Identifikation von Angriffen auf Basis von Verkehrsanomalien

Thomas Gamer, Jannis Breitwieser

Institut für Telematik
Universität Karlsruhe (TH)

Angriffe im Internet, beispielsweise durch Internetwürmer oder Viren, gehören für die meisten der immer zahlreicher werdenden Internetnutzer mittlerweile zum Alltag. Während Bedrohungen wie Distributed Denial of Service-Angriffe (DDoS) früher häufig noch persönlich motiviert waren, wird heute immer häufiger versucht, durch Erpressungen finanziellen Gewinn zu erzielen oder Webseiten von Konkurrenten anzugreifen, denen durch die Nicht-Erreichbarkeit hohe Einnahmeverluste entstehen.

Aufgrund der ständigen Bedrohung durch Angriffe im Internet existieren zahlreiche Lösungsansätze zu deren Erkennung. Internetwürmer oder Viren werden beispielsweise meist mit Signatur-basierten Systemen wie z.B. Snort [1] identifiziert. Zur Erkennung von DDoS-Angriffen werden Anomalie-basierte Systeme benötigt [2,3], da sich DDoS-Angriffe, die gezielt Dienste durch Fluten ausschalten, in den meisten Fällen protokollkonform verhalten und dadurch nicht über Signaturen erkannt werden können. Will man Wurmausbreitungen nicht erst am Netzrand, sondern bereits im Netzinneren erkennen, wird ebenfalls auf Anomalie-basierte Systeme zurückgegriffen, da diese das Verkehrsverhalten und nicht Paketmuster untersuchen und somit nicht jedes einzelne Paket betrachtet werden muss.

Das in [3] entwickelte System ist ein Beispiel für ein solches Anomalie-basiertes Angriffserkennungssystem, das für den Einsatz im Netzinneren konzipiert wurde. Distack, ein Framework zur Anomalie-basierten Angriffserkennung durch verteilte Instanzen, ist eine Weiterentwicklung des in [3] vorgestellten Systems und bietet die Möglichkeit, verschiedene Verfahren zur Erkennung von Verkehrsanomalien als *Module* zu integrieren. Diese Module können anschließend über eine XML-basierte Konfiguration z.B. mit Sampling- oder Filtering-Modulen zu so genannten *Channels* kombiniert werden. Die Channels wiederum können hierarchisch konfiguriert werden, d.h. bestimmte Anomalie-Erkennungsverfahren werden erst dann gestartet, wenn ein anderes Verfahren bereits eine Anomalie erkannt hat. Dies ermöglicht eine grobgranulare Angriffserkennung mit geringem Ressourcen-Aufwand solange keine Anomalie beobachtet wird, beispielsweise durch ein einfaches Schwellenwert-basiertes Verfahren auf IP-Schicht. Bei Erkennung einer Anomalie können dann weitere Erkennungsverfahren gestartet werden, die zwar mehr Aufwand benötigen, dafür aber auch eine feingranularere Erkennung liefern. Hierbei kann es sich beispielsweise um die Erkennung von Protokollanomalien auf Transport- oder Anwendungsschicht handeln.

Zusätzlich zur einfachen Integration von Anomalie-Erkennungsverfahren in Form von Modulen stellt das Framework ein internes Nachrichtensystem zur Verfügung, das eine Kommunikation zwischen den Modulen ermöglicht. Dieses Nachrichtensystem arbeitet datenzentriert. Zur Signalisierung zwischen Detektionsinstanzen existiert außerdem eine wohldefinierte Schnittstelle, die beliebige externe Kommunikationsmethoden zur Verfügung stellen kann.

Als Ausgabe liefert ein solches Erkennungssystem eine Menge von erkannten Verkehrsanomalien. Zur Einleitung von Gegenmaßnahmen, zur einfacheren Interpretation oder zur Kommunikation zwischen Detektionsinstanzen eignet sich die Ausgabe des erkannten Angriffs jedoch wesentlich besser. Daher wird ein Identifikationsverfahren benötigt, welches aus den erkannten Anomalien auf den zugehörigen Angriff schließen kann. Ein solches Verfahren sollte einfach um weitere Anomalien und Angriffe erweiterbar sein sowie die Tatsache berücksichtigen, dass das Erkennungssystem im Netzinneren arbeitet – das entwickelte Verfahren muss also ressourcenschonend arbeiten. Zusätzlich muss bedacht werden, dass unterschiedliche Instanzen auch über eine unterschiedliche Menge an Anomalie-Erkennungsverfahren verfügen können.

