Prof. Dr. Gabi Dreo Rodosek
Lehrstuhl für Kommunikationssysteme und Netzsicherheit
Direktorin des Forschungszentrums CODE

Universität der Bundeswehr München
Universität der Bundeswehr München · 85577 Neubiberg · Germany

Telefon +49 89 6004-2826
Telefax +49 89 6004-3898
E-Mail Gabi.Dreo@unibw.de

Deutscher Bundestag
z.Hd. Herr Ministerialrat Harald Georgii

Dorotheenstr. 88
10117 Berlin

30.09.2016

Sachverständigengutachten, 1. Untersuchungsausschuss der 18. Wahlperiode
Geschäftszeichen PA 25-5452-09
Beweisbeschluss SV-13

Sehr geehrte Herr Ministerialrat Georgii,

ich darf Ihnen hiermit die schriftliche Ausarbeitung des Sachverständigengutachtens übermitteln.

Mit freundlichen Grüßen

Prof. Dr. Gabi Dreo Rodosek

Forschungszentrum Cyber Defence
Universität der Bundeswehr München

der Bundeswehr Universität München

Sachverständigengutachten

Beweisbeschluss SV-13
1. Untersuchungsausschuss der 18. Wahlperiode

Darstellung der technischen Gegebenheiten bei der paketvermittelten Übertragung von Telekommunikationsdaten auf der Ebene sogenannter „Autonomer Systeme" (AS), die in einem sogenannten globalen „Internet Exchange Point"(IXP) verbunden sind, einschließlich der technischen Hintergründe und der technischen Entwicklung der IP-Übertragungsverfahren sowie Darstellung der technischen Beschaffenheit der gemäß § 27 Abs. 2 TKÜV bzw. zur strategischen Überwachung von Ausland-Ausland-Telekommunikationsverfahren auszuleitenden Daten und der Möglichkeiten der regionalen Zuordnung dieser ausgeleiteten Kommunikationsdaten

Prof. Dr. Gabi Dreo Rodosek
Lehrstuhl für Kommunikationssysteme und Netzsicherheit
Direktorin des Forschungszentrums CODE
Universität der Bundeswehr München

30. September 2016

Zusammenfassung

Das Internet besteht aus einer stetig wachsenden Anzahl voneinander unabhängiger Autonomer Systeme (AS). In der Einleitung werden grundlegende Begriffe erläutert und auf den Verbund aus AS und den Internet Exchange Points (IXP) eingegangen. Das Deutsche Forschungsnetz wird dabei als beispielhaftes AS und der DE-CIX in Frankfurt als ein Vertreter der IXP aufgegriffen. Ferner wird auf die Wegefindung und -wahl im Internet eingegangen und in diesem Zusammenhang das Border Gateway Protocol (BGP) detailliert erläutert.

Es werden Möglichkeiten dargestellt, die eine Klassifizierung des Datenverkehrs an einem IXP erlauben. Insbesondere werden hierbei technische Gegebenheiten betrachtet und die Problemstellung hinsