(19) Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 10 2005 049 561 A1** 2007.04.19

(12) # Offenlegungsschrift

(21) Aktenzeichen: **10 2005 049 561.3**
(22) Anmeldetag: **12.10.2005**
(43) Offenlegungstag: **19.04.2007**

(51) Int Cl.[8]: ***H04L 12/26*** (2006.01)
H04L 9/00 (2006.01)

(71) Anmelder:
Deutsche Telekom AG, 53113 Bonn, DE;
Technische Universität Berlin, 10623 Berlin, DE

(72) Erfinder:
Müller, Achim, 10825 Berlin, DE; Albayrak, Sahin, 14195 Berlin, DE; Bub, Udo, 10405 Berlin, DE; Feil, Peter, 10585 Berlin, DE; Solarski, Marcin, 10553 Berlin, DE

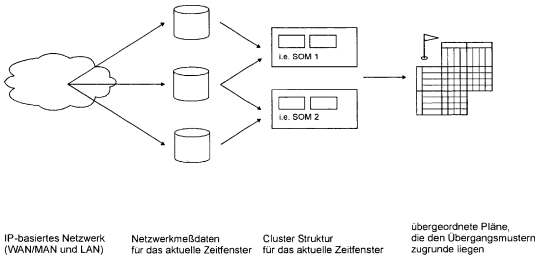
(56) Für die Beurteilung der Patentfähigkeit in Betracht zu ziehende Druckschriften:
DE 199 29 166 A1
US 61 78 450 B1
EP 15 80 957 A2
EP 15 07 360 A1
WO 02/21 802 A1

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

Rechercheantrag gemäß § 43 Abs. 1 Satz 1 PatG ist gestellt.

(54) Bezeichnung: **Verfahren zur automatischen Erkennung von Anomalien in Weitverkehrsnetzen (WAN) und lokalen Netzen (LAN)**

(57) Zusammenfassung: Die vorliegende Erfindung bezieht sich allgemein auf das Gebiet der Netzwerktechnik und betrifft die Erkennung von Anomalien, wie insbesondere Netzwerkangriffen, in Weitverkehrsnetzen und lokalen Netzen. Erfindungsgemäß werden in kontinuierlichen Zeitabständen Netzwerkmessdaten erfasst, die geeignet sind, fortlaufend eine symbolische Darstellung der in dem zu überwachenden Netzwerk auftretenden Netzwerkaktivitäten zu generieren. Für die symbolische Darstellung der Netzwerkaktivitäten werden die erfassten Netzwerkdaten mittels Techniken des "unüberwachten Lernens" komprimiert. Im Anschluss daran werden an sich bekannte Techniken der Planerkennung (SOM) auf die in symbolischer Form erfassten Netzwerkdaten angewendet, um die dem Lernverhalten zugrunde liegenden Regeln zu erkennen. Auf der Grundlage dieses Lösungsansatzes lassen sich zukünftig Abweichungen vom "normalen" Netzverhalten nahezu in Echtzeit erkennen.



Beschreibung

[0001] Die vorliegende Erfindung bezieht sich allgemein auf das Gebiet der Netzwerktechnik und betrifft die Erkennung von Anomalien in Weitverkehrsnetzen und lokalen Netzen. Darunter wird insbesondere die Erkennung von Netzwerkangriffen durch böswillige Benutzer/Angreifer verstanden, die beispielsweise versuchen, einen oder mehrere Server durch einen Dienstverweigerungs-Angriff/Denial-of-Service-Attac/DoS-Angriff lahm zu legen.

Stand der Technik

[0002] In www.heise.de/ix/ artikel/2205/04/107 Artikel Christoph Puppe, Jörn Maier ist ausgeführt, dass die ersten Denial of Service Angriffe, die bekannt geworden sind, auf die Webseiten von Yahoo, CCN eBay Amazon und anderer Diensteanbieter erfolgten. Am 09.02. 2000 begann „Mafiaboy" seine Angriffe und sorgte eine Woche lang für massive Störungen. Am 10. März 200 brach das Vertrauen der Anleger in die Internet-Aktien zusammen und an der Börse wurden 2,7 Billionen US-Dollar vernichtet.

[0003] Die Folgen einer solchen Attacke sind neben einem immensen materiellen Verlust immer auch ein Vertrauensverlust der Kunden und Anleger in das angegriffene Unternehmen. Sicherheitsfragen von Computersystemen sind daher, insbesondere auch aus den o.a. Gründen, in der letzten Zeit immer wichtiger geworden. Das liegt darin begründet, dass Einerseits die Anzahl der Angriffe auf Netzwerke, und insbesondere auch auf Server, ständig zunimmt, und dass Andererseits immer mehr Computer sowohl mit Weitverkehrsnetzen (WAN), als auch mit lokalen Netzwerken (LAN), verbunden sind.

[0004] Primitive DoS-Angriffe belasten die Dienste eines Servers, beispielsweise HTTP, mit einer größeren Anzahl von Anfragen als dieser bearbeiten kann. Im Ergebnis wird der betroffene Server die Bearbeitung einstellen oder die regulären Anfragen so langsam beantworten dass die Beantwortung abgebrochen wird.

