# Die globale Sicht auf das Internet

*Artikel IAS*

Stand: 2008/04/08 -12-

Dominique Petersen <dominique.petersen@internet-sicherheit.de>
Prof. Dr. Norbert Pohlmann <norbert.pohlmann@informatik.fh-gelsenkirchen.de>
Institut für Internet-Sicherheit, http://www.internet-sicherheit.de
Fachhochschule Gelsenkirchen
Neidenburger Str. 43, D - 45877 Gelsenkirchen

## Inhaltsverzeichnis

# Einführung

Jeder von uns hat die Situation schon erlebt: Sie stehen im Stau und alles, was Sie sehen, ist eine Autokolonne vor und hinter Ihnen. In dieser Situation fehlt, ohne Hilfsmittel, die richtige Sicht auf das Problem. Es gibt keine direkten Informationen darüber, warum der Stau entstanden ist, wie lang er ist, an welcher Stelle man sich selber innerhalb des Staus befindet oder die wichtigste Information, wann der Stau wieder vorbei ist. Da dies ein Problem ist, dem sich tagtäglich Tausende von Autofahrern gegenüber sehen, wurden Lösungen entwickelt, um den Informationsmangel zu beseitigen. In Deutschland existiert ein engmaschiges Netz von Zählschleifen, die das Verkehrsaufkommen und die Verkehrssituation auf den Autobahnen erfassen. Per Radiodurchsagen werden grundlegende Informationen über Staus verbreitet, es stehen Stauinformationen per SMS, Telefon und Internet bereit, moderne Navigationssysteme verarbeiten die Stauinformationen direkt bei der Routenplanung. Durch diese Hilfsmittel werden die Autofahrer von der lokalen, eingeengten Sicht befreit und können aufgrund von globalen Informationen frühzeitig Entscheidungen treffen, wie an der nächsten Ausfahrt abzufahren und einen alternativen Weg zu nutzen.

Diese Situation lässt sich auf die Sichtweise, die Netzbetreiber heutzutage auf ihr eigenes Netz und das Internet haben, übertragen. In der Regel gibt es nur eine lokale Sicht, also einen Überblick über die eigenen Netzsegmente und die übertragenen Kommunikationsdaten. Treten hier Probleme auf, die erkannt werden, können diese schnell und gezielt behoben werden. Zeichnet es sich allerdings ab, dass ein Problem nicht innerhalb der eigenen Handlungsdomäne aufgetreten ist oder die notwendige Sichtweise fehlt, wird es schwierig. Es ist oft nicht klar, woher das Problem kommt, bei deren Beseitigung wir auf Dritte angewiesen sind.

Es fehlt eine globale Sicht, um das Problem zu erkennen und die richtigen Lösungen auszuwählen. Eine solche globale Sichtweise ist im Internet deshalb schwer zu erreichen, da sich niemand gerne in die Karten schauen lässt. Die genaue interne Netzstruktur, Kommunikationsverbindungen und Topologien werden oft von den Netzbetreibern vertraulich behandelt.

Zusätzlich müssen, um eine globale Sichtweise zu erreichen, einige Herausforderungen gelöst werden: Kommunikationsdaten sind im Prinzip datenschutzrelevant, die Datenmengen sind

enorm, die Datenraten sind zum Teil so groß, dass sie nicht immer „live“ analysiert werden können, und eine längerfristige Speicherung der Kommunikationsdaten, um langfristige Entwicklungen zu beobachten, erscheint unmöglich. Außerdem stellt sich die Frage, wer sich für die Erstellung einer globalen Sichtweise verantwortlich fühlt.

Dennoch hat sich das Internet in den letzten Jahren zu einem allgegenwärtigen Medium entwickelt, das aus sehr großen Bereichen der Wirtschaft, der Forschung und dem Privatleben nicht mehr wegzudenken ist. Deshalb ist die Analyse und Kenntnis des Mediums Internet in seiner Gesamtheit von besonderer Bedeutung, um dessen Entwicklungen bewerten und die zukünftige Funktionsweise aller enthaltener Dienste gewährleisten zu können.

Die stetig steigende Bedeutung des Internets für unsere vernetzte Wissens- und Informationsgesellschaft macht es notwendig, den Status über die Grenzen der einzelnen Netzbetreiber hinweg zu analysieren und zu kennen. Erst die genaue Kenntnis des Normalzustands macht es möglich, Anomalien zu erkennen, die die Funktionalität des Internets beeinflussen.

Mit Hilfe des Sonden-basierten Internet-Analyse-Systems, welches bereits seit einigen Jahren als Forschungs- und Entwicklungsprojekt des Instituts für Internet-Sicherheit an der Fachhochschule Gelsenkirchen in Zusammenarbeit mit dem Bundesamt für Sicherheit in der Informationstechnik (BSI) realisiert wird, sollen lokale und vor allem globale Sichtweisen erstellt und analysiert werden, um Frühwarnungen generieren zu können.

Besondere Schwerpunkte des Projektes sind eine datenschutzverträgliche Sammlung von Informationen, sowie eine Optimierung der Informationsmenge, um auch längerfristig Informationen speichern zu können und somit die Analyse von Trends und Entwicklungen über große Zeiträume zu ermöglichen.

## Ziele und Aufgabe des Internet-Analyse-Systems

Die Aufgabe des Internet-Analyse-Systems ist zum einen die Analyse von lokalen Kommunikationsdaten in definierten Teilnetzen des Internets und zum anderen die Erstellung

einer globalen Sichtweise auf die Teilnetze oder das Internet durch die Zusammenführung der vielen lokalen Sichten durch den Netzbetreiber selber oder einer unabhängigen Partei, wie dem Institut für Internet-Sicherheit.

Die Funktionen des Internet-Analyse-Systems lassen sich in die vier Teilbereiche Aufbau der Wissensbasis (Musterbildung), Beschreibung des Ist-Zustandes, Alarmierung und Prognostizierung unterteilen (siehe Abbildung 1).

