Forschungsprojekt
Frühwarnsystem für Bot-Netze

Abschlußbericht

Dr. Alexander K. Seewald, Doz. Dr. Wilfried N. Gansterer
23.1.2009

1 Hintergrund und Zielsetzung

Große Netze gekaperter Computer unbeteiligter und unschuldiger Nutzer des Internet, sogenannte „Bot-Netze“, sind mittlerweile zur Hauptquelle für diverse negative Entwicklungen im Bereich des Internet (Viren, malware, spam, phishing, etc.) geworden. Das Ziel dieses Projektes war es, ein Frühwarnsystem für Bot-Netze zur Sicherung des österreichischen Internet zu konzipieren, zu entwickeln, und der Öffentlichkeit verfügbar zu machen.

Die Verbreitung der Malware zum Aufbau von Bot-Netzen und zur Infizierung neuer Computer kann beispielsweise über *unsolicited bulk e-mail* (im Weiteren vereinfachend mit dem Begriff *Spam* zusammengefasst) erfolgen. Bisherige Forschungsinitiativen im Bereich der Anti-Spam-Forschung konzentrieren sich hauptsächlich auf die Erkennung und Abwehr dieser unerwünschten oder potentiell schädlichen E-Mail Nachrichten. In diesem Projekt wurde als wichtige Ergänzung dazu die (vorbeugende) Identifizierung und Früherkennung der derzeitigen Hauptquelle von Spam eingehend untersucht.

2 Aktivitäten

Es wurde eine Webpage eingerichtet, auf der der Projektfortschritt laufend überblicksmäßig dokumentiert wurde.

Nach einer umfassenden Recherche des aktuellen Standes der Forschung in diesem Bereich war eine erste zentrale Komponente des Projektes der Aufbau eines „Darknets“ und einer zugehörigen Serverinfrastruktur in Kooperation mit dem ZID der Universität Wien. Dieses Darknet dient vor allem der Sammlung von authentischen Daten (Zugriffe, Verbindungen, Netzwerkaktivität) über aktive Bot-Netze.

Nach dem erfolgreichen Aufbau des Darknets wurde das Hauptaugenmerk auf die Konzeption und Entwicklung des Bot-Netz-Warnsystems gelegt. Die zugrunde liegende „Intelligenz“ unseres Systems basiert auf der Analyse der aus unserem Darknet gewonnenen Daten. Nach der Analyse der Daten wurden entsprechende Modelle für die Erkennung von Spam-Bot-Netzen entwickelt und validiert.

Wie im Deliverable 2.4 des Projektplanes vorgesehen, haben wir eine installierbare Gesamtversion des Bot-Netz-Warnsystems zusammengestellt (GPL v3, Debian Linux (Etch)) und zum Download auf http://botnetz-tracker.seewald.at zur Verfügung gestellt (siehe den Link in der Mitte der Seite). Die zugehörige Installationsanleitung ist in einer README-Datei enthalten. Zusätzlich steht auf

http://botnetz-tracker.seewald.at eine Live-Version des Systems zur Verfügung, die zwei Funktionen bietet: Einerseits kann online getestet werden, ob der Rechner, von dem aus die Seite aufgerufen wird, momentan Teil eines der (dzt. bekannten) Bot-Netze ist, andererseits findet sich eine online-Visualisierung der weltweiten Bot-Netz-Aktivität der letzten 24 Stunden.

Einer der schwierigsten Aspekte des in diesem Projekt adressierten Problems ist der Zugang zu verifizierten Referenzdaten aus dem Netzwerkverkehr, die für die Generierung und Validierung der erzeugten Klassifikationsmodelle erforderlich sind. Auch in diesem Projekt konnte diese Schwierigkeit nicht völlig überwunden werden, weil mehrere diesbezügliche Versuche, über Kontakte zu einschlägigen Unternehmen und Organisationen, die über entsprechende Referenzdaten verfügen, Zugriff auf solche Daten zu erhalten, leider nur eingeschränkt erfolgreich waren.

Aufgrund von in manchen Bereichen nur unzureichend vorhandenen Referenzdaten konnten wir bis dato nur einen Teil der ursprünglich geplanten Funktionalität realisieren. Die zuvor erwähnte installierbare Version des Bot-Netz-Warnsystems erlaubt den Aufbau einer Blacklist mit allen bis jetzt getrackten Bots. An den Aktivitäten eines Bot-Netzes kann unser System den Zugriff selbst erkennen und auch das zugehörige Spam-Bot-Netz identifizieren, weitere Informationen liegen derzeit aber nicht vor. Das System macht auch keinen Unterschied zwischen Aktivitäten außerhalb und innerhalb von Bot-Netzen. Dies liegt darin begründet, dass die Aktvität von Botnetzen bei weitem überwiegt. Obwohl wir sehr wohl vorläufige Resultate hinsichtlich der Zugriffsmuster auf unser Darknet haben, sind diese noch zu unklar bzw. unzureichend validiert für eine Integration den Praxisbetrieb des fertigen Bot-Netz-Warnsystems. Wir schließen deshalb das Projekt mit der vorliegenden konsolidierten Version des Bot-Netz-Warnsystems ab.

3 Dokumentation der Projektergebnisse

Die Dokumentation der Projektergebnisse erfolgt über drei Schienen:

1. Auf der zuvor erwähnten Webpage http://botnetz-tracker.seewald.at ist sowohl Information über den Projektverlauf wie auch das aus dem Projekt resultierende Produkt, das Bot-Netz-Warnsystem, vorhanden. Gemeinsam mit diesem Abschlussbericht liegen damit alle Deliverables des Projektes vor.
2. In einer ersten Publikation auf einer wissenschaftlichen Tagung wurden erste für das Projekt relevante Ergebnisse (gemeinsam mit schon vor Projektbeginn erzielten Ergebnissen) zusammengefasst:
   - A.K. Seewald: *Towards Automating Malware Classification and Characterization*. Im Konferenzband der 4. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik (german-language proceedings), Saabrücken, April 2008, pp. 291-302.

   In diesem Beitrag wurde die Verwendung von Techniken des maschinellen Lernens zum Erlernen von globalen und lokalen Eigenschaften von Malware basierend auf bestimmten Verhaltensmustern der Malware untersucht.
3. Darüber hinaus arbeiten wir gerade an einem weiteren wissenschaftlichen Artikel, der die Ergebnisse dieses Forschungsprojektes in allen Details zusammenfassen wird. Im Zuge der Vorbereitungen dieses Artikels haben wir zusätzliche Analysen durchgeführt (ein paar Details dazu sind weiter unten zusammengefasst), Literaturrecherchen durchgeführt und ein Rahmenwerk für die erhaltenen Ergebnisse erarbeitet. Wir planen, diesen Artikel als Dokumentation dem Projektabschlussbericht beizulegen, sobald er begutachtet und veröffentlicht wurde (Umfang derzeit ca. 20 Seiten).

Um die vorliegenden Referenzdaten zum Spambot-Typ bestmöglich zu nutzen, haben wir beschlossen, die Zugriffsmuster – d.h., die Abfolge, in der eine bestimmte IP auf unser Darknet zugreift – zu analysieren. Fig. 1 zeigt eine Visualisierung von 288 zufällig ausgewählten Zugriffen mittels Sammon Mapping und Edit Distanz.

