# Einsatz der graphbasierten Meldungsstrukturanalyse in domänenübergreifenden Meta-IDS

Nils gentschen Felde (felde@nm.ifi.lmu.de)
Munich Network Management Team
Ludwig-Maximilians-Universität München
Oettingenstr. 67, D-80538 München

**Abstract:** Diese Arbeit beschreibt den Einsatz der Meldungsstrukturanalyse in einem so genannten Meta-IDS, das Ereignismeldungen mehrerer Partner in Koalitionsumgebungen zusammenführt und auswertet. Es handelt sich um ein graphbasiertes Anomalieerkennungsverfahren, das unabhängig von den zugrunde liegenden Meldungslieferanten und Sicherheitsvorschriften arbeitet. In diesem Beitrag wird die Leistungsfähigkeit des Verfahrens unter praxisrelevanten Bedingungen eines Koalitionsszenarios untersucht. Dabei wird einerseits auf real erfasste Ereignismeldungen, andererseits auf durch eine Simulation der Ausbreitung eines Internet-Wurms entstehende Meldungen zurückgegriffen. Es wird gezeigt, dass ein kooperativer Ansatz zur Anomalieerkennung für einen Teil der kooperierenden Partner einen zeitlichen Vorteil bei der Meldung von Auffälligkeiten mit sich bringt.

## 1 Einleitung

Bei der Betrachtung verteilt arbeitender Intrusion Detection Systeme besteht durch den Zusammenschluss kooperierender Partner die begründete Hoffnung, durch Zusammenführung und kooperative Auswertung von Meldungen über potentiell sicherheitsrelevante Ereignisse (sog. *Ereignismeldungen*) eine Steigerung der Möglichkeiten zur Erkennung von Vorfällen durch die größere Menge verfügbarer Informationen zu erzielen. Ein so genanntes *Meta-Intrusion Detection System (Meta-IDS)* verfolgt eben diesen Ansatz der kooperativen Intrusion Detection und hat die Aufgabe, alle anfallenden Meldungen lokaler IT-Sicherheitswerkzeuge (IDS, Paketfilter, Virenscanner, Integritätsprüfprogramme etc.) aus heterogenen Netzwerk- und Sicherheitsstrukturen zusammenzuführen und gemeinsam auszuwerten, wie auch in anderen Arbeiten beschrieben [1] [3]. Dabei werden a priori weder Annahmen über die Beschaffenheit der lokalen Sicherheitsvorkehrungen noch über die von ihnen durchzusetzenden Sicherheitsrichtlinien der teilhabenden Domänen gemacht. Einzige Voraussetzung ist, dass die eintreffenden Meldungen auf ein gemeinsames Datenmodell (z. B. IDMEF [2]) zurückgreifen.

Auf Grund der fehlenden Informationen über zugrunde liegende lokale IDS-Installationen sowie über durchzusetzende Sicherheitsrichtlinien kann zur Erkennung verdächtiger Aktivitäten nur ein auf Anomalieerkennung basierender Ansatz verwendet werden. Der vorliegende Beitrag untersucht die Einsatzfähigkeit der so genannten *Meldungsstrukturanalyse*, insbesondere unter verschiedenen praxisrelevanten Randbedingungen einer Koalitionsumgebung mit mehreren Domänen.

## 2 Meldungsstrukturanalyse im Meta-IDS

Das in dieser Arbeit vorgestellte prototypische System unterstützt die oben genannte Anforderung. Es wurde in einer Multi-Domänen-Konfiguration sowohl im operationellen Einsatz als auch unter Verwendung simulierter Daten einer Internet-Wurmausbreitung untersucht. Die untersuchte Konfiguration beinhaltet lokale IDS-Installationen in einer Forschungseinrichtung, einer Universität und einer Fachhochschule und ist in Abbildung 1 dargestellt. Weitere Details zum untersuchten System finden sich in [3] und [4].