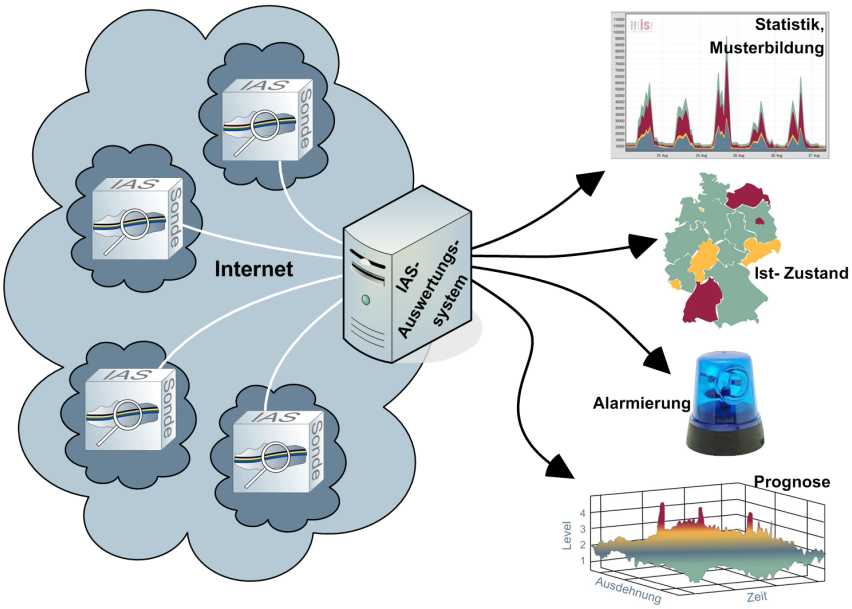


Abbildung 1: Aufgaben des Internet-Analyse-Systems

Hauptaufgabe der Musterbildung ist eine umfassende Analyse und Interpretation der Kommunikationsparameter des Internetverkehrs, mit dem Ziel, Technologietrends, Zusammenhänge und Muster zu erkennen, die unterschiedliche Zustände und Sichtweisen des Internets darstellen. Auf der Grundlage dieser Wissensbasis werden bei aktuellen Messwerten Anomalien gesucht und die Ursachen für die Zustandsänderungen analysiert und interpretiert. Dabei ist es wichtig herauszufinden, ob die Zustandsanomalien natürlichen Ursprungs sind, wie beispielsweise durch eine Technologieveränderung, oder ob ein mutwilliger Angriff zu Grunde liegt. Falls eine mutwillige Attacke vorliegt, werden die Muster identifiziert, die den Angriff charakterisieren.

Über exakte Kenntnis des aktuellen Zustands einer Kommunikationsleitung und Zuhilfenahme historischer, das heißt zuvor erfasster Informationen (Wissensbasis), kann bei signifikanten Änderungen des Verkehrsaufkommens oder der Kommunikationsdaten eine Warnmeldung generiert werden, aufgrund derer Maßnahmen zum Schutz und Erhalt der Funktionsfähigkeit des Internets ergriffen werden können.

Eine weitere wichtige Funktion ist die visuelle Darstellung des Internet-Zustands, analog zu einer Wetter-, oder Staukarte (siehe Abbildung 2).

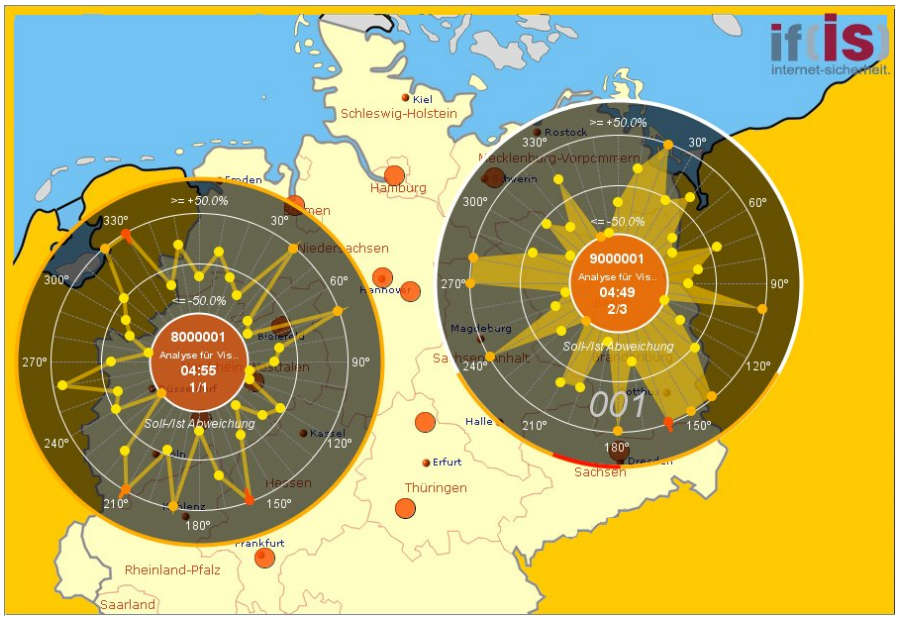

Abbildung 2: Visualisierung vom Zustand des Internets

Dabei steht es im Vordergrund, dass dringliche Entscheidungen – insbesondere bei Gefahren – einfacher und schneller als bisher getroffen und komplexe Sachverhalte gegenüber Dritten verdeutlicht werden können. Hierbei wird nicht nur bei Gefahren gewarnt, sondern auch der positive Zustand dargestellt, wenn die überwachten Netze in Ordnung sind. Neben der vereinfachten Sichtweise auf komplexe Strukturen soll das Visualisierungssystem unterstützend wirken, um Ursachen wie Spam-Attacken oder böswillige Malware-Angriffe frühzeitig zu signalisieren. Mithilfe flexibler Zustandsindikatoren lassen sich daraufhin Präventivmaßnahmen schaffen und Risiken gezielter minimieren.

Durch die Untersuchung und Analyse der extrapolierten Profile, Technologietrends,

Zusammenhänge und Muster ist es durch einen Evolutionsprozess der gewonnenen Ergebnisse möglich sein, Prognosen über Zustandsänderungen des Internets zu treffen (z.B mit Hilfe von neuronalen Netzen). Auf diese Weise können Angriffe und wichtige Veränderungen bereits frühzeitig erkannt und die Schadenswirkung und Kapazitätsengpässe prognostiziert werden [2].

## Funktionsweise des Internet-Analyse-Systems

Das Internet-Analyse-System besteht aus Sonden, die den Netzwerkverkehr an Kommunikationsleitungen unterschiedlicher Netze passiv abgreifen und Kommunikationsparameter auf verschiedenen Kommunikationsebenen zählen. In einem Auswertungssystem werden die Kommunikationsparameter unter verschiedenen Gesichtspunkten ausgewertet und übersichtlich dargestellt. Die Abbildung 3 zeigt die Zusammenhänge zwischen den am Internet-Analyse-System beteiligten Komponenten auf.