[0005] Noch effektiver in seiner Wirkung ist ein Angriff, der darauf beruht, einen Programmfehler auszunutzen um eine Fehlerfunktion in der Serversoftware auszulösen, welche dann im Ergebnis zum Absturz des Servers führt. Einem solchen Angriff kann jedoch die Grundlage durch Verbesserung der Softwareprogramme bzw. durch Vermeidung von Programmfehlern entzogen werden.

[0006] Erfolgt der Angriff koordiniert von einer größeren Anzahl anderer Systeme, spricht man von einem Distributet-Denial-of-Service-Angriff, DDoS-Angriff. Strategie dieser Angriffe ist es, mit Backdoor Programmen, welche sich alleine auf anderen Rechnern im Netzwerk verbreiten, dem Angreifer weitere Wirte zum Ausführen seiner Angriffe zur Verfügung zu stellen.

[0007] Eine besondere Form des Angriffes ist die distributed reflected denial of service Attacke DRDoS-Attacke. Bei dieser Art des Angriffes adressiert der Angreifer seine Datenpakete nicht direkt an das Opfer, sondern an regulär arbeitende Internet-Dienste. Er trägt jedoch als Absenderadresse die Adresse des Opfers ein (IP-Spoofing). Die Antworten auf diese Anfragen stellen dann für das Opfer den eigentlichen DoS-Angriff dar. Der Ursprung des Angriffs ist für das Opfer durch diese Verfahrensweise praktisch nicht mehr ermittelbar.

[0008] Es ist auch prinzipiell festzustellen, das nach wie vor die Anfälligkeit vor DoS-Angriffen eine der gravierendsten Schwachstelle des Internet ist.

[0009] Die Schwachstellen des Internets, die DoS-Angriffe erst ermöglichen, sind systemimmanent und liegen insbesondere darin begründet, das es im Internet nicht vorgesehen ist, jeden Kommunikationspartner eindeutig zu identifizieren, eine zentrale Verwaltung der Datentransfers zu bieten oder jedem Teilnehmer einen Anteil an der vorhandenen Bandbreite zu garantieren.

[0010] Für das Erkennen von Sicherheitsproblemen in einem Rechnernetz gibt es zwei unterschiedliche Lösungsansätze. Der erste Lösungsansatz wird als wissensbasiertes Verfahren bezeichnet. Dem wissensbasierten Verfahrensansatz liegt die Annahme zugrunde, dass dem System alle möglichen Angriffsarten bekannt sind. Für jeden Angriff gibt es eine definierte Signatur. Ein System zur Überwachung des Datenverkehrs sucht nach diesen Signaturen und identifiziert sie.

[0011] In US-Patentschrift 5,278,901 wird ein derartiges signaturbasiertes System beschrieben.

[0012] Vorteil dieses Verfahrens ist es, dass sehr selten ein falscher Alarm ausgelöst wird, d. h. die Fehlerrate ist verhältnismäßig niedrig. Der wesentliche Nachteil dieses Lösungsansatzes wird darin gesehen, dass systembedingt auch nur die Angriffe erkannt werden können, deren Signaturen bereits gespeichert sind. Gegenwärtig sind jedoch so viele Angriffe zu verzeichnen, dass die Anzahl der unterschiedlichen Signaturen sehr schnell zunimmt. Viele Signaturen sind auch so strukturiert, dass sie sich schwierig über einen Algorithmus darstellen lassen. Zwei Beispiele für Produkte, die auf dem wissensbasierten Lösungsansatz beruhen, sind NetRanger von Cisco Systems (siehe www.alliancedaracom.com/manifacturers/cisco-systems/security_vpn/ids.asp) und RealSecure von Internet Security Systems Inc. (siehe www.rio-

co.co.uk/products/intrusion_detektion/Datasheet_RealSecure.pdf)

[0013] Der zweite Lösungsansatz wird als verhaltensbasierter Ansatz bezeichnet. Er geht von der Annahme aus, dass sich das „Verhalten" eines Systems verändert, wenn ein Angriff auf das System erfolgt. Bei dem verhaltensbasierten Ansatz wird für das System ein Normalprofil definiert. Ausgehend von dem Normalprofil wird nach Abweichungen gesucht.

[0014] Zwei Beispiele für neue, gerade angekündigte Produkte, die auf dem verhaltensbasierten Lösungsansatz beruhen sollen, sind Cisco Guard XT 5650 (siehe www.cisco.com/en/US/products/ps5894) bzw. Cisco Traffic Anomaly Detector XT 5600 (siehe www.cisco.com/en/US/products/ps5892).

[0015] Für die Realisierung dieses verhaltensbasierten Lösungsansatzes können verschiedene Verfahren, z. B. statistische Verfahren, regelbasierte Systeme und neuronale Netze eingesetzt werden. Eine Art von künstlichen neuronalen Netzen stellen beispielsweise selbstorganisierende Karten (Self-Organizing Maps/SOMs oder Self-Organizing Feature Maps) dar, die nach Teuvo Kohonen auch als Kohonennetze bezeichnet werden.