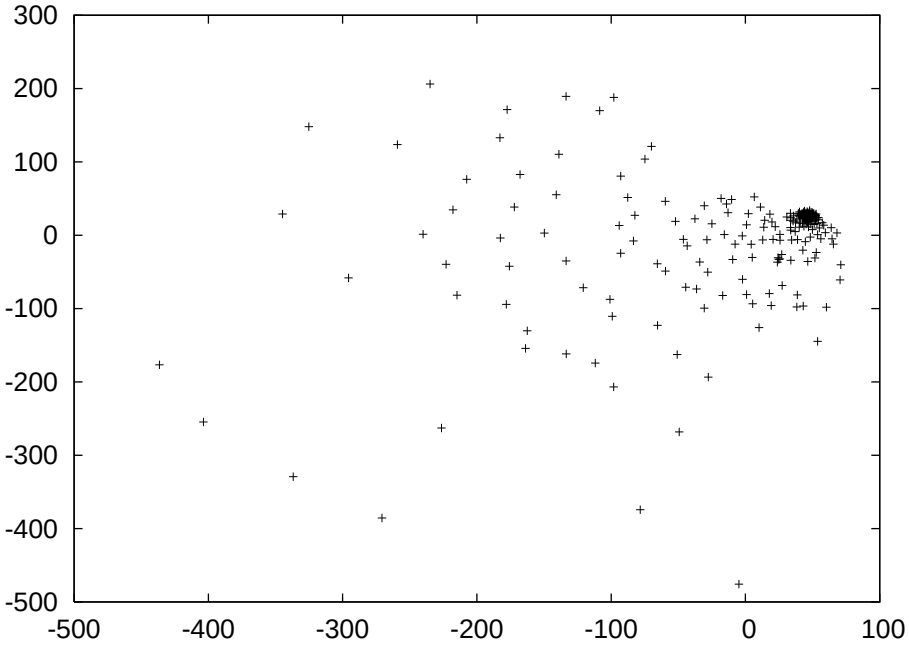

Fig. 1: Visualisierung von 288 zufällige ausgewählten Zugriffsmustern

Wie sich zeigt, ist die Verteilung der Zugriffe keinesfalls zufällig, sondern weist deutliche Muster auf. Aus diesem Grund haben wir die Zugriffsmuster genauer analysiert.

Fig. 2 visualisiert die Ähnlichkeit mittels Edit-Distanz für 32 Zugriffsmuster, die eindeutig einem Spambot-Typ zugeordnet werden konnten. Es zeigen sich Cluster von ähnlichen Zugriffsmustern, die jedoch nicht einem bestimmten Spambot-Typ zugeordnet werden können. Vielmehr zeigen vollkommen unterschiedliche Spambots faktisch die gleichen Zugriffsmuster. Dies weist möglicherweise auf einen gemeinsamen zugrunde liegenden Kontrollmechanismus hin.

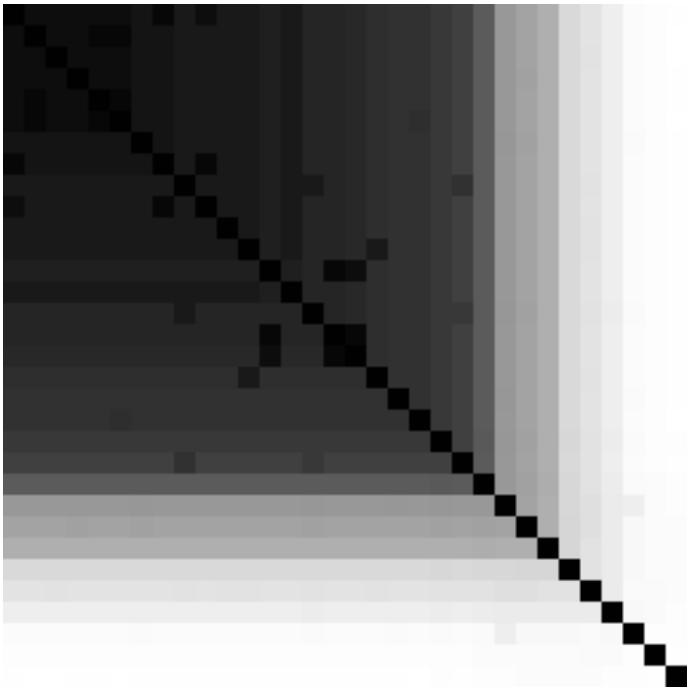

Fig. 2: Ähnlichkeitsmatrix zwischen Zugriffsmustern via Edit-Distanz (dunkel = ähnlich)

4 Fazit

Das hiermit abgeschlossene Netidee-Projekt „Frühwarnsystem für Bot-Netze“ hat ein hochaktuelles und sehr schwieriges Problem adressiert. Obwohl manche Aspekte der Gesamtproblematik ungelöst blieben und, wie zu erwarten war, ein paar der offenen Fragen nicht in voller Allgemeinheit beantwortet werden konnten, ist hervorzuheben, dass im Rahmen dieses Projektes sehr wichtige Schritte gesetzt werden konnten. Das entwickelte Bot-Netz-Warnsystem steht der Öffentlichkeit zur Verfügung und bietet vor allem auch dem Internet-Laien neue Möglichkeiten, Sicherheitsgefährdungen seines Computers zu erkennen. Damit wurde ein wichtiger Baustein zur Verbesserung der Sicherheit im österreichischen Internet entwickelt.

Wir danken der Internet Privatstiftung Austria für die Unterstützung dieses Projektes im Rahmen der Netidee-Initiative.