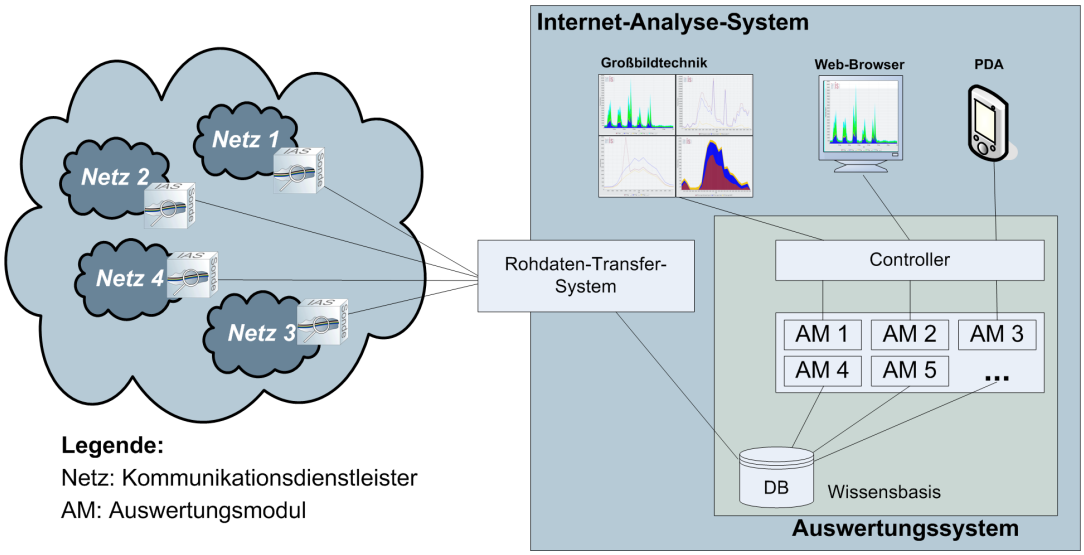


Abbildung 3: Komponenten des Internet-Analyse-Systems

Damit das Internet-Analyse-System aussagekräftige Ergebnisse liefern kann, benötigt es eine möglichst große Menge an Rohdaten, d.h. sehr viele Zähler von unterschiedlichen Kommunikationsparametern auf allen Kommunikationsebenen über die Zeit. Auf diesen Rohdaten gründen sich alle Analysen, die das Auswertungssystem durchführt. Sie bestehen aus aggregierten Zählerständen, die dem System von diversen Netzen zur Verfügung gestellt werden. Auf der linken Seite der Abbildung 3 ist das Internet dargestellt, das aus einem

Zusammenschluss zahlreicher Netze besteht. In jedem dieser Netze übernehmen Telekommunikations-Dienstleister (ISPs, Unternehmen, Hochschulen, …) z.B. die Aufgaben, einen Internetzugang für Endnutzer, Inhalte oder sonstige Dienste bereit zu stellen. Das Internet-Analyse-System bezieht seine Rohdaten von Sonden, die von den Netzbetreibern eingesetzt und betrieben werden. Übermittelt werden die Rohdaten über das eigens dafür spezifizierte Secure Raw Data Transfer Protocol (RDTPs), welches hochwertig verschlüsselt und signiert (4096 Bit RSA).

Die Sonden können die Rohdaten an ein oder mehrere Auswertungssysteme senden. Jede Organisation ist in der Lage, die Kommunikation mit dem Internet mit seinem Auswertungssystem eigene Analysen durchzuführen. Um eine globale und repräsentative Sichtweise des Internets zu erreichen, müssen Sonden in unterschiedlichen Netztypen, wie Global Tier One Provider, Transit Provider, Eyeball Internet Service Provider (DSL Provider), Content Provider und Business Networks sowie unterschiedlichen Regionen betrieben werden (siehe auch in Internet-Deutschland [1]).

## Aufgaben und Arbeitsweise der Sonden

Aufgabe der Sonden ist es, aus einem Kommunikationsdatenstrom Informationen zu extrahieren, die Aufschluss über den Zustand und die Nutzung der Kommunikationsstrecke und des oder der dahinter liegenden Netze geben. Hierbei sollen alle Informationen erhalten bleiben, die nötig sind, um eine missbräuchliche Nutzung, eine Fehlkonfiguration, Trendentwicklungen oder eine Angriffssituation zu erkennen. Gleichermaßen soll aber die Menge der Informationen auf ein notwendiges Minimum beschränkt werden, damit die Informationen auch rückwirkend über längere Zeiträume betrachtet und analysiert werden können. Ein weiterer wichtiger Punkt für den Betrieb der Sonden ist, dass keine datenschutzrechtlich relevanten Informationen in den Extrakten der Sonden, den so genannten Rohdaten, enthalten sind.

Technisch wird die Netzwerkverbindung passiv abgegriffen und die Kommunikationsparameter der unterschiedlichen Protokolle auf den Kommunikationsebenen gezählt. Die Zählergebnisse werden in festgelegten Zeitintervallen an das Rohdaten-Transfer-System übertragen.

# Prinzip der Rohdatenerfassung

Abbildung 4 verdeutlicht das Prinzip der Rohdatenerfassung durch die Sonden.