[0016] Mittels der oben genannten Verfahren können unterschiedliche Überwachungsziele, wie beispielsweise die Benutzer des Systems, die Leistungsparameter des Rechnernetzes, die CPU-Takte usw. definiert werden. Der Hauptvorteil gegenüber dem wissensbasierten Ansatz ist, darin zu sehen, dass beim verhaltensbasierten Lösungsansatz auch unbekannte Angriffsarten erkannt werden können. Es entfällt bei diesem Lösungsansatz somit der Zwang eine Datenbank permanent mit den bekannten Signaturen aktualisieren zu müssen.

[0017] Der Nachteil dieses Lösungsansatzes beruht darin, dass es auch ohne einen Angriff zu Abweichungen kommen kann, die beispielsweise durch Änderungen der Benutzeraktivitäten, neue Software, neue Maschinen und neue Benutzer ausgelöst werden.

[0018] Die Schwierigkeit bei diesem Lösungsansatz besteht daher darin, normales Nutzerverhalten von einem böswilligen Angriff zu unterscheiden.

[0019] Der verhaltensbasierte Ansatz kann daher aufgrund des angewendeten Lösungsprinzips zu einer sehr hohen Anzahl von Fehlalarmen führen, die das Verfahren für den Anwender unattraktiv machen.

[0020] Eine weitere Lösung, die auf dem verhaltensbasierten Lösungsansatz beruht, ist aus DE 698 17 176 T2 bekannt. Bei dieser Lösung handelt es sich um ein Verfahren und eine Vorrichtung zur Eindringdetektion, vorrangig in einem Rechnersystem, das auf Ereignismustern basiert und insbesondere die Erkennung von Abweichungen von einem „normalen" Prozessverhalten und somit die Erkennung von Angriffen gegen diesen Prozess betrifft. Diese Lösung basiert

a) auf einem Trainingsmodus, dem eine Tabelle charakteristischer, den Prozess darstellender Muster zugrunde liegt, die das normale Verhalten eines Modellprozesses in dem Rechner definieren, durch Ausführen der Schritte

– Erstellen einer ersten Ereignissequenz durch Filtern eines ersten durch den Modellprozess generierten Ereignisdatenstroms

– Verwenden des Teiresias-Algorithmus zum Extrahieren von Ereignissequenzmustern, aus der ersten Ereignissequenz, wobei die Muster den Modellprozess darstellen,

– Speichern der den Prozess darstellenden Muster und

b) auf einem Betriebsmodus, in dem aus einem realen Prozess charakteristische Muster durch Ausführen der folgenden Schritte extrahiert werden;

– Erstellen einer zweiten Ereignissequenz durch Filtern eines durch den realen Prozess generierten Ereignisdatenstromes,

– Vergleichen der ersten Ereignissequenz mit den gespeicherten, den Prozess darstellenden Mustern und

– Anzeigen des Ergebnisses des Vergleichsschrittes.

[0021] Bei dem Teiresias Algorithmus, siehe auch US 6,108,666A und DE 698 17 176T2 handelt es sich um einen Suchalgorithmus der 1996 von IBM entwickelt wurde und der ursprünglich dazu vorgesehen war, genetische Sequenzdaten nach wiederkehrenden genetischen Muster zu durchsuchen. Zukünftig soll dieser Algorithmus im Rahmen eines Anti-Spam-Filters eingesetzt werden.

[0022] Weiterhin sind Verfahren bekannt, die auf die automatische Bildung von Hypothesen über Intentionen aus beobachtetem Verhalten ausgerichtet sind (Verfahren zur Planerkennung). Unter Intentionen werden dabei solche Ziele/Pläne eines Akteurs bzw. einer Gruppe von Akteuren verstanden, die von ihm bzw. von ihnen innerhalb der nächsten Zeit angestrebt werden und die ihm auch in diesem Zeitraum erreichbar erscheinen. Nicht unter den Begriff der Intention fallen also längerfristige Ziele und Ziele, über deren Umsetzung sich der Akteur noch keine Gedanken gemacht hat (siehe www.dfki.uni-sb.de/vitra/papers/sport89/node9.html).

[0023] Die existierenden Verfahren gehen im Kern zumeist davon aus, dass Daten und Messungen auf

Rechnersystemen – sogenannten Hosts – durchgeführt werden. Diese Ansätze sind für einen direkten Einsatz in den Netzwerk-Infrastrukturen selber nur bedingt geeignet, da dort auf einer höheren Aggregationsebene Daten zusammengefasst und analysiert werden müssen. Dies erfordert wesentlich komplexere und auf die besondere Situation in den Netzen ausgerichtete Formen von Erkennungsmechanismen.

