
Blacklist-Nutzung zeigt Internet-Missbrauch in Echtzeit

# Ausgefragt

**Sebastian Ganschow, Christian J. Dietrich, Norbert Pohlmann**

Anti-Spam-Blacklists erhalten zum Teil viele Tausend DNS-Abfragen pro Sekunde, die sich direkt auf den weltweiten E-Mail-Verkehr beziehen. Betreiber von DNSBL-Servern haben damit einen guten Blick auf das Geschehen – müssen dafür aber regelrechte Datengebirge bezwingen.

IP-Blacklisting ist ein wichtiger Schutzmechanismus im Kampf gegen Internet-Missbrauch, vor allem Spam. Eine Reihe von Blacklists lässt sich kostenlos über das Internet nutzen. Üblicherweise kommt das DNS-Protokoll für deren Abfrage zum Einsatz (DNSBL). Außer dem eigentlichen Inhalt der Blacklists, also den Spam-Quellen in Form von IP-Adressen, sind auch die Abfragen ein interessantes Analyseobjekt – sowohl die abfragenden Hosts als auch die abgefragten Adressen betreffend.

Wer Zugriff auf die Abfragedaten hat, kann unter anderem den Anteil des Inhalts einer Blacklist bestimmen, der überhaupt abgefragt wird. Darüber hinaus kann man die Menge an positiven Antworten, also Treffern auf der Blacklist, ins Verhältnis setzen zu allen Anfragen („Trefferquote“ oder „Hit Rate“). Ferner lässt sich bestimmen, wie sich etwa das Spam-Aufkommen entwickelt oder welche Nutzerzahl die Blacklist hat – allesamt wichtige Eckdaten für Wirksamkeitsbetrachtungen und Optimierungen.

In einem Zeitraum von fast einem halben Jahr wurden die Anfragen an zwei öffentliche Blacklists untersucht, darunter die vom Projekt „NiX Spam“ der *iX*. Sie umfasste zu Beginn der Untersuchung im Juli 2007 etwa 78 000 IP-Adressen aus 166 Ländern, zum Ende (Januar 2008) bereits über 400 000 Adressen. Zum Vergleich diente die „Blackholes“-Liste (www.five-ten-sg.com). Für beide betreiben die Autoren DNS-Slave-Server, sodass sie einen repräsentativen Ausschnitt aus den Anfragen an die Blacklists „mitschneiden“ können.

## Sklavenbetreiber lesen mit

Die NiX-Spam-Liste verfügte am Ende des Erfassungszeitraums über zehn DNS-Server, von denen zwei während der Messungen hinzukamen. Mit jedem hinzukommenden Server sollte die Anzahl der Anfragen pro Server zumindest vorübergehend abnehmen. Die Anzahl der Anfragen stieg jedoch praktisch über den gesamten Messzeitraum hinweg unbeirrt: Zu Beginn lag die Zahl der Requests pro Tag bei rund 5,5 Mio. Zuletzt liefen täglich etwa 9,5 Mio. Anfragen auf. Die Trendlinie deutet an, dass von einem weiteren Anstieg auszugehen ist.

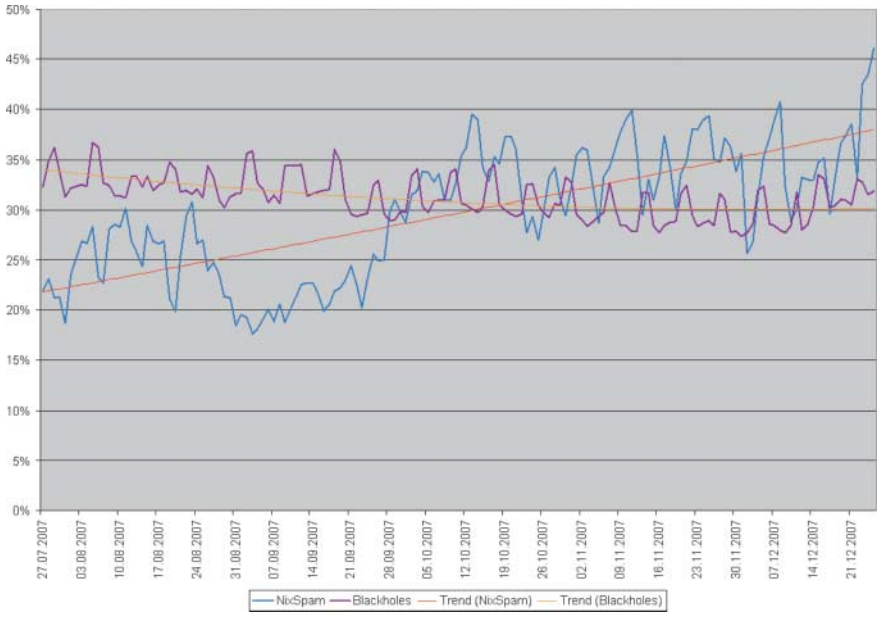
Während die abgefragten IP-Adressen zeigen, wo Spam und erwünschte E-Mails herkommen, enthalten die Adressen der Abfragenden selbst Hinweise darauf, wer die Nutzer der Blacklist sind. Bei NiX Spam sind es pro Tag etwa 10 000 verschiedene Adressen und damit schätzungsweise mindestens ebenso viele Mailserver-Betreiber. Da die vorliegende Messung lediglich an einem Server stattfand, blieben eventuelle weitere Anwender der anderen Server unberücksichtigt. Da die DNS-Anfragen jedoch per Round Robin praktisch gleichmäßig an alle Server gehen, dürfte das Ergebnis dennoch realistisch sein.