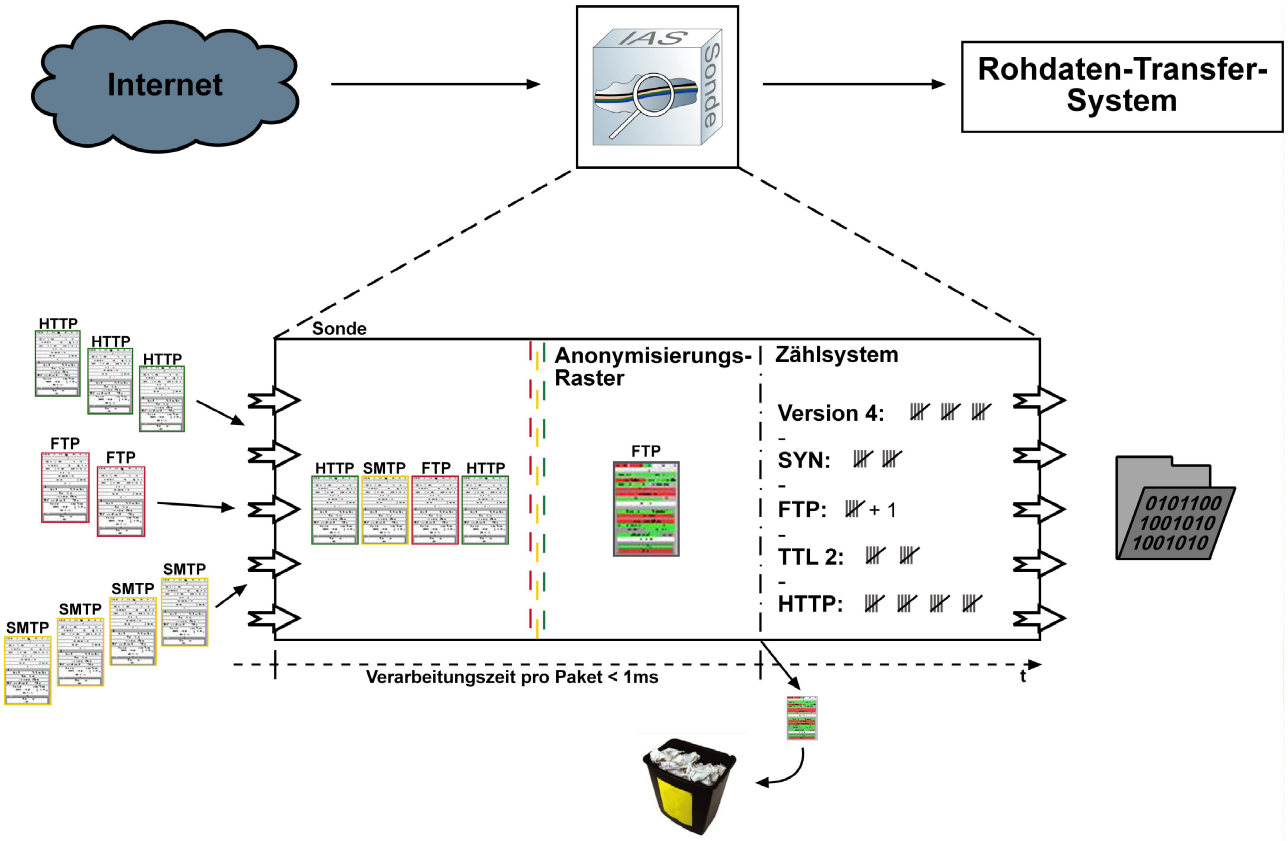


Abbildung 4: Prinzip der Rohdatenerfassung

Die Rohdatenerfassung gliedert sich in drei Teile. Links befindet sich schematisch das Internet. Es sind Pakete von drei unterschiedlichen Anwendungssitzungen dargestellt. Zusammengehörige HTTP-Pakete, eine FTP-Sitzung und eine SMTP-Sitzung. In der Mitte der Abbildung 4 befindet sich die Sonde. Die Pakete der drei Anwendungen werden in ihrer zufälligen Reihenfolge nacheinander von der Sonde passive abgriffen und ausgewertet. Das abgegriffene Paket wird durch mehrere Analyseklassen geschleust, die jeweils für ein bestimmtes Protokoll zuständig sind. Sie werten fest definierte Kommunikationsparameter in den Protokoll-Header auf den unterschiedlichen Kommunikationsebenen aus, die datenschutzrechtlich nicht relevant sind. IP-Adressen oder private Daten werden direkt verworfen. Je nachdem, wie die Headerinformationen des Paketes ausgefüllt sind, werden die im Zählsystem zugeordneten Zähler inkrementiert. Ähnlich wie bei einer Strichliste wird die Häufigkeit bestimmter

Headerinformationen festgehalten. Beispielsweise wird in der Abbildung 4 das Abgreifen des FTP-Paketes durch die Inkrementierung des FTP-Zählers um 1 festgehalten. Bei den Rohdaten handelt es sich also um Aggregate von Zählern, d.h. um Zähler von aufgetretenen Kommunikationsparametern auf den verschiedenen Kommunikationsebenen über einen definierten Zeitraum. Das Paket, in der Abbildung 4 ein FTP Paket, wird sofort nach der Kommunikationsparameter-Auswertung physikalisch, d.h. irreversibel und spurenlos, von der Sonde gelöscht [3].

| ID | Description | | Count |
|---|---|---|---|
| 131134 | IP (Protocol Number 6) | : | 18.854.151 |
| 131145 | IP (Protocol Number 17) | : | 1.123.149 |
| 327708 | TCP (Flags: SYN) | : | 334.435 |
| 327723 | TCP (Flags: FIN/ACK) | : | 480.697 |
| 327724 | TCP (Flags: SYN/ACK) | : | 275.779 |
| 545857 | HTTP (Request Method POST) | : | 2.026 |
| 545861 | HTTP (Request Method GET) | : | 293.616 |
| 545863 | HTTP (Request Method HEAD) | : | 18.992 |

Abbildung 5: Zählsystem der Sonde

Eine Wiederherstellung des Kontext eines Paketes oder auch nur eines Kommunikationsparameters ist weder möglich noch nötig. In definierbaren Zeitanständen werden die Zählerstände (Rohdaten) von den Sonden an das Rohdaten-Transfer-System übertragen werden. Hierbei handelt es sich ausschließlich um die vollständig anonymen Informationen, wie sie in Abbildung 5 zu sehen sind. Rechts hinter dem Doppelpunkt befinden sich die Zählerstände für die links spezifizierten Headerinformationen. Jede Zeile steht für einen Zähler. Auf der linken Seite des Doppelpunktes steht die Zähl-Bedingung (Auftreten des entsprechenden Kommunikationsparameters), rechts die Anzahl der Pakete, die den Kommunikationsparameter während des definierten Mess-Zeitraums enthalten haben. Zeile zwei der gezeigten Rohdaten beschreibt z.B., dass 1.123.149 Pakete mit der IP-Protokoll-Nummer 17 (UDP) in der gezählten Zeit aufgetreten sind. Die Zählbedingungen und deren Kodierungen sind in einer versionisierten XML-Datei spezifiziert.

Das Rohdaten-Transfer-System fungiert als Server, mit dem sich die Sonden verbinden können, um ihre Rohdaten für einen definierten Zeitraum zu übertragen. Es handelt sich hier um eine unidirektionale Verbindungsmöglichkeit. Das bedeutet, ein Verbindungsaufbau ist nur von der Sonden-Seite aus möglich. Eine Sonde kann die Rohdaten an ein oder mehrere Rohdaten-

Transfer-Systeme übertragen. Beispiel einer typischen Konfiguration ist, dass alle 5 Minuten die Rohdaten in einer Größe von ca. 50 KByte an das eigene und an ein zentrales Rohdaten-Transfer-System gesendet wird.

