zuerst fällt ins Auge, dass die Zahl der Clients im OpenFT-Netz deutlich kleiner ist als im Gnutella-Netz. Eine überraschende Beobachtung ist, dass trotzdem das Verhältnis zwischen versuchten (also angeforderten, aber auf Grund von Nicht-Erreichbarkeit des Hosts oder Verbindungsabbruch nicht zu Ende geführten) und erfolgreichen Downloads im OpenFT-Netz mit ca. 77% deutlich höher liegt als im Gnutella-Netz (dort ca. 34%). Das deutet darauf hin, dass die Nutzer von OpenFT ihren Client weitaus seltener hinter einem NAT-Gerät betreiben (bzw. öfter ein Port-Forwarding eingerichtet haben) als die von Gnutella (wodurch es seltener zu Hosts mit nicht erreichbaren Adressen aus privaten Subnetzen kommt). Weiterhin ist zu beobachten, dass im OpenFT-Netz gefundene Dateien seltener den Downloadkriterien entsprechen (also von Windows-Viren

befallbar sind). Die spiegelt sich in der Zahl der „Qualifying responses" wider. Gleichzeitig erzielt eine Suche im OpenFT-Netz aber eine prozentual wesentlich höhere Anzahl ($\frac{\#queries}{\#responses} \approx 250\%$) von Ergebnissen als im Gnutella-Netz, wo $\frac{\#queries}{\#responses} \approx 94\%$ ist, was auf eine sehr große Verteilung der gesuchten Dateien im OpenFT-Netz hindeutet. Abbildung 4 gibt eine Übersicht über den Anteil von infizierten Dateien an den Gesamtergebnissen der jeweiligen Suchanfrage.

Tabelle 4: Gesamtübersicht über Messungen in Gnutella und OpenFT

	Limewire	OpenFT
Data collection days	45	37
Start date	1.4.06	9.4.06
Number of queries	34.268.803	12.347.509
Number of responses	32.788.921	30.538.152
Qualifying responses	2.468.327	381.851
Attempted downloads	228.722	22.231
Successful downloads	78.004	17.758
Unique clients	383.601	14.432

Abbildung 4: Anteil Infizierter Dateien in Gnutella und OpenFT



Der Anteil infizierter Binaries im Gnutella-Netz liegt meist über 30% (Abbildung 4), und stellt mit teilweise sogar über 50% einen signifikanten Anteil dar. Im OpenFT-Netz ist die Situation komplett anders: hier ist eine Quote von über 5% nur genau einmal zu beobachten, und zwar durch den Wurm „Poom.A". Dies mag an der geringen Verbreitung von OpenFT in der Windowswelt liegen; alle der hier untersuchten Schadprogramme sind Windows-spezifisch.

5.2.1 Häufigste Malware

Für das Gnutella-Netz sind die zehn häufigsten Schädlinge in Tabelle 5 dargestellt. Die zehn häufigsten Schädlinge im OpenFT-Netz finden sich in Tabelle 6. Der Wurm Poom.A ist im OpenFT-Netz deutlich stärker vertreten als bei Gnutella und steht auch mit 4512 Antworten an der Spitze der Malware-Liste im OpenFT-Netz. Betrachtet man Poom.A als „Ausreisser", so lässt sich feststellen, dass die Infektionsrate im OpenFT-Netz deutlich kleiner ist als im Gnutella-Netz. Allerdings rangieren Trojan.VB-100, Worm.Alcan.D und Worm.P2p.Poom.A in beiden Netzen auf den ersten Plätzen.

Tabelle 5: Die zehn häufigsten Schädlinge im Gnutella-Netz

Name	Files	Responses
Trojan.VB-100	19841	774216
Worm.Alcan.D	5978	140428
Worm.VB-16	334	5329
Worm.P2P.Poom.A	372	5120
Worm.SomeFool.P	83	2196
Trojan.Downloader.Istbar-176	331	818
Worm.VB-26	190	557
Trojan.JS.Startpage.C	212	447
Worm.Wupeer.A	159	182
Worm.P2P.Selmo.A	65	66

Tabelle 6: Die zehn häufigsten Schädlinge in OpenFT

Name	Files	Responses
Worm.P2P.Poom.A	101	4512
Trojan.VB-100	168	512
Worm.Alcan.D	71	395
RAR	71	361
Trojan.Downloader.Istbar-176	51	206
Worm.SomeFool.P	24	149
DOS.HLLC.Slam.6000	1	114
Worm.SomeFool.Gen-1	12	107
Trojan.Downloader.Istbar-172	11	50
Trojan.Downloader.Delf-286	3	47