Aufgabenstellung

[0024] Ziel der vorliegenden Erfindung ist es, unter Zugrundelegen des verhaltensbasierten Lösungsansatzes Denial-of-Service-Angriffe in IP-basierten Netzwerken nahezu in Echtzeit zu erkennen, wobei die unterschiedlichen Anforderungen an eine automatische Erkennung von Denial-of-Service-Angriffen in Einklang zu bringen sind.

[0025] Ausgehend vom verhaltensbasierten Lösungsansatz beruht die Erfindung auf der Einführung eines kontinuierlichen Lernprozesses mit dem sich „normales Verhalten" von Netzwerkressourcen verstehen lässt. Das wird durch ein Verfahren erreicht, bei dem zwei an sich bekannte Lösungsansätze erstmalig im Rahmen einer neuen eigenständigen Lösung zusammengeführt werden.

[0026] Erstens werden in kontinuierlichen Zeitabständen Netzwerkmessdaten erfasst, die geeignet sind, fortlaufend eine symbolische Darstellung der in dem zu überwachenden Netzwerk auftretenden Netzwerkaktivitäten zu generieren. Für die symbolische Darstellung der Netzwerkaktivitäten werden die erfassten Netzwerkdaten mittels Techniken des „unüberwachten Lernens" komprimiert.

[0027] Unter Techniken des unüberwachten Lernens wird hier maschinelles Lernen ohne im Voraus bekannte Klassen verstanden. Beispiele für solches unüberwachtes maschinelles Lernen sind die automatische Gruppenbildung (Clustering) oder die Komprimierung von Daten zur Dimensionsreduktion. Ausgehend von den gemäß der Prinzipien des so verstandenen „unüberwachten maschinellen Lernens" ermittelten und komprimierten Informationen zu den Netzwerkaktivitäten bzw. zum Netzwerkstatus werden diese komprimierten Netzwerkmessdaten in symbolischer Form durch Clusterstrukturen dargestellt.

[0028] Zweitens werden an sich bekannte Techniken der Planerkennung auf die in symbolischer Form, beispielsweise durch SOMs dargestellten Netzwerkdaten angewendet. Damit sollen die dem Lernverhalten zugrunde liegenden Regeln erkannt und modelliert werden. Auf der Grundlage dieses Lösungsansatzes lassen sich zukünftig Abweichungen vom „normalen Netzverhalten nahezu in Echtzeit erkennen. Ein Beispiel für die Erfassung von Netzwerkmessdaten in Cluster-Strukturen und Plan- Datenbanken ist in Fig. 1 dargestellt.

[0029] Zum Sammeln der Netzwerkmessdaten gibt es grundsätzlich zwei unterschiedliche Arten der Messung.

[0030] Bei der passiven Messung erfolgt das Sammeln der Netzwerkmessdaten durch Abtasten (Sampling) und durch Verarbeiten von Datenpaketen aus dem Benutzerverkehr. Die aktive Messung erfolgt durch Abtasten und Verarbeiten von Testpaketen, die zu Messzwecken zusätzlich in das Netzwerk injiziert werden. Um eine Minderung der Netzwerkleistung festzustellen, werden beispielsweise folgende Messgrößen ermittelt:

– Konnektivität (connectivity)
– Einfache Verzögerung (one-way-delay) und Umlaufverlust (one-way-loss)
– Netzlaufzeit (round-trip-delay)
– Laufzeitvariation (delay variation)
– Verlustmuster (loss patterns)
– Umordnung der Pakete (packet reordering)
– Massentransportkapazität (bulk transport capacity) und
– Link-Bandbreitenkapazität (link bandwidth capacity)

[0031] Netzwerkmessdaten, die nach den oben beschriebenen Verfahren ermittelt wurden, werden in vordefinierte Kategorien eingeordnet, oder mit Hilfe von „unüberwachtem maschinellen Lernen" in Gruppen/Clustern eingeteilt. Dabei handelt es sich jedoch nicht um die Zuordnung zu vordefinierten Kategorien, sondern um „gelernte" Kategorien, die sich automatisch aufgrund unterschiedlicher Netzaktivitäten als Beobachtungsschwerpunkte erwiesen haben.

[0032] Danach werden die erfassten und zugeordneten Daten als Trainingsmaterial zur Erkennung von Plänen bei einem speziellen Typ von Verkehrs- oder Protokolldaten verwendet. Dabei kann es sich beispielsweise um den mittleren Durchsatz von Edge-Routern in einem Netzwerksegment oder auch um die durchschnittliche Antwortzeit von Webservern bei http-Anfragen in einer Hosting-Umgebung handeln. Voraussetzung ist, dass die gewählten Datenströme Aspekte der realen Welt wiedergeben und damit geeignet sind, alle relevanten Netzwerkaktivitäten zu repräsentieren, wobei genügend Messungen vorliegen müssen, um hinreichend stabile Clusterstrukturen zu erkennen.

[0033] Innerhalb welchen Zeitraumes eine Clusterstruktur stabil bleiben muss, richtet sich nach der Art der Anwendung, d. h. danach, ob eine Leistungsminderung von Routern auf der Grundlage von Clustern auf Verkehrsspitzen-Szenarien hin beobachtet werden soll, oder ob es auf das Klick- und Surfverhalten

von Internetnutzern ankommt.