Da die Rohdaten nur eine statistische Formulierung der eigentlichen Kommunikationsdaten sind, würde es auch ausreichen, wenn nicht jedes Paket betrachtet wird, sondern z.B. nur jedes 10. Paket. Dieser Aspekt kann bei sehr hohen Kommunikationsdatenraten eine pragmatische Lösung sein, ohne statistisch gesehen ein anderes Ergebnis zu erhalten.

## Auswertung der gesammelten Rohdaten

Die eigentliche Auswertung und Verarbeitung der gesammelten Informationen finden in diversen Analysemodulen (AM) des Auswertesystems statt. In Abbildung 3 sind diese durch „AM1“ – „AM5“ gekennzeichnet. Die Module beziehen die Informationen ausschließlich aus der Wissensbasis (Rohdaten und Auswertungsergebnisse). Ziel der diversen Module ist das Erstellen von Profilen, Statistiken und Zusammenhängen sowie das Erkennen von Schwellwert-Überschreitungen und die grafische Aufbereitung der Rohdaten und Auswertungsergebnisse.

Da es sich bei den Rohdaten um vollständig anonyme Informationen handelt, könnten sie zusätzlich zwischen verschiedenen Netzbetreibern ausgetauscht oder an zentraler Stelle gesammelt werden, um sie als Basis für eine globale Sichtweise und Analysen eines IT-Frühwarnsystems verwenden zu können.

## Benutzungsoberfläche des Internet-Analyse-Systems

Es sind viele Arten denkbar, die Ergebnisse des Internet-Analyse-System anzuzeigen. In Abbildung 3 sind beispielhaft angeführt: eine Großbildtechnik, ein Web-Client, sowie ein PDA. Die Großbildtechnik dient der laufend aktualisierten Anzeige gewisser Statistiken, Profile und aktuelle Zustände. Über einen intelligenten Client können mit dem Auswertesystem weiterreichende Analysen durchgeführt und Ergebnisse protokolliert werden. Zusätzlich können z.B. über einen PDA Warnmeldungen des Systems mobil empfangen werden, um sich sofort einen ersten Überblick über die Gefahrenlage verschaffen zu können. Zurzeit dient dem Internet-Analyse-System sowohl zur Visualisierung als auch zur Administration ein Standalone-Client mit intuitivem Frontend. Über diesen Client können alle Zählerstände zu frei wählbaren Zeiträumen graphisch visualisiert werden. Um der Menge an Kommunikationsparametern Herr zu werden, existiert ein Favoritensystem, in welchem wichtige Parameter zur schnellen Übersicht oder zur detailierten Analyse zur Verfügung stehen. Darüber hinaus kann man bequem individuelle Reporte sowohl manuell generieren, als auch automatisch per E-Mal zyklisch zuschicken zu lassen, um einen stetigen Überblick über besondere Kommunikationsparameter zu haben. Es ist ebenfalls leicht möglich, über ein Plugin-System weitere Funktionalitäten hinzuzufügen.

## Ergebnisse des Internet-Analyse-Systems

In diesem Abschnitt werden zur Veranschaulichung einige Ergebnisse dargestellt, um eine Idee der Möglichkeiten des heutigen Zustands des Internet-Analyse-System zu vermitteln. Es werden zurzeit ca. 1.200.000 unterschiedliche Zähler von Kommunikationsparametern für die verschiedenen Kommunikationsebenen und -protokollen berücksichtigt. Diese große Zahl macht deutlich, wie komplex die Ergebnisse sein können. Hierbei existieren Zählsysteme für aktuelle Technologien, wie Ethernet, IP, ICMP, TCP und UDP, HTTP, SMTP, DNS, SIP und viele weitere Kommunikationsprotokolle.

## *Transportprotokoll-Verteilung*

In der Abbildung 6 wird die Verteilung der genutzten Protokolle der Transportschicht im Fachbereich Informatik der FH Gelsenkirchen über einen Zeitraum von mehreren Tagen dargestellt.