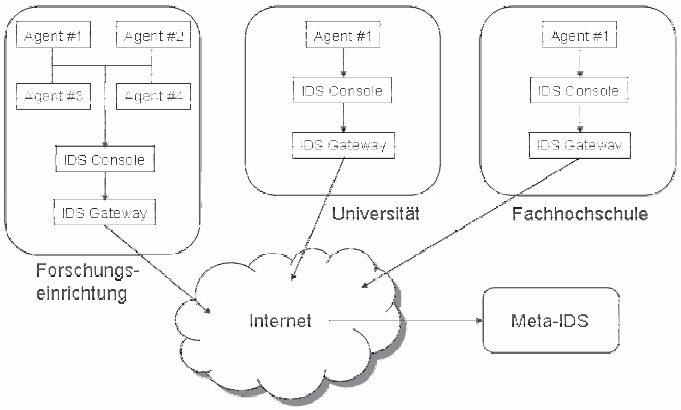


Abbildung 1: Struktur des Meta-IDS

Beim hier betrachteten zentralen Analyseverfahren handelt es sich um eine graphbasierte Meldungsstrukturanalyse, die über eine zuvor festzulegende Intervalllänge $\Delta t$ (hier: $\Delta t = 300$ Sekunden) alle eintreffenden Ereignismeldungen sammelt und entsprechend der Inhalte dieser Meldungen einen gerichteten Graphen $G = (V, E)$ pro Intervall erstellt. Die Knoten des Graphen sind mit Netzwerkadressen assoziiert. Zwei Knoten $v_1, v_2 \in V$ sowie eine gerichtete Kante $(v_1, v_2) \in E$ existieren in $G$ genau dann, wenn von der Analyseeinheit innerhalb $\Delta t$ eine Meldung erfasst wurde, die ein sicherheitsrelevantes Ereignis repräsentiert, deren (mutmaßliche) Quelle die Adresse $v_1$ und deren (wahrscheinliches) Ziel die Adresse $v_2$ ist. Die entsprechende Kante wird durch die in der Meldung angegebene Priorität bzw. den Schweregrad (*Severity*) gewichtet. Weiter wird ein zusätzlicher Pseudoknoten $v_3$, der mit dem betroffenen Netzdienst (bzw. der Portnummer) assoziiert ist, und zwei bidirektional gerichtete Kanten $(v_1, v_3)$, $(v_2, v_3) \in E$ eingeführt. Dies dient dazu, den Auslöser einer Anomalie besser identifizieren zu können.

Jeder Graph wird am Ende des entsprechenden Intervalls mittels Clustering-Algorithmen eindeutig in disjunkte, stark verknüpfte Teilgraphen zerlegt, wobei die mit Adressen und mit Portnummern assoziierte Knoten gleich behandelt werden. Das Clustering eines Graphen repräsentiert die typische Struktur im Meldungsaufkommen des Systems. Plötzliche Änderungen dieser Strukturen werden als Anomalie aufgefasst und gemeldet. Dazu ist eine Metrik erforderlich, die Ähnlichkeiten von aufeinander folgenden Clusterings $C_i$ und $C_{i+1}$ bewertet. Hier eignet sich beispielsweise ein Vergleichsmaß, das die Differenz zwischen zwei Clusterings $\delta_i$ durch die Anzahl der Elementaroperationen (Abspalten bzw. Hinzufügen eines Knotens aus einem bzw. in ein Cluster) definiert, die notwendig sind, um die Clusterings ineinander zu überführen. Vergleicht man das aktuelle