[0034] Beispielsweise wird bei einer selbstorganisierenden Karte nach einer zu vernachlässigenden Trainingszeit aus den Messergebnissen eine erste Clusterstruktur C1 ermittelt. Die Clusterstruktur bleibt über einen gewissen Zeitraum stabil und beginnt dann zu zerfallen bzw. in eine andere Clusterstruktur C2 überzugehen, die wiederum nach einer gewissen Zeit zerfällt.

[0035] Bemerkenswert an diesem Prozess ist folgendes:
Unüberwachtes maschinelles Lernen setzt in der Regel voraus, dass die zu beobachtende Domaine eine gewisse Stabilität aufweiset. Das unterstützt die weitere Verwendung der identifizieren Cluster als Eingabe für darauf aufbauende Planerkennungsalgorithmen, welche die identifizierten Cluster nach wiederkehrenden Mustern durchsuchen. Das geschieht allerdings auf einer abstrakten Ebene innerhalb der vorliegenden Clusterstrukturen.

[0036] Wird die beobachtete Welt instabil, beginnt der gesamte Trainingszyklus von neuem. Bei einem IP-basierten Datennetzwerk kann die Clusterbildung vor allem dadurch instabil werden, dass in der realen Welt eine größere Veränderung eintritt, die durch die sich verändernden Messergebnisse manifestiert wird.

[0037] Im vorstehend beschriebenen Beispiel einer Router- oder Serverbelastung kann eine solche Veränderung schon dadurch hervorgerufen werden, dass eine beliebte Fernsehsendung ausgestrahlt wird und dadurch größere Benutzergruppen zeitgleich von der Internet- zur Fernsehnutzung übergehen. Das aus diesem Vorgang resultierende Ergebnis wäre eine plötzliche Änderung der Verkehrs- oder Lastmuster, die je nach der geographischen Verteilung der Benutzer oder ihrer individuellen Präferenzen bei der Internetnutzung wiederum Auswirkungen auf die jeweilige Clusterstruktur haben. Andererseits ist eine Stabilität zu beobachten, wenn große Benutzergruppen kohärentes Verhalten zeigen, was sich beispielsweise in typischen Internetzeiten, wie z. B. um 22.00 Uhr oder in der späten Nacht, nachweisen lässt.

[0038] Im Hinblick auf die erfindungsgemäße Architektur lassen Perioden stabiler Clusterstrukturen darauf schließen, dass keine größeren Veränderungen eingetreten sind. Dagegen kann eine Instabilität, wie bereits beschrieben, natürliche Ursachen haben oder aber auf böswilligem Verhalten beruhen, wie etwa einem verteilten Denial-of-Service-Angriff; DDoS-Angriff. Absichtliches Überfluten von Netzwerksegmenten oder Massenanfragen an Zielserver lösen eine kurzfristige Änderung der Clusterstruktur aus, sofern sie nicht zeitgleich mit einer natürlichen Stabilitätsänderung der Clusterstruktur zusammenfallen.

[0039] Da die zugrunde liegende reale Welt, die durch Messungen in einem IP-basierten Netz abgebildet wird, typischen Mustern folgt, beispielsweise, wenn ganze Benutzergruppen während eines definierbaren Zeitraumes zum Fernsehkonsum übergehen, lässt sich der Übergang der Clusterstruktur als planbasiertes Verhalten erklären. Unter einem Plan wird hier regelbasiertes Verhalten verstanden, welches durch eine Gruppe autonomer Agenten verursacht wird, deren Verhalten wohldefinierte Ziele zugrunde liegen. Dabei werden diese Ziele jedoch nicht offen gelegt. Weiterhin wird davon ausgegangen, dass die genaue Definition der Agenten ebenfalls unbekannt ist. Lediglich die Existenz dieser Agenten und die Tatsache, dass das Verhalten der Agenten als planbasiert betrachtet werden darf, ist als gesichert anzusehen.

Ausführungsbeispiel

[0040] Nachfolgend wird ein Beispiel für planbasiertes Verhalten von Benutzergruppen anhand des Alltags von Studenten beschrieben.

[0041] Betrachten wir, wie allgemein in Fig. 2 abgebildet, den Alltag junger Studenten mit einer Vorliebe für abendliche Kino- oder Clubbesuche und nächtlichem Surfen im Internet. Die diesem Beispiel zugrunde liegende reale Welt sei eine große Universitätsstadt, wie beispielsweise Berlin. Die diesem Szenario zugrunde liegende reale Welt wird durch die bereits beschriebenen Messungen des Datenverkehrs in ausgewählten Edge-Routern und durch beobachtbare http-Anfragen erfasst. Die Messungen erfolgen dabei beispielsweise im Abstand von 100 Millisekunden und können ca. 200 Werte je Messdurchlauf umfassen.

