HSR
HOCHSCHULE FÜR TECHNIK
RAPPERSWIL

Kurzfassung der Studienarbeit

Abteilung	**Informatik**
Namen der Studenten	**Adrian Ruoss** **Christian Höhn**
Semester	**2004/05**
Titel der Studienarbeit	**SPAM DCC and Honeypot**
Examinator	**Prof. Dr. Peter Heinzmann**

Im Rahmen dieser Studienarbeit galt es, als Teil der Anti-SPAM Lösung für die SWITCH Community, ein Distributed Checksum Clearinghouse (DCC) System zu evaluieren, zu realisieren und in einen Pilotbetrieb überzuführen. Ferner sollten anhand von E-Mail Experimenten mit Decoy/Honeypot-Adressen und durch die Analyse von Adressen auf einer SPAM-Adessen-CD Hinweise zur Arbeitsweise von SPAM-Organisationen gewonnen werden.

Nach eingehendem Studium des DCC Prinzips wurden bei SWITCH zwei private DCC-Server aufgesetzt und in Betrieb genommen. Dieses Pilot-DCC-System wird nun von verschiedenen Universitäten und von der ETHZ genutzt. Messungen belegen, dass DCC 30% bis 65% aller SPAM-Mails detektiert.
Um die Qualität der Anti-SPAM-Massnahmen bewerten zu können, sind klar definierte Messpunkte und Paramter nötig. Im Rahmen der vorliegenden Arbeit wurden solche defniert. Die SpamAssassin Log Files der Pilotkunden zeigen beispielsweise, dass mit DCC die SPAM-Erkennungsrate um 1% bis 5% besser wird als ohne DCC. Die Rate hängt jedoch nicht davon ab, ob öffentliche oder lokale DCC-Server verwendet werden. Der Einsatz privater DCC-Server verbessert aber die Antwortzeiten. Mit durchschnittlich 49ms erreicht das Pilot-DCC-System eine um rund 80% schnellere Antwortzeit als öffentliche DCC-Server, was die Bearbeitungszeit für eine E-Mail merklich verkürzt.
Anhand des Decoy / Honeypot E-Mail Experimentes konnte aufgezeigt werden, dass E-Mail Adressen, welche auf statischen Webseiten und Foren / Newsgorups platziert sind, innerhalb einer Woche die ersten SPAM E-Mails erhalten, worauf pro Woche ca. 5 bis 10 SPAM E-Mails an diese Adressen geschickt werden.
Die Analyse der Adressen auf einer über EBay gekauften CDROM mit SPAM-Adressen hat gezeigt, dass von den angepriesen 200 Mio.Adressen, nach Abzug von mehrfach aufgeführten Adressen und nach Elimination von ungültigen Adressen (durch Überprüfung der Rejects) nur etwa 4.5% aller angepriesenen Adressen auch als SPAM-Adressen genutzt werden können.
Nach Abschluss dieser Studienarbeit, wird SWITCH voraussichtlich die DCC-Server aus dem Testbetrieb in den produktiven Betrieb überführen.