Clustering mit dem unmittelbar vorhergehenden, so erhält man den Wert $\delta_i$ für das aktuell betrachtete Zeitfenster (textit„timeframe“). Überschreitet $\delta_i$ ein aus der Historie der Clustering-Differenzen resultierendes Akzeptanzintervall um deren geglätteten Mittelwert $\overline{\delta}$ (d.h. $\delta_i \notin [\overline{\delta} - d \cdot \sigma_{\overline{\delta}}, \overline{\delta} + d \cdot \sigma_{\overline{\delta}}]$, mit der Standardabweichung $\sigma_{\overline{\delta}}$ des geglätteten Mittelwerts der Clustering-Differenzen $\overline{\delta}$), so liegt eine abnormale Struktur im Meldungsaufkommen vor. Der Faktor $d$ (hier: $d = 4.5$) sowie weitere variable Parameter des Verfahrens (z. B. Alterungsfaktoren für Mittelwert und Standardabweichung) werden heuristisch ermittelt. Alle aktuell ermittelten Clusterdifferenzen, die innerhalb des Akzeptanzintervalls liegen, werden als normale Zustände im zu beobachtenden Netz interpretiert. Ein Verstoß gegen das Akzeptanzintervall hingegen wird als Anomalie gedeutet und berichtet.

## 3 Einsatz des Verfahrens im domänenübergreifenden Betrieb

Das beschriebene Meta-IDS wurde in einem realen Koalitionsszenario mit drei unabhängigen Domänen, die jeweils lokale IDS-Instanzen betreiben, eingesetzt. Die aus den Domänen eintreffenden Ereignismeldungen entstammen unterschiedlichen Quellen: Netzwerk-IDS, Logfile-Analyzer für Paketfilter-Logs sowie für Protokolldateien von Betriebssystemen und Serverapplikationen (SMTP, HTTP, FTP) in der demilitarisierten Zone (DMZ) des Internet-Übergangs.

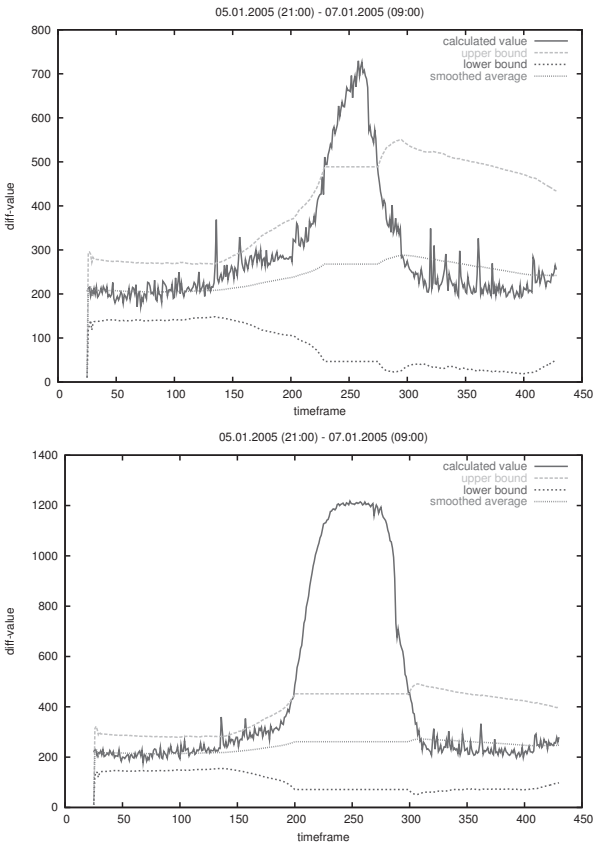


Abbildung 2: Vergleich der Erkennungsleistung mit einem Koalitionspartner (oben) und drei Koalitionspartnern (unten)

Zur Bewertung der Leistungsfähigkeit des Analyseverfahrens steht die Zeitnähe des Analysevorgangs, d.h. die Zeitdifferenz zwischen Auslösung und Meldung eines abnormalen Systemverhaltens, im Mittelpunkt.