[0042] Nach einer Zeit intensiven Internetsurfens durch eine Vielzahl von Studenten, die in einen Zeitraum von ca 22:00 Uhr bis 04:00 Uhr fällt, begeben sich die meisten Studenten etwa zwischen 04:00 Uhr und 05:00 Uhr zur Ruhe um etwas Schlaf zu finden, ehe um 09:00 Uhr die nächste Vorlesung beginnt.

[0043] Wenn in der gewählten Stadt signifikant viele Studenten diesem Verhaltensmuster folgen, ist nach einem Retraining der SOMs eine Änderung der Clusterstruktur zu beobachten. Da die Cluster selber aber noch keinen Aufschluss über die Pläne, die den Änderungen zugrunde liegen, zulassen, werden in der zweiten Schicht die den Änderungen zugrunde liegenden Pläne mittels Analysetechniken ermittelt. Das erfolgt vorzugsweise in zwei Phasen. In der ersten Phase werden die Clusterstrukturen auf symbolischer Ebene betrachtet und eine lange Trainingsphase begonnen, um die in den symbolischen Darstellungen verborgenen Pläne zu erkennen. Im zweiten

Schritt wird dann die Gültigkeit der erkannten Pläne überwacht und eine Anomalie als unerwarteter und nicht plankonformer Übergang in der Clusterbildung definiert, der nicht aufgrund des jeweiligen Plan- oder Clusterwissens mithilfe der verwendeten Messgrößen zur Quantifizierung von Entfernungen zwischen vorangegangenen und gegenwärtigen Zuständen oder Plänen zu erklären ist. Dabei sind zunächst voreingestellte Schwellwerte zur Bewertung herausgefilterter Anomalien zu verwenden, die in Abhängigkeit von der Erkennungsgenauigkeit echter Angriffe bzw. der Häufigkeit falscher Alarme automatisch angepasst werden.

[0044] Die Überwachung der Gültigkeit der erkannten Pläne und die Definition einer Anomalie als Übergang in der Clusterbildung erfolgt dabei in Form eines Kompromisses zwischen Genauigkeit und Geschwindigkeit. Ein Schlüsselaspekt für eine funktionierende Lösung der zu beachten ist, beruht darauf, dass die gewählten Werkzeuge, wie z. B. selbstorganisierende Karten, über eine inhärente Flexibilität verfügen, die der beobachteten realen Welt bzw. den beobachteten Netzwerkkomponenten angepasst werden muss. Bei geringer Granularität der Netzwerkmessungen oder Clusterzahlen ist die Aussagekraft des verwendeten Modells möglicherweise gering. Dies lässt sich beispielsweise anhand steigender DoS-Angriffe in einem Netzwerksegment veranschaulichen. Der Einfluss eines DoS-Angriffes auf ein Netzwerksegment ist anfangs schwach und mit groben Modellen, die auf einer geringen Anzahl von Netzwerkmessungen und einer geringen Anzahl von nachgewiesenen Clustern basieren, nicht darstellbar. Innerhalb eines kurzen Zeitraumes, gemessen vom Beginn des DoS-Angriffes, lassen sich Abweichungen, die auf den DoS-Angriff schließen lassen, nur bei feingranularen Modellen erkennen, die auf einer wesentlich größeren Anzahl von Netzwerkmessungen und einer größeren Anzahl von nachgewiesenen Clustern basieren. Andererseits steigt mit dem gewählten Genauigkeitsgrad der Cluster- oder Planmodelle der Umfang der benötigten Rechenressourcen, die zur Erfüllung der Echtzeitanforderungen benötigt werden.

[0045] Es ist also ein Kompromiss zwischen Erfassungsgenauigkeit und Rechenzeit zu finden.

[0046] Fig. 3 stellt anhand eines Diagramms einen Kompromiss zwischen Erfassungsgenauigkeit und Rechenzeit graphisch dar.

[0047] Die Entwicklung eines Arbeitsmodells, das in der Lage ist, DoS-Angriffe in einem kommerziellen IP- gestützten Netzwerk zu erkennen, erfordert bei der Verwendung von Algorithmen, wie beispielsweise SOM-Berechnungen und Algorithmen zur Planerkennung, ein genaues Verständnis von Leistungsanfragen an das Netzwerk. Aus der Bottom-up-Sicht können öffentlich zugängliche Testdaten aus bereits referenzierten Projekten verwendet werden, um die Geschwindigkeit bestehender Anwendungen zu überprüfen. Bei einem Top-down-Ansatz werden dagegen, ausgehend von einem Referenznetzwerk, das als Modell für Teile eines Kommunikations-Netzwerkes dienen kann, mithilfe der theoretisch gültigen Skalierbarkeitsgesichtspunkte Anforderungen an Rechenleistung an eine kommerzielle Anwendung abgeleitet, um die erforderliche Mindestleistung zu bestimmen.

[0048] Entscheidend ist, dass DoS-Angriffe und anderer Netzwerkprobleme mit der Erfindung nur dann erfolgreich erkannt werden können, wenn eine gewisse Mindestgenauigkeit eingehalten wird.