Grundlage des entwickelten Identifikationsverfahrens ist ein *Anomalie-Angriffsmodell*, welches bekannten Angriffen im Internet jeweils die für sie typischen notwendigen und hinreichenden Anomalien zuordnet sowie innerhalb eines Angriffstyps eine hierarchische Struktur definiert, falls dies möglich ist. Auf Basis dieses Modells wurde ein zweistufiges Identifikationsverfahren entworfen, mit

dessen Hilfe sowohl die Ablaufsteuerung der verschiedenen Anomalie-Erkennungsverfahren als auch die schrittweise Identifikation eines Angriffs unter Beachtung von Ressourcenbeschränkungen möglich ist. Die erste Stufe wird durch einen regelbasierten Mechanismus realisiert, welcher aufgrund der notwendigen Definition von Regelsätzen initial zwar aufwändig zu konfigurieren ist, anschließend aber vergleichsweise wenig Identifikationsfehler aufweist. Die im Fall eines Misserfolgs gestartete zweite Stufe nutzt ein Klassifikationsverfahren, das auch bei kleineren Abweichungen zu den Regelsätzen mit hoher Wahrscheinlichkeit den korrekten Angriff identifizieren kann. Durch den Einsatz eines Klassifikationsverfahrens lassen sich außerdem bisher unbekannte Angriffe erkennen, da diese durch einen betragsmäßig großen Klassifikationsfehler auffallen.

Die erste Stufe der Identifikation – der so genannte *regelbasierte Koordinator* – bildet das Kernstück des Identifikationsverfahrens. Er übernimmt zum einen die Ablaufsteuerung des gesamten Prozesses, zum anderen ist er dafür zuständig, Angriffe mit Hilfe von extern konfigurierten Regelsätzen schrittweise zu identifizieren. Bei den eingesetzten Anomalie-Erkennungsverfahren wird zwischen *Initialverfahren* und *Konditionalverfahren* unterschieden. Initialverfahren besitzen keine Vorbedingungen und werden daher eingesetzt, um den Verkehr permanent zu beobachten und bei einem Angriffsverdacht einen Identifikationsvorgang zu starten. Konditionalverfahren hingegen werden maximal einmal während eines Identifikationsvorgangs eingesetzt, wenn all ihre Vorbedingungen erfüllt sind und noch Regelsätze vorhanden sind, welche den Einsatz des Verfahrens rechtfertigen. Durch dieses Vorgehen wird sichergestellt, dass nur die Erkennungsverfahren ausgeführt werden, die abhängig vom aktuellen Zustand der Identifikation noch einen Informationsgewinn ermöglichen. Falls zu einem Zeitpunkt mehrere Konditionalverfahren ausgeführt werden können, wird anhand ihres Ressourcenbedarfs, des zu erwartenden Mehrwerts und der aktuellen Auslastung der Detektionsinstanz entschieden, welche und wie viele der Verfahren als nächstes parallel gestartet werden.

Falls der regelbasierte Mechanismus am Ende eines Identifikationsvorgangs nicht in der Lage ist, basierend auf den erkannten Anomalien sowie dem Anomalie-Angriffsmodell einen Angriff zu identifizieren, wird anschließend ein Quader-Klassifikationsverfahren gestartet. Dieses Verfahren erhält als Eingabe alle verfügbaren Informationen der ausgeführten Anomalie-Erkennungsverfahren sowie bestimmte Referenzquader, die jeweils einen aus den Regelsätzen abgeleiteten bekannten Angriff darstellen. Das Klassifikationsverfahren berechnet auf Basis der Nearest-Neighbor-Distanz den wahrscheinlichsten Angriffstyp bzw. kann aus einem sehr hohen Abstand zu allen Quadern auf einen unbekannten Angriff schließen. Nach erfolgreicher Identifikation wird eine Angriffsbeschreibung generiert, welche möglichst viele nützliche Informationen über den vorliegenden Angriff enthält und daher auch zur Kommunikation mit anderen Instanzen genutzt werden kann.

Die Funktionsweise des entwickelten Identifikationsverfahrens wurde anhand verschiedener Simulationen mit dem diskreten Ereignissimulator OMNeT++ in Kombination mit dem INET-Framework evaluiert. Diese Evaluierung zeigte zum einen, dass neue Anomalie-Erkennungsverfahren einfach in das Identifikationsverfahren integriert werden können. Zum anderen zeigen die Ergebnisse der Simulationen, dass sich Angriffe zuverlässig identifizieren lassen, falls die Anomalie-Erkennungsverfahren korrekte Ergebnisse liefern. Unvollständige Ergebnisse können teilweise durch das Klassifikationsverfahren noch ausgeglichen werden.

Zukünftige Arbeiten sollten sich damit beschäftigen, wie Ressourcenbedarf und Mehrwert eines Erkennungsverfahrens bestimmt werden können und wie sich die Verfügbarkeit mehrerer unterschiedlicher Erkennungsverfahren für dieselbe Anomalie auf die Ablaufsteuerung und Identifikation auswirkt. Außerdem sollte untersucht werden, inwiefern das eingesetzte Klassifikationsverfahren selbstlernende Methoden nutzen kann um das zugrunde liegende Klassifikationsmodell den aktuellen Gegebenheiten anzupassen oder im Fall von unbekannten Angriffen weiterzuentwickeln.

Literaturreferenzen

[1] Martin Roesch. Snort. http://www.snort.org, 2001.

[2] Robin Sommer. Bro: An Open Source Network Intrusion Detection System. In: *Proceedings of the 17. DFN-Arbeitstagung über Kommunikationsnetze*, 2003, S. 273–288.

[3] Gamer, Thomas. *A System for in-Network Anomaly Detection*. In: *Proceedings of Kommunikation in Verteilten Systemen*, Informatik aktuell, Springer , Feb. 2007, S. 275–282.