5.2.2 Kritische Suchanfragen

Bestimmte Suchanfragen führten zu besonders vielen auf Malware zurückzuführenden Antworten. Tabelle 7 stellt die Situation im Gnutella-Netz dar, Tabelle 8 die Situation im OpenFT-Netz.
Dieses Verhalten ist von Interesse, da die Viren sich offenbar in der Benennung der von ihnen infizierten Dateien an momentanen Trends orientieren, um möglichst viele Benutzer dazu zu bringen, sie herunterzuladen. Bemerkenswert ist, dass im Gnutella-Netz viele Anfragen, die eigentlich Mediendateien als Ergebnis haben sollten (scary movie 4, ice age 2, lost, ...) offenbar infizierte Binärdateien als Ergebnis generieren. Die Situation im OpenFT-Netz ist grundlegend verschieden. Hier erhält man mit größter Wahrscheinlichkeit infizierte Dateien, wenn man nach ausführbaren Programmen sucht. Dieser Unterschied erklärt sich einerseits durch den geringen Anteil von Malware im OpenFT-Netz: es sind schlicht wenige Viren vorhanden, die sich in ihrer Namensgebung an aktuelle Trends anpassen können. Andererseits kann man auch über das Verhalten von Benutzern spekulieren, die erst gar keine befallbaren Dateien herunterladen, wenn sie auf der Suche nach Mediendateien waren (z.B. weil das von ihnen verwendete Betriebssystem keine Windows-Binaries ausführen kann).

Tabelle 7: Kritische Suchanfragen im Gnutella-Netz

Query	Files	Responses
scary movie 4	64	19003
ice age 2	145	17020
2006	1706	12008
lost	237	10550
silent hill	65	10117
ice age	145	9388
sex	248	7600
prison break	78	6704
hostel	41	5571
nero	180	5406

Tabelle 8: Kritische Suchanfragen im OpenFT-Netz

Query	Files	Responses
crack	94	872
adobe	16	422
sims	14	407
limewire	94	280
games	13	230
windows xp	7	222
macromedia	7	188
dreamweaver	3	165
zip	80	164
the sims	14	155

5.2.3 Host-Charakteristiken

Zur Analyse der Malware-Verteilung ist es wichtig festzustellen, in welchem Umfang die einzelnen Clients Malware verbreiten. Interessanterweise existieren sowohl im Gnutella, als auch im OpenFT-Netz Knoten, die besonders viel Schadsoftware verteilen. Die Tabellen 5.2.3 und 5.2.3 stellen die Verteilung im Gnutella und OpenFT-Netz dar.

Im folgenden werden an einigen Stellen die IP-Adressen von Clients im Netz verwendet. Dabei ist zu beachten, dass durch Verwendung von Network Address Translation (NAT) die Möglichkeit besteht, dass eine solche IP-Adresse aus einem privaten Subnetz stammt. Dadurch kann es passieren, dass mit einem Client hinter einem NAT-Gerät keine Verbindung aufgebaut werden kann, wenn dieser kein Forwarding der für das Netz spezifischen Ports eingerichtet hat. Es lässt sich beobachten, dass im Gnutella-Netz eine große Zahl von Malware von Nodes mit einer Privaten Adresse angeboten wird, und infizierte Hosts meistens mehr als nur einen Virus anbieten. Die Situation im OpenFT-Netz ist grundlegend verschieden: hier wird die Malware fast nie von Nodes mit Adressen aus privaten Netzen angeboten, und auch die Anzahl der verschiedenen Schädlinge liegt hier bei höchstens zwei.

Tabelle 9: Die zehn häufigsten Virenverteiler im Gnutella-Netz

IP	Infected Files	Infected Responses	Malware	Clean Files	Clean Responses
192.168.1.11	3982	72418	4	231	419
65.34.187.196	4782	18257	2	27	63
192.168.1.100	7593	14664	10	3827	7787
192.168.1.2	5624	9393	12	4046	9551
192.168.1.101	4736	9219	8	1789	3543
192.168.1.47	1593	7851	2	1778	3275
192.168.1.3	4710	7418	9	1582	2449
192.168.0.10	3471	6421	2	434	631
85.167.159.53	3248	5823	1	4	7

Tabelle 10: Die zehn häufigsten Virenverteiler im OpenFT-Netz

IP	Infected Files	Infected Responses	Malware	Clean Files	Clean Responses
24.185.43.12	101	4512	1	55	185
200.193.133.181	12	91	1	0	0
68.199.111.60	24	83	1	3	9
24.108.148.47	1	63	1	8	342
67.70.44.58	1	50	1	11	186
24.73.2.236	23	44	2	99	124
200.63.211.100	3	42	2	2	10
69.157.73.229	1	33	1	23	486
67.8.149.155	2	32	2	11	210
69.253.47.181	1	30	1	0	0

6 Impact und mögliche Gegenmaßnahmen

Der tatsächlichen Auswirkungen der zunehmenden Virenverteilung sind schwer einzuschätzen: zwar scheint eine relativ große Zahl von Hosts infizierte Dateien anzubieten, jedoch muss der Nutzer diese auch herunterladen und ausführen, um sich einer konkreten Gefahr durch Malware auszusetzen. Benutzer, die nur Mediendateien wie Musik oder Filme tauschen, bleiben von der hier analysierten Malware verschont.