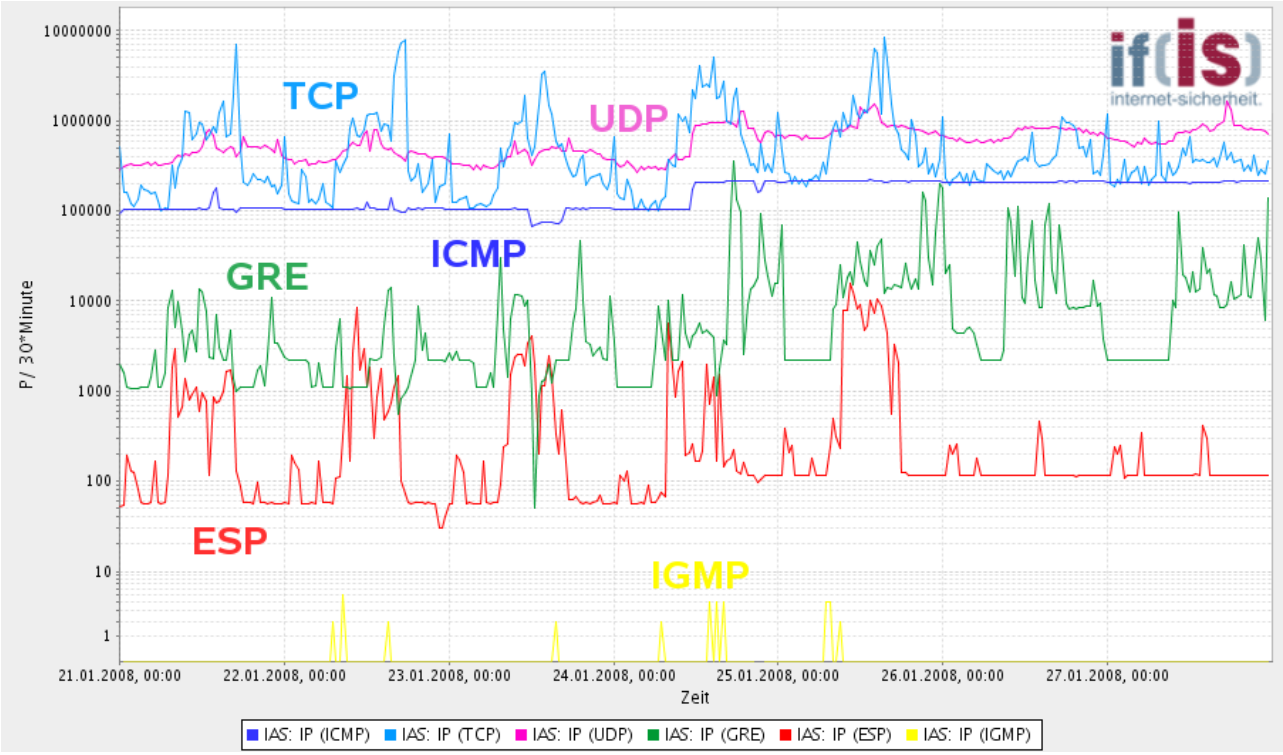


Abbildung 6: Protokolle der Transportschicht FB Informatik

Aus der Abbildung 6 kann sehr leicht erkannt werden, dass eine Anomalieerkennung erfolgen kann. Das Internet-Analyse-System kennt aus der Vergangenheit das Profil, die Standardabweichung und kann daraus sehr einfach einen Hinweis über untypisches Verhalten anzeigen.

Zusätzlich kann die Nutzung bestimmter Protokolle ermittelt werden. Hieraus können Kapazitätsplanungen, beispielsweise für den Einsatz von Virtual-Private-Networks (ESP-Protokoll), erfolgen.

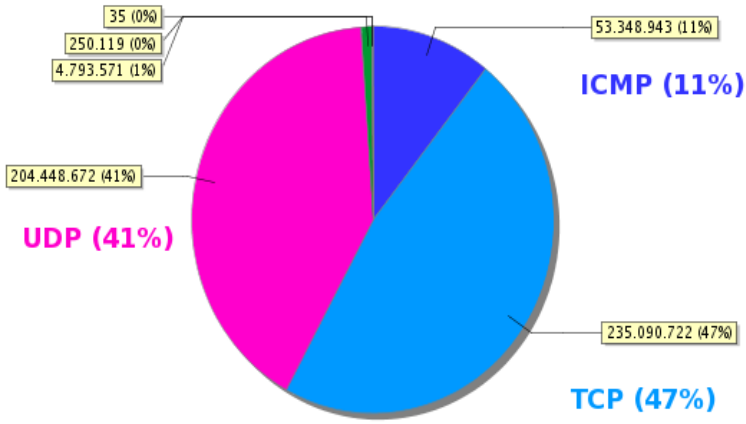


Abbildung 7: IP-Protokollverteilung FB Informatik

Auch Abhängigkeiten von Protokollen lassen sich erkennen. So scheint UDP proportional zu TCP zu sein, was auf die Abhängigkeiten von Anwendungsprotokollen (wie HTTP und SMTP) und der damit verbundenen Namensauflösung (DNS) zurückgeführt werden kann (in Abbildung 6 nicht mehr zu erkennen). Durch die stetig zunehmende Benutzung von Realzeit-Anwendungen, wie VoIP und Video-Streaming, und der Einbindung von entfernten Computer in das private Netz der FH Gelsenkirchen mittels Virtual Private Network (VPN) ist UDP mittlerweile im prozentualen Gebrauch mit TCP nahezu gleich auf (siehe Abbildung 7). Ohne VPN und den Realzeit-Anwendungen beträgt das Verhältnis von TCP zu UDP im Normalfall 89% zu 7%.

## *Browserverteilung (Technologie-Trend)*

In der Abbildung 8 ist die Verteilung von unterschiedlichen Browsern im Fachbereich Informatik der FH Gelsenkirchen über einen Zeitraum von einem Tag dargestellt.

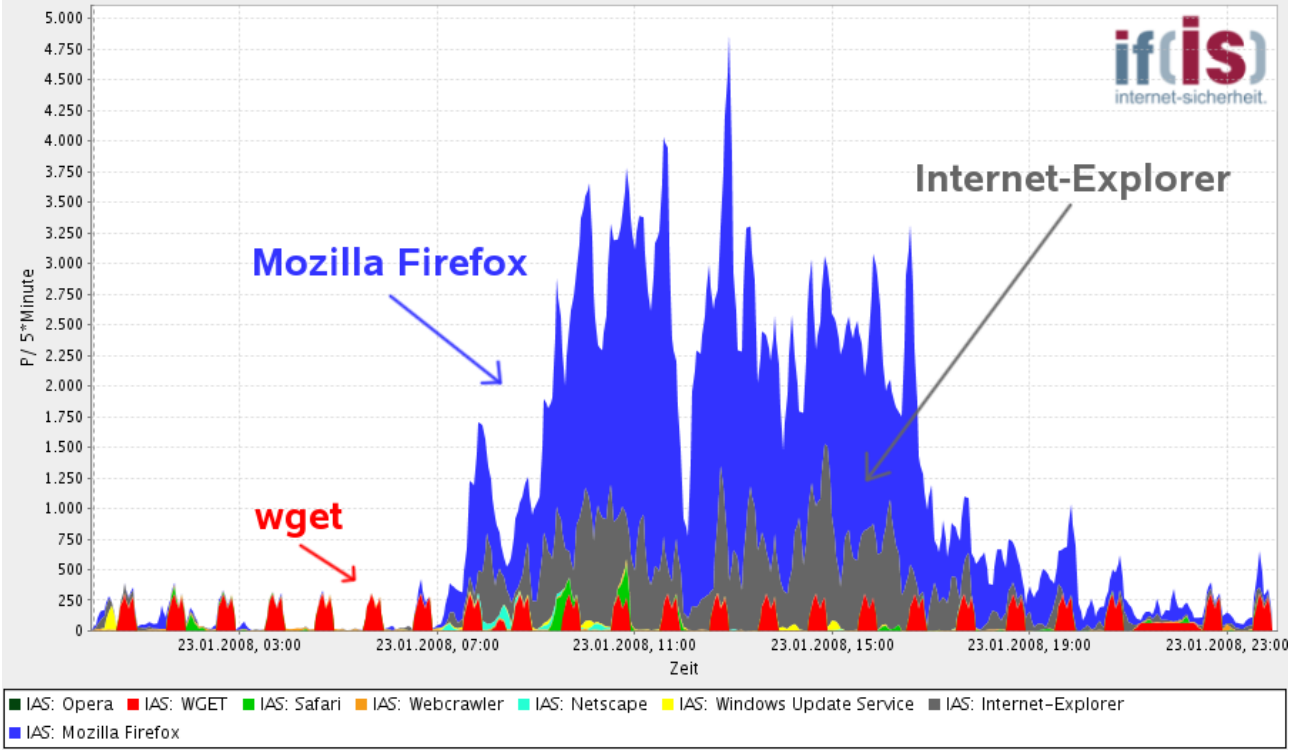


Abbildung 8: Browserverteilung über die Zeit FB Informatik

Hier kann man das Tagesprofil der unterschiedlichen Browser erkennen. Deutlich sind die Unterschiede zwischen manueller Nutzung (z.B. Firefox) und automatischer Nutzung (z.B. wget) über den Tagesverlauf verteilt zu erkennen.