DNS-Resolver können ferner Antworten zwischenspeichern, sodass vermutlich nicht alle Anfragen tatsächlich bis zum Server der Blacklist durchdringen. Unter Vernachlässigung des Resolver-Caching entspricht jede Anfrage an die Blacklist einer SMTP-Session (zumeist mit einer empfangenen oder abgewiesenen E-Mail). Interessanterweise blieb die Anzahl der anfragenden IP-Adressen über den gesamten Messzeitraum hinweg relativ konstant, während die Zahl der Anfragen deutlich wuchs. Eine mögliche Erklärung dafür ist ein Anstieg des gesamten E-Mail-Volumens.

Befindet sich eine abgefragte IP-Adresse auf der Blacklist, gibt der DNSBL-Server eine positive Antwort. Die Zahl verschiedener IP-Adressen

gibt Aufschluss über die Menge mailsendender Hosts. Da Spam über 90 Prozent aller Mails ausmacht, tauchen vor allem solche IP-Adressen auf, die den Spam-Versendern zur Verfügung stehen. Die Abfragen an den NiX-Spam-Slave beziehen sich durchschnittlich auf 1,2 Mio. Adressen täglich. Da manche davon mehrmals auftauchen, sind es insgesamt deutlich mehr Abfragen.

Am 24. Dezember 2007 etwa fragten NiX-Spam-Anwender exakt 1 158 148 verschiedene IP-Adressen ab. Davon standen nur rund 10 % (112 980) auf der Blacklist. Die Liste umfasste damals rund 300 000 IP-Adressen. Nur gut ein Drittel der gelisteten IP-Adressen wurde also an diesem Tag überhaupt abgefragt. Einige Monate zuvor, am 1. August 2007, wurden „nur“ 781 472 verschiedene IP-Adressen im Laufe des Tages abgefragt. Die Zahl verseuchter, für den Spam-Versand missbrauchter Systeme steigt anscheinend nach wie vor.




**Die NiX-Spam-Liste nimmt laufend neue Spam-Fallen in Betrieb und steigert dadurch den Anteil positiver Antworten (Abb. 1).**

## Blackholes als Vergleichsbasis

Die zum Vergleich herangezogene Blackholes-Liste verfügte im Untersuchungszeitraum konstant über 14 DNS-Server. Anders als bei „NiX Spam“ zeigt der Trend der Anfragen abwärts. Zugleich hat sich die Zahl der anfragenden IP-Adressen erhöht. Zwischen der Anzahl der abgefragten IP-Adressen und der positiven Antworten gibt es – unerklärlicherweise – kaum Abweichungen.

Bei Blackholes liegt die Zahl der abgefragten unterschiedlichen IP-Adressen pro Tag im Durchschnitt bei rund 1 Mio. Am 24. Dezember 2007 standen davon etwa 320 000 auf der Liste. Dies entspricht rund 32 %, mehr als dreimal so viel wie bei NiX Spam. Über den gesamten Dezember 2007 hinweg belief sich die Anzahl an verschiedenen IP-Adressen auf 4 744 440, während der gesamten fünf Monate zeigten sich sogar 17 742 550 – offenbar vor allem aus Bereichen dynamisch immer wieder neu vergebener Adressen.

Anfang August belief sich die Anzahl an verschiedenen, abgefragten IP-Adressen der Blackholes auf 1 006 668. Hiervon standen etwa 228 000 auf der Blacklist, die sich damit potenziell spammenden Hosts zuordnen ließen.