Weiterhin ist bei der Virenverteilung im FastTrack-Netz zu beachten, dass die gestellten Suchanfragen das Ergebnis maßgeblich beeinflussen. Um eine wirklich exakte Messung vornehmen zu können, sollten Dateien angefragt werden, die auch im Durchschnitt oft von den Benutzern des Netzes gesucht werden.
Da die Suchbegriffe jedoch von www.download.com stammen, kann nicht direkt sichergestellt werden, dass die gesuchten Dateien auch im KaZaA-Netz von besonderer Beliebtheit sind; schließlich könnten viele Benutzer von KaZaA das Netz nur zum Download von Dateien nutzen, die sie über direkte HTTP oder FTP-Downloads nicht finden.

Dennoch ist es natürlich wünschenswert, der Ausbreitung von Malware auch in Tauschbörsen entgegenzuwirken.

In [KAM 06] wird vorgeschlagen, Malware anhand der Dateigröße zu erkennen und entsprechende Suchanfragen zu filtern. Dieses Vorgehen scheint auch für den momentanen Bestand an Viren seinen Zweck zu erfüllen, ist aber recht naiv: die Autoren der speziell für Peer-to-Peer-Netze ausgelegten Viren könnten sehr einfach ihre Malware so

gestalten, dass sie alle möglichen (realistischen) Dateigrößen annehmen kann. Dadurch würde ein solches Verfahren schnell ausgehebelt werden, da bei Aufnahme aller Dateigrößen in die Liste der verdächtigen Dateien zu viele False-Positives entstünden.

Ein weiterer einfacher Ansatz wäre, jeden Client mit einem Virenscanner auszustatten. Dies ist zwar ein technisch recht leicht zu realisierendes Mittel, jedoch könnte der Benutzer von Peer-to-Peer-Software auch selbst dafür sorgen, indem er einen aktuellen Virenscanner installiert, der automatisch alle Dateien, die auf der Festplatte des Computers gespeichert werden, durchsucht.

Ein etwas komplizierterer Ansatz ist ein verteiltes (evtl. vertrauensbasiertes) Modell zur Bewertung der Gefährlichkeit von Dateien im Netzwerk. Dies ist jedoch algorithmisch und technisch relativ komplex, da viele mögliche Schwierigkeiten bestehen:

- „Böswillige" Clients können behaupten, infizierte Dateien seien nicht infiziert. Andererseits können Denial-Of-Service-Attacken durch Clients, die alle Dateien als infiziert kennzeichnen, auftreten.
- Es müssen gemeinsame Kriterien zur Bewertung der Gefährlichkeit von Dateien gefunden werden
- Es muss einen verteilten Algorithmus zum Aktualisieren der Virensignaturen geben. Das bringt das Problem mit sich, dass Schadsoftware veränderte Virensignaturen ins Netz einschleusen könnte.

Alles in allem scheint ein solcher Ansatz schwierig, jedoch nicht unmöglich. Ein vertrauensbasiertes System (wie z.B. in [KaSG 03] beschrieben) könnte bei einigen Problemen Abhilfe schaffen.

Der Vorteil, den ein solches verteiltes System zur Bewertung mit sich bringen würde ist, dass man es erweitern könnte, um auch andere Aspekte in einem Netz zu beurteilen („Qualität" von Dateien, Up-/Download-Verhältnis von Nutzern etc.). Insbesondere in sozialen Netzen, die nicht nur das Tauschen von Dateien gestatten, sondern Interaktionen zwischen den Nutzern ermöglichen (wie z.B. Auktionsplattformen), könnte ein verteiltes Bewertungs- und Abstimmungssystem von großem Nutzen sein.

7 Zusammenfassung

Malware in Peer-to-Peer-Netzen verbreitet sich mehr und mehr. Der angerichtete Gesamtschaden lässt sich zwar momentan nur schwer beziffern, aber zumindest für die betroffenen Netze selbst scheint Malware ein nicht unerhebliches Problem zu sein. Die Benutzer der Netze laden mit einer Wahrscheinlichkeit von ca. 30% (im Fall von OpenFT nur um höchstens 5%) infizierte Dateien herunter (ausgenommen natürlich, sie laden nicht infizierbare Inhalte wie Filme oder Musik herunter).

Weiterhin zeigt die Untersuchung, dass die über Peer-to-Peer-Netze verbreitete Malware oft zum Spamversand dient. Spam ist heutzutage ein ernstzunehmendes Problem.

Gegenmaßnahmen obliegen bisher nur dem Nutzer selbst. Die gängigen Tauschbörsen bieten keine effektive Möglichkeit, sich vor Malware zu schützen. Interessanterweise scheint es jedoch hinreichend viele Nutzer zu geben, die keinen (oder einen schlechten) Virenscanner einsetzen und so zur Verbreitung von Malware in Peer-to-Peer-Netzen beitragen.

Literatur

[KAM 06] Andrew Kalafut, Abhinav Acharya, Minaxi Gupta. A Study of Malware in Peer-to-Peer Networks. IMC'06, October 25-27.

[SJB 06] Seungwon Shin, Jaeyeon Jung, Hari Balakrishnan. Malware Prevalence in the KaZaA File-Sharing Network. IMC'06, October 25-27.

[KaSG 03] Sepandar D. Kamvar, Mario T. Schlosser, Hector Garcia-Molina. The Eigen-Trust Algorithm for Reputation Management in P2P Networks. WWW2003, May 20–24, 2003.