Patentansprüche

1. Verfahren zur automatischen Erkennung von Anomalien in Weitverkehrsnetzen (WAN) und lokalen Netzen (LAN) und insbesondere von Denial-of-Service Angriffen unter Zugrundelegung des verhaltensbasierten Lösungsansatzes,
gekennzeichnet durch die Schritte
– automtisches Erfassen von Netzwerkmessdaten im Rahmen eines vorgegebenen Zeitregimes, die geeignet sind, Netzwerkaktivitäten so abzubilden, dass sie die reale Welt im Netzwerk wiederspiegeln,
– komprimieren der Netzwerkmessdaten mittels Techniken des unüberwachten Lernens (Dimensionsreduktion),
– automatisches ermitteln von Clusterstrukturen/Kategorien aus den kontinuierlich erfassten komprimierten Netzwerkmessdaten mittels Methoden des unüberwachten Lernens während eines ständigen kontinuierlichen Trainings- und Lernprozesses,
– automatisches ermitteln der den Clusterstrukturen/Kategorien zugrunde liegenden Pläne mittels Algorithmen zur Planerkennung während eines ständigen kontinuierlichen Trainings- und Lernprozesses,
– automatisches überwachen der Gültigkeit der erkannten Pläne und Clusterstrukturen/Kategorien sowie Herausfiltern von Anomalien im Sinne eines unerwarteten und nicht plankonformen Übergangs in mindestens einer Clusterstruktur/Kategorie, der sich nicht dem bisher ermittelten Plan- oder Clusterwissen zuordnen lässt,
– automatisches bewerten der herausgefilterten Anomalien anhand voreingestellter Schwellwerte, die in Abhängigkeit von der Erkennungsgenauigkeit echter Angriffe bzw. der Häufigkeit falscher Alarme automatisch angepasst werden.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass die Netzwerkmessdaten
– Konnektivität (connectivity),
– einfache Verzögerung (one-way-delay) und Umlaufverlust (one-way-loss),
– Netzlaufzeit (round-trip-delay),

– Laufzeitvariation (delay variation),
– Verlustmuster (loss patterns),
– Umordnung der Pakete (packet reordering),
– Massentransportkapazität (bulk transport capacity) und
– Link-Bandbreitenkapazität (link bandwidth capacity)
für den ständigen kontinuierlichen Trainingsprozess erfasst werden.