Die Funktionsfähigkeit des Analyseverfahrens wurde zunächst ausschließlich anhand realer Phänomene untersucht. Im aufgezeichneten Meldungsaufkommen des Koalitionsszenarios im Zeitraum einer Woche verzeichnete die Meldungsstrukturanalyse verschiedene Anomalien, die nach eingehender Untersuchung auf verschiedene tatsächliche Vorgänge zurückgeführt werden

konnten. Dazu gehörte die Ausbreitung verschiedener, im untersuchten Zeitraum aktiver Internet-Würmer (z. B. Beagle/Bagle), aber auch massive Zugriffe auf einzelne Rechner, die offensichtlich auf Fehlkonfigurationen zurückzuführen waren. Außerdem konnten Fehlkonfigurationen der IT-Sicherheitswerkzeuge in den Analyseergebnissen der Meldungsstrukturanalyse identifiziert werden. Weitere Informationen finden sich in [4].

Um die Zeitnähe des Analyseverfahrens zu untersuchen, wurden zusätzlich simulierte Meldungen eines Paketfilters in das Meta-IDS eingebracht, die bei der Ausbreitung eines Internet-Wurmes (hier: CodeRed v.2, Code Red II und Scalper) entstehen würden. Es galt, die Zeitdifferenzen zwischen dem Beginn der Injizierung der simulierten Meldungen und der Indikation eines abnormalen Systemverhaltens in verschiedenen Konstellationen der Kooperation der drei verfügbaren Koalitionspartner unter verschiedenen praxisrelevanten Randbedingungen einer Koalitionsumgebung zu bestimmen. Eine besonders bedeutsame Randbedingung ist das so genannte *Rate-Limiting* bei der Erzeugung von Meldungen eines Firewall-Paketfilters. Meldungen von Firewalls stellen bei der Untersuchung von Wurmausbreitungen die für den Analysevorgang bedeutsamste Informationsquelle dar. Im hier betrachteten Fall werden Meldungen mit einer maximalen Rate von 1Hz erzeugt, wobei verschiedene Verwerfungsstrategien untersucht wurden.

Abbildung 2 (oben) zeigt die Clusterdifferenz für den Meldungsbestand einer einzelnen Domäne bei injizierter Wurmausbreitungsstrategie von CodeRed v.2 und einem Rate-Limiting von 1Hz (*„calculated value“*), die Grenzen des Akzeptanzintervalls (*„upper-/lower bound“*) und den geglätteten Mittelwert (*„smoothed average“*) der Clusterdifferenzen. Der untere Teil von Abbildung 2 hingegen veranschaulicht die gemeinschaftliche Analyse der Meldungen aller drei Koalitionspartner unter gleichen Bedingungen.

In beiden gezeigten Fällen ist es eindeutig möglich gewesen, die injizierte Wurmausbreitung, die sich durch die Verstöße der oberen Grenze des Akzeptanzintervalls ausdrückt, zu erkennen (die Wurmausbreitung begann ca. im Timeframe 150 (06.01.05, 9:00 Uhr); eine Anomalie wurde im Timeframe 223 (15:48 Uhr) bzw. 194 (13:20 Uhr) erkannt).

Abbildung 3 veranschaulicht Messungen im angesprochenen Szenario mit Meldungen aus drei unabhängigen Domänen. Hier sind die einzelnen Erkennungszeitpunkte einer um 09:00 Uhr beginnenden, simulierten Wurmausbreitung aller möglichen Konstellationen einer Kooperation unter den bereits angesprochenen Bedingungen angeführt.

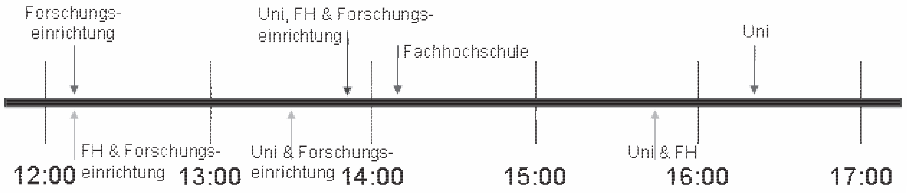


Abbildung 3: Erkennungszeitpunkte bei 1Hz Rate-Limiting

## 4 Diskussion der Ergebnisse

Der Einsatz der graphbasierten Meldungsstrukturanalyse führte zunächst grundsätzlich zur Erkennung von Phänomenen, die beim Einsatz konventioneller Verfahren erst bei der Offline-Auswertung der Meldungs-Datenbank entdeckt worden wären.