## Trefferquote ist nicht alles

Ein wichtiger Indikator dafür, ob sich das Abfragen einer Blacklist überhaupt lohnt, ist der Anteil der positiv beantworteten Anfragen. Je höher die Trefferwahrscheinlichkeit, desto mehr Spam-Nachrichten könnten mithilfe der Blacklist erkannt werden und desto effektiver ist die Liste. Andererseits ist die „Hit Rate“ in Verbindung mit der Gesamtanzahl an Anfragen auch ein Indiz für die Menge an Spam, die versendet wird.

Der Trefferanteil bei NiX Spam stieg über den gesamten Zeitraum der Untersuchung hinweg deutlich (Abb. 1). Er hat sich während des Zeitraums fast verdoppelt und lag zuletzt bei weit über 40 %. Eine hohe Hit Rate bedeutet in erster Linie, dass viele abgefragte Adressen tatsächlich auf der Liste stehen. Dies geht allerdings nur so lange gut, wie die Zahl der fehlerhaft gelisteten IP-Adressen (False Positives) sehr gering bleibt. Für die Bewertung der Qualität einer Blacklist muss daher

### iX-TRACT

- Die bei Anti-Spam-Blacklists abgefragten IP-Adressen sind mindestens ebenso interessant wie der eigentliche Datenbestand, da sie ein realistisches Bild vom laufenden E-Mail-Geschehen vermitteln.
- Die Abfragen bei Blacklists geben sowohl die räumliche Verteilung von Spam-Quellen als auch den zeitlichen Ablauf von Spam-Angriffen wieder.
- Spam-Versender nutzen laufend über eine Million IP-Adressen, meist über verseuchte PCs. Sie verwenden einen Großteil davon nur für äußerst kurze Zeit.

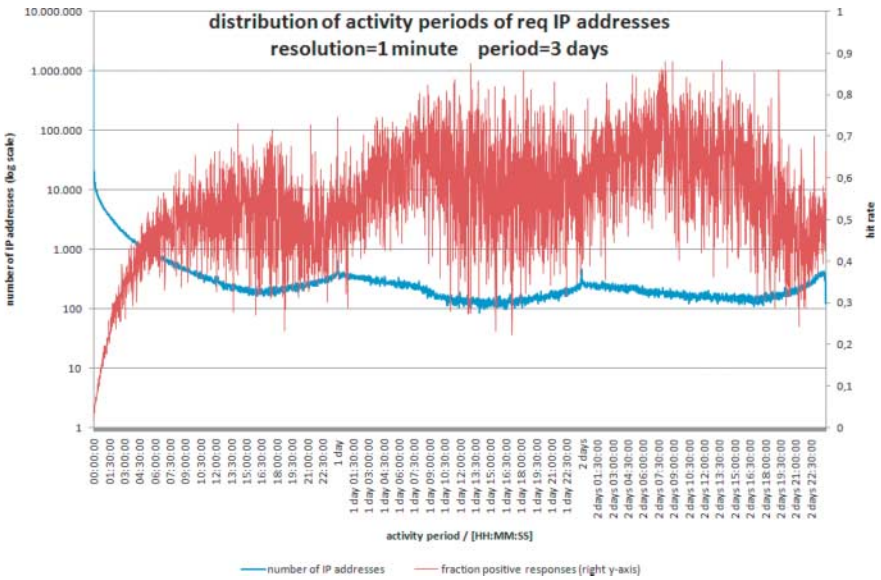
neben der Trefferquote immer auch der False-Positive-Anteil berücksichtigt werden. Eine automatische Messung der False-Positive-Rate gestaltet sich jedoch mindestens so schwierig wie die Definition von „False Positive“ an sich und ist derzeit Gegenstand der Forschung.

Das Sinken der Zugriffszahlen auf die Blackholes könnte auf deren relativ hohe False-Positive-Rate zurückzuführen sein. Al Iverson hat etwa zum gleichen Zeitpunkt, zu dem unsere Messungen begannen, angefangen, verschiedene Blacklists zu analysieren und die Blackholes schlecht bewertet (www.dnsbl.com).