Es folgen 2 Blatt Zeichnungen

Anhängende Zeichnungen

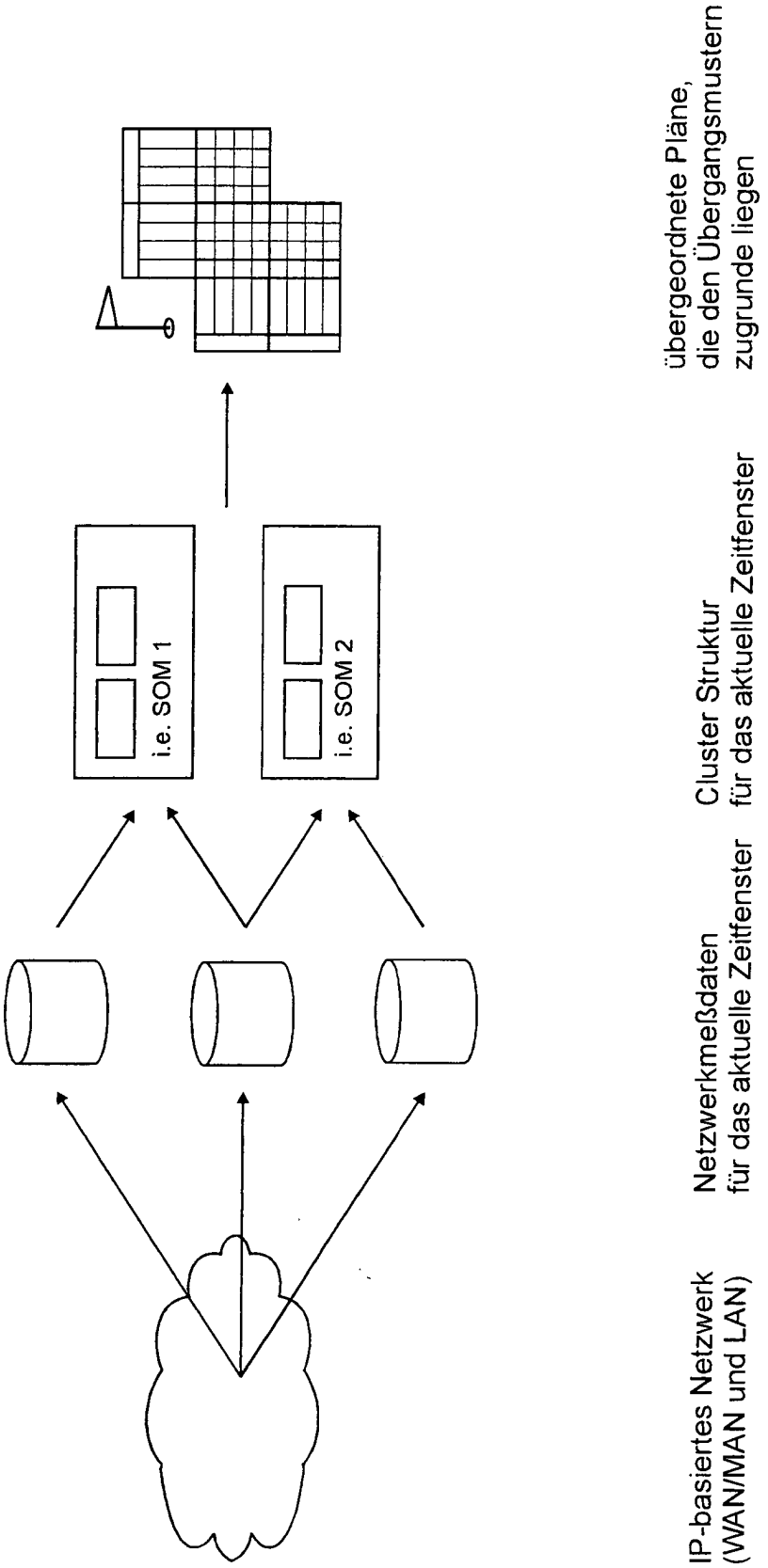
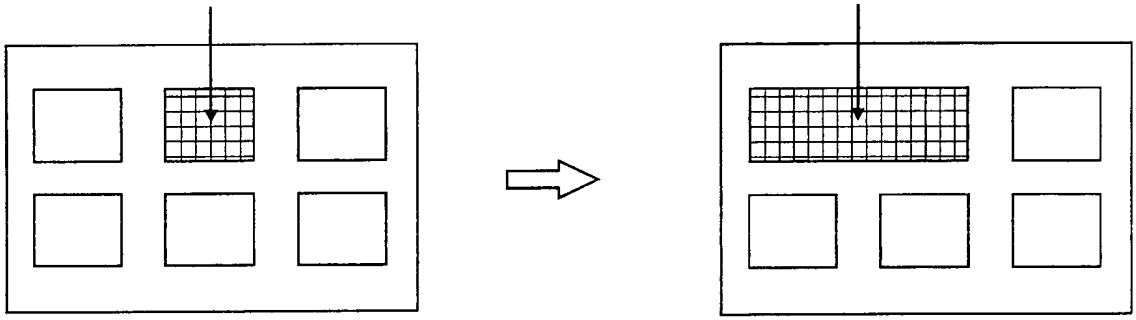


Fig. 1

Symbolische Darstellung
eines zusammenhängenden
Datenclusters

Symbolische Darstellung
zweier verschmolzener
Datencluster



SOM gültig für das aktuelle Zeitfenster
(keine wesentl. Änderungen)

veränderte SOM für ein
nachfolgendes Zeitfenster

Fig. 2

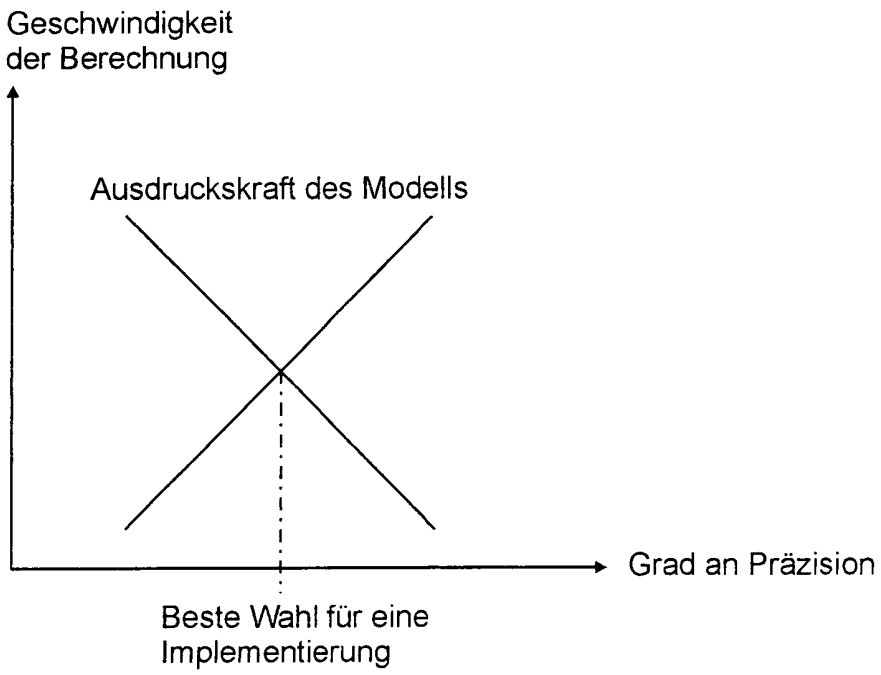


Fig. 3