Beim Einsatz des Verfahrens in einer Koalitionsumgebung ergeben sich verschiedene Konsequenzen. Der Hauptunterschied der Domänen untereinander bezüglich der entstehenden Paketfilter-Meldungen besteht in der Größe der betrachteten IP-Adressbereiche. Daher überrascht es keineswegs, dass bei der kombinierten Betrachtung der Meldungen eine Effizienzsteigerung ausschließlich für Domänen auftritt, die einen kleinen, vom Wurm betroffenen Adressbereich abdecken, wohingegen Domänen mit großem betroffenem Adressbereich einen Verlust bezüglich der Erkennungseffizienz hinnehmen müssen. Daraus folgt, dass eine Kombination von lokaler und kooperativer Auswertung sicherheitsrelevanter Ereignisse die besten Ergebnisse liefert, da so einem etwaigen Zeitverlust bezüglich der Erkennungsleistung des Systems bei der Betrachtung des kombinierten Meldungsaufkommens vorgebeugt werden kann.

## 5 Zusammenfassung und Schlussfolgerungen

Diese Arbeit beschreibt die Untersuchung der Leistungsfähigkeit eines Anomalieerkennungsverfahrens im Einsatz in einem domänenübergreifenden Meta-IDS. Dabei stehen vor allem praxisnahe Randbedingungen im Vordergrund. Zur Bewertung wurden sowohl in der Realität erfasste Meldungen über potentiell sicherheitsrelevante Ereignisse (*Ereignismeldungen*), als auch Meldungen einer Simulation der Ausbreitung von Internet-Würmern herangezogen. Durch den Vergleich der Analyseergebnisse bei separaten und bei gemeinsamen Meldungsvolumina konnte für den beschriebenen Anwendungsfall eine Steigerung der Erkennungsleistung für den Fall nachgewiesen werden, dass die betrachtete Domäne nur über einen verhältnismäßig kleinen, vom Wurm betroffenen IP-Adressraum verfügt. Einem etwaigen Verlust bei der Erkennungsleistung kann durch kombinierte Betrachtung lokaler und gemeinsamer Auswertungen entgegengewirkt werden.

**Danksagung:**
Der Autor dankt den Mitgliedern des Münchener Netzwerk-Management Teams (MNM Team) für hilfreiche Diskussionen und wertvolle Kommentare zu früheren Versionen dieses Artikels. Das MNM Team ist eine Forschungsgruppe der Münchener Universitäten und des Leibniz-Rechenzentrums der Bayerischen Akademie der Wissenschaften unter der Leitung von Prof. Dr. Heinz-Gerd Hegering.

## Literatur

[1] J. Mirkovic, M. Robinson, P. Reiher and G. Kuenning: „Alliance Formation for DDoS Defense“. Proceedings of the New Security Paradigms Workshop, ACM SIGSAC, 2003.

[2] The Internet Engineering Task Force - Intrusion Detection Working Group: „The Intrusion Detection Message Exchange Format“, 27. Januar 2005. `http://www.ietf.org/internet-drafts/draft-ietf-idwg-idmef-xml-14.txt`

[3] M. Jahnke, M. Lies, S. Henkel, M. Bussmann und J. Tölle. „Komponenten für kooperative Intrusion-Detection in dynamischen Koalitionsumgebungen“. Proceedings of the GI Workshop on Detection of Intrusions and Malware & Vulnerability Assessment (DIMVA2004), Dortmund, 2004.

[4] N. O.v.d. gentschen Felde: „Leistungsfähigkeit von Anomalieerkennungsverfahren in domänenübergreifenden Meta-IDS“. Diplomarbeit, Universität Bonn, Institut für Informatik, Abt. IV, 2005.