Bei einem Spam-Anteil von zurzeit über 90 % treffen sämtliche Statistiken zwangsläufig vor allem Aussagen über unerwünschte E-Mails – so auch diejenige über Quellregionen (siehe Tabellen). Bei beiden Blacklists belegen die USA und Russland die ersten beiden Plätze. Neun Länder finden sich in den „Top 10“ beider Blacklists. Ein Drittel der Anfragen für IP-Adressen aus den USA beantwortet NiX Spam positiv. Bei den IP-Adressen aus Russland ist es sogar die Hälfte der Anfragen. Der Anteil für aus Deutschland empfangene Spam-Mails liegt nur bei 10 %. Daraus lässt sich schließen, dass die Nutzer dieser Blacklist größtenteils Spam aus den USA, Russland und Korea empfangen.

Der Anteil positiv beantworteter Anfragen liegt bei Blackholes wesentlich niedriger. Auf Anfragen für IP-Adressen aus den USA liefert sie nur in rund einem Viertel der Fälle eine positive Antwort. Immerhin ein Drittel der Anfragen zu deutschen IP-Adressen beantwortet sie jedoch positiv (NiX Spam: 10 %).

E-Mails aus China sind offenbar meist Spam. Das Ergebnis deckt sich mit den Aussagen von Anti-Spam-Dienstleistern. In absoluten Zahlen werden die AS von den türkischen, italienischen und polnischen Telecoms (TTnet, Telecom Italia und TP) hier am häufigsten abgefragt.




**Drei Tage im Überblick: Die Trefferquote (rot) steigt mit der Lebensdauer eines Eintrags, doch die meisten Einträge sind nur für extrem kurze Zeit aktiv (Abb. 2).**

Unter den bei Blackholes abgefragten Adressen fallen die AS von Arcor und Korea Telecom aus dem Rahmen, denn der Anteil an positiven Antworten ist dort mit jeweils fast 99 % sehr hoch. Fast jede IP-Adresse, zu der die Blackholes-Liste konsultiert wurde, stand auch tatsächlich auf der Blacklist. Man könnte aufgrund dieses Ergebnisses auf gewisse Abneigungen des Blacklist-Betreibers schließen. Die zweite Reihe bilden die beiden autonomen Systeme von Chinanet (AS 4837 und 4134). Der relative Anteil an positiv beantworteten Anfragen liegt in beiden Fällen bei rund 87 %. Sie verursachen ein deutlich höheres absolutes Aufkommen an Anfragen als Arcor und Korea Telecom.

Auf der NiX-Spam-Liste spielt Arcor dagegen keine große Rolle. Die abgefragten Adressen werden dort derzeit dominiert von der Koreanischen Telecom (11 Mio. Abfragen, 80 % davon positiv beantwortet), TTnet (10 Mio./44 %) und Telefonica del Peru (5 Mio./51 %). Die Deutsche Telekom taucht, was die Abfragen angeht, mit 3,2 Mio. pro Tag erst auf Platz 10 auf, und davon werden nur 730 000 (23 %) positiv beantwortet, ein besonders geringer Wert.

## Sehr geringe Lebensdauer

Bei NiX Spam gelistete IP-Adressen fallen nach drei Tagen wieder aus der Blacklist, wenn von ihnen nicht erneut Spam ausgeht. Bisher basierte dieser Wert allein auf Experimenten. Eine Statistik über die Verteilung der Zeiträume, in denen IP-Adressen abgefragt werden, zeigt, dass tatsächlich nur rund 8 % der abgefragten IP-Adressen überhaupt über einen Zeitraum von mehr als drei Tagen auftauchen – sprich: E-Mails (meist Spam) versenden. Anders ausgedrückt: 92 % aller abgefragten IP-Adressen werden in einem Zeitraum von drei Tagen abgefragt. Drei von vier abgefragten IP-Adressen scheinen sogar nur maximal einen Tag lang aktiv zu sein.

Noch deutlicher zeigt sich dieses Phänomen hinsichtlich der Aktivitätszeiträume innerhalb eines Tages (Abb. 2). Die x-Achse stellt den Aktivitätszeitraum, die y-Achse die Anzahl abgefragter IP-Adressen in logarithmischer Skalierung dar. Wird eine IP-Adresse beispielsweise je einmal um 12:00 und um 13:00 Uhr (und dann im Messzeitraum nicht mehr) abgefragt, gilt sie als eine Stunde lang aktiv.

**NiX-Spam-Abfragen nach Ländern**

| Land | Anfragen | positive Antworten |
|---|---|---|
| USA | 36 687 286 | 11 750 439 |
| Russland | 15 070 834 | 8 406 100 |
| Deutschland | 14 957 047 | 1 494 028 |
| Südkorea | 14 067 659 | 8 604 614 |
| Brasilien | 8 592 407 | 2 611 582 |
| Großbritannien | 8 200 419 | 2 454 146 |
| Polen | 8 049 853 | 2 796 431 |
| Türkei | 7 747 089 | 2 016 771 |
| Spanien | 6 848 259 | 2 371 162 |
| Italien | 6 695 144 | 1 682 333 |

**Top 10 der bei der NiX-Spam-Blacklist abgefragten IP-Adressen nach Ländern**

**Blackholes-Abfragen nach Ländern**

| Land | Anfragen | positive Antworten |
|---|---|---|
| USA | 19 700 000 | 4 760 000 |
| Russland | 6 325 000 | 1 122 000 |
| Türkei | 4 441 000 | 1 573 000 |
| Deutschland | 4 035 000 | 1 368 000 |
| Brasilien | 3 852 000 | 1 194 000 |
| Italien | 3 778 000 | 1 135 000 |
| Großbritannien | 3 253 000 | 1 289 000 |
| Polen | 3 226 000 | 1 542 000 |
| China | 3 109 000 | 2 729 000 |
| Spanien | 2 774 000 | 915 000 |

**Top 10 der abgefragten IP-Adressen bei Blackholes nach Ländern**

Die Darstellung verdeutlicht, dass ein großer Teil der Anfragen eine sehr kurze Aktivität aufweist. Von den pro Tag rund 1,35 Mio. verschiedenen, abgefragten IP-Adressen sind rund 496 000 IP-Adressen (entsprechen 37 %) weniger als eine Minute lang aktiv. In fast 93 % der Fälle tauchen diese Adressen auch tatsächlich nur ein einziges Mal innerhalb eines Tages auf. Weitere 6 % der „kurzlebigen" Adressen tauchen innerhalb der ersten Minute zweimal auf. Ein lokales Maximum befindet sich zwischen 23 und 24 Stunden, offenbar durch dauerhaft betriebene Mailserver verursacht. Während eines längeren Zeitraums, etwa über drei Tage, zeigt sich ein weiteres Muster.

Die Erkennungsrate fällt erwartungsgemäß bei sehr kurzlebigen IP-Adressen entsprechend gering aus. Da eine IP-Adresse zunächst als Spam-Quelle auffallen und in die Blacklist gelangen muss, haben kurzlebige oder einmal verwendete IP-Adressen gute Chancen, nicht erkannt zu werden. Für Betreiber von Blacklists bedeutet das Ergebnis jedoch die Herausforderung, die Trägheit der Systeme zu minimieren. Keine leichte Aufgabe, wenn man nicht riskieren möchte, zu viele Adressen fälschlicherweise zu blockieren.

Ein ebenfalls zu beobachtendes, noch unklares Phänomen sind „selbstabfragende" IP-Adressen: Gelegentlich entspricht die abgefragte IP-Adresse der abfragenden. Das fand während des Messzeitraums über 150 000-mal statt, etwas zu häufig etwa für gelegentliche Experimente von Endanwendern. Eventuell gehen solche selbstbezüglichen Abfragen von Bots aus, die prüfen, ob ihr Wirt bereits unangenehm aufgefallen ist – oder von unvorsichtig konfigurierten Mailservern.

## Fazit und Ausblick

Die Verhaltensanalyse von Blacklists kann Licht in das Dunkel bringen, wenn es um ein Gefühl für die Nutzer einer Blacklist sowie Eigenschaften der gelisteten Adressen geht. Langfristig lassen sich möglicherweise aus dem Abfrageverhalten einer Blacklist Optimierungen wie zur Lebensdauer einer IP-Adresse auf der Blacklist ableiten. Derzeit stellt darüber hinaus die automatische Messung des False-Positive-Anteils einer Blacklist noch eine Forschungsaufgabe dar. (un)


SEBASTIAN GANSCHOW

arbeitet als Senior Network Engineer bei der Dr. Bülow & Masiak GmbH in Marl. Er unterstützt das Institut für Internet-Sicherheit an der FH Gelsenkirchen im Forschungsbereich E-Mail-Sicherheit.

CHRISTIAN DIETRICH

ist Mitarbeiter des Instituts für Internet-Sicherheit an der FH Gelsenkirchen und für den Forschungsbereich E-Mail-Sicherheit verantwortlich.

PROF. DR. NORBERT POHLMANN

ist Informatikprofessor für Verteilte Systeme und Informationssicherheit sowie Leiter des Instituts für Internet-Sicherheit an der FH Gelsenkirchen.