Bemerkenswert ist, dass es sich keineswegs um Statistiken von HTTP-Anfragen handelt, die auf einen Webserver bezogen sind, wie heute üblich. Es handelt sich um Statistiken, die sich direkt auf eine Kommunikationsverbindung beziehen.

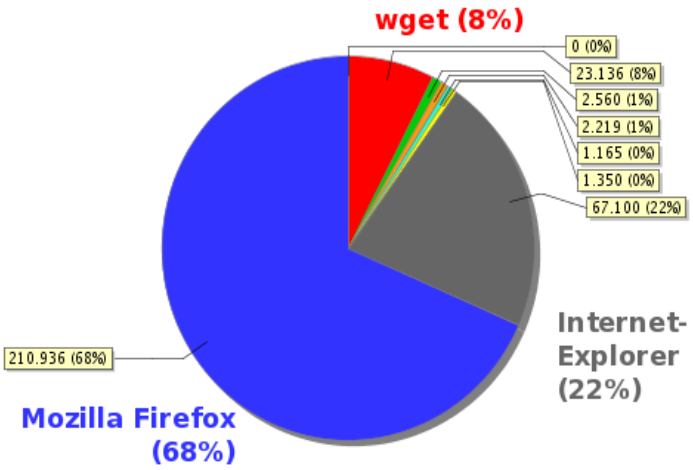


Abbildung 9: Browserverteilung FB Informatik

Bei der Abbildung 9 handelt es sich um die Browserverteilung im Fachbereich Informatik, was den hohen Mozilla Firefox-Anteil (68%) erklärt. Bei anderen Sondenpartnern, die allgemein ausgerichtet sind, beträgt das Verhältnis vom Internet Explorer zum Mozilla Firefox 72% zu 21%.

## *BKA-Wurm (Januar 2007)*

Ende Januar und Anfang Februar 2007 verbreitete sich der BKA-Wurm erneut stark. Der BKA-Wurm ist per E-Mail-Anhang verschickt worden und trat in Wellen auf. Abbildung 10 zeigt eine dafür charakteristische Verteilung.

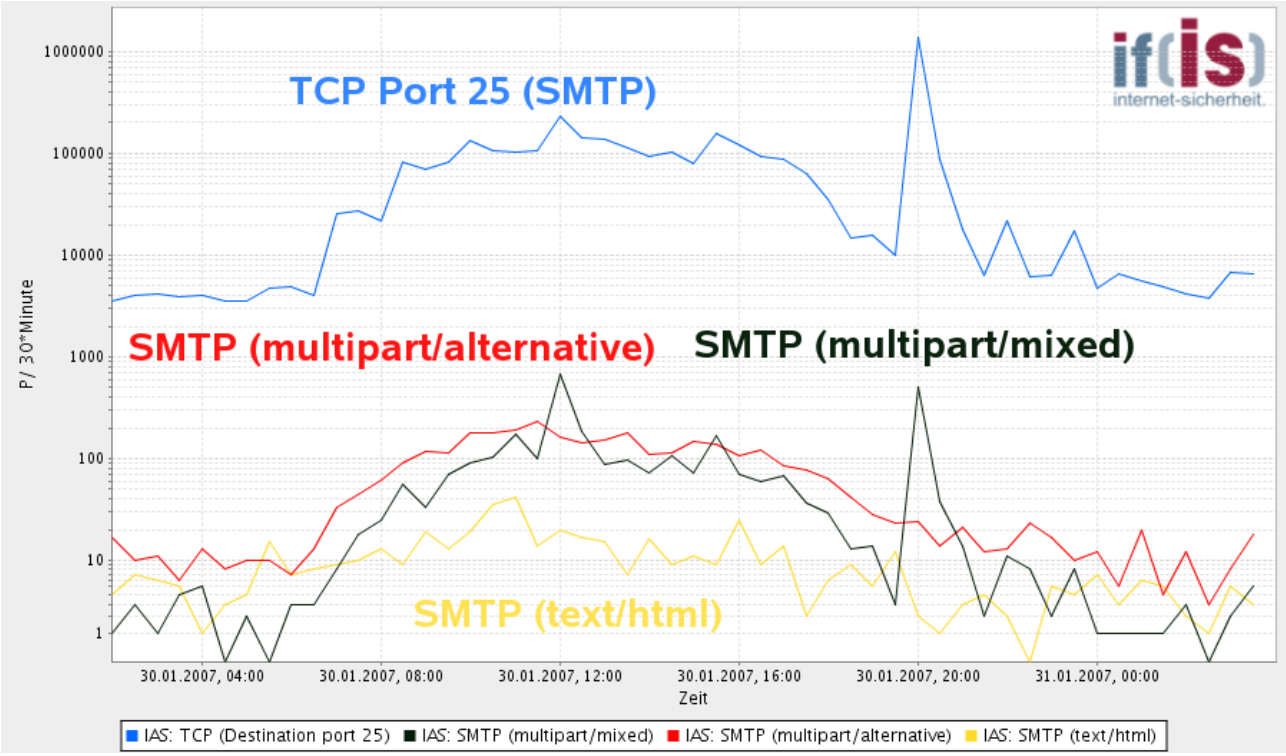


Abbildung 10: BKA-Wurm Verbreitung am 30.01.2007 um 20:00

Der normale Mail-Verkehr fängt mit dem Arbeitsbeginn gegen 7:30 an, zeigt gegen Mittag das stärkste Aufkommen und flacht zum Nachmittag hin langsam ab. Durch die Analyse des Quelltextes einer solchen gefälschten BKA-Mail und Vergleich mit bestehenden Kommunikationsparametern des Internet-Analyse-Systems wurden anhand des Deskriptors „SMTP (Multipart/Mixed)“ in Verbindung mit den Anfragen auf „TCP Destination Port 25“ (SMTP) die BKA-Wurm-Wellen erkannt. Kennzeichnend für diese Mails sind zudem das Ausbleiben anderer typischer Mail-Deskriptoren, wie „SMTP (Multipart/Alternative)“ und „SMTP (Multipart/Text/HTML)“. Bekannt ist, dass die BKA-Wellen in dieser Zeit hauptsächlich konzentriert um 15:00 und um 20:00 losgeschickt wurden, was in der Abbildung 10 am 30.01.2007 um 20 Uhr gut zu erkennen ist. Weitere Wellen konnten im Januar und Februar so erfolgreich identifiziert werden.

## Nutzen des Internet-Analyse-Systems

Sondenbetreiber mit eigenem Auswertungssystem werden bei auftretenden Problemen vom Internet-Analyse-System informiert und können die umfangreichen Hilfsmittel des Systems bei den Analysen von Problem nutzen. Eine übersichtliche Darstellung des aktuellen Zustandes des eigenen Netzes an ausgewählten Knotenpunkten und ein aufschlussreiches Reportingsystem stellen die Basis für einen zuverlässigen Netzbetrieb dar.

Ist das eigene System Bestandteil eines zentralen Auswertungssystems, werden Probleme bei einem Teilnehmer rechtzeitig weitergegeben, so dass alle frühzeitig Gegenmaßnahmen ergreifen können. Des Weiteren kann die eigene lokale Sicht mit einer globalen verglichen werden, um eigene Probleme schnell zu identifizieren oder auf drohende Gefahren, die bereits bei den Partnern aufgetreten sind, präventiv reagieren zu können.

## Ausblick

Mit Hilfe des Sonden-basierten Internet-Analyse-Systems können kontinuierlich Rohdaten gewonnen werden, die den Internetverkehr statistisch widerspiegeln. Durch die Auswertung der Rohdaten auf den unterschiedlichen Kommunikationsebenen wie Vermittlungsebene, Transportebene und Anwendungsebene können weit reichende Informationen abgeleitet werden.

Durch die Analyse der Ergebnisse unterschiedlicher Sonden ist es möglich, eine globale Sichtweise des Internets darzustellen und Warnstufen im Fall von Problemen, wie infrastrukturellen Ausfällen oder Angriffen, zu definieren. Weitere Analysen der Rohdaten erlauben die Prognosen von Trends in der Benutzung von Protokollen, Netzwerkdiensten und Angriffen.

Ein weiterer wichtiger Prozess ist die genaue Analyse der Rohdaten, um die Ergebnisse mit den Informationen anderer Netzanalyse-Tools zu korrelieren. Mit den hergestellten Zusammenhängen können Muster erkannt werden und somit Ergebnisse in Form von Statistiken, Mustern, Angriffsszenarien und Prognosen erstellt werden. Diese Ergebnisse können wiederum in den Sichtungs- und Auswertungsprozess einfließen, um so immer früher immer genauere

Ergebnisse zu erlangen. Hierbei ist der Umfang der Rohdatenbasis von entscheidender Bedeutung: Je mehr Rohdaten analysiert werden können, sowohl in Bezug auf die Anzahl der unterschiedlichen Sonden und deren Positionierung, als auch in Bezug auf die Zeitspanne der Informationen, desto genauere Ergebnisse sind zu erwarten. Insbesondere gilt dies für die Vorhersageanalysen, da Algorithmen, die geeignet sind entsprechende Analysen vorzunehmen, eine besonders große Initialdatenmenge benötigen.

Das Institut für Internet-Sicherheit hat die dritte Entwicklungsphase erfolgreich abgeschlossen. Das System hat sich im Betrieb bewährt und ist in verschiedenen Institutionen und Ländern produktiv im Einsatz. Es werden weiterhin Partner gesucht, die sich dem Verbund anschließen und weitere Sonden und Auswertungssysteme betreiben, um eine noch genauere globale Sichtweise zu erzeugen. Organisationen, die nur eine Sonde betreiben wollen, erhalten vom zentralen Auswertungssystem z.B. wöchentlich oder monatlich einen aufschlussreichen Bericht über Ihren Netzwerkverkehr.

Neben dem Betrieb des Internet-Analyse-Systems wird das Projekt in der nächsten Entwicklungs- und Forschungsphase weiter vorangetrieben. Es werden weitere Protokolle realisiert und Bestehende erweitert, um zusätzliche Informationen zu erhalten. Außerdem werden die Auswertungsmodule und die Angriffserkennung ausgebaut.

Im Bereich der Forschung werden mit statistischen Methoden und Datamining- Algorithmen intelligente Verfahren optimiert, die den Evaluierungsprozess beschleunigen sollen. Durch die kontinuierliche Analyse und das Einfließen der gewonnenen Erfahrungen in den Analyseprozess steigt die Qualität der Ergebnisse stetig an.

Das Internet ist seiner Natur nach äußerst flexibel, komplex und täglichen Veränderungen unterworfen. Um Statistiken über das Internet auch für beliebige zukünftige Technologien zur Verfügung stellen zu können, muss das Internet-Analyse-System in seiner grundlegenden Form ebenfalls absolut flexibel sein.

Für ein umfangreiches IT-Frühwarnsystem können weitere Systeme aufgrund der innovativen Idee der anonymen Zählwerte leicht in das Internet-Analyse-System integriert werden werden.

Derzeit werden im Institut für Internet-Sicherheit weitere Komponenten, wie ein Verfügbarkeitssystem, in dem Drohnen aktiv in das Internet ausschwärmen und den Zustand messen, und einem Logdaten-Auswertungssystem, bei welchem Log-Dateien auf kritische Meldungen überwacht und Warnungen im Internet-Analyse-System visualisiert werden, entwickelt.

Es gilt also, dieses System flexibel und flächendeckend einzusetzen, damit genau wie im Stau auf der Autobahn eine globale Sicht auf die Strukturen und Gefahren des Internets möglich wird.

## Literatur

[1] S. Dierichs, N. Pohlmann: „Netz-Deutschland“, iX - Magazin für professionelle Informationstechnik, Heise-Verlag, 12/2005

[2] N. Pohlmann: “Internetstatistik”, Proceedings of CIP Europe 2005, Hrsg.: B.M. Hämmerli, S.D. Wolthusen; Gesellschaft für Informatik, Bonn 2005

[3] M. Proest: „Internet-Analyse - Ein Blick in die Dunkelheit“, Konferenz: Internet-Sicherheit 2005; http://www.internet-sicherheit.de/center-berichte.html

Weitere Informationen:
Institut für Internet-Sicherheit, http://www.internet-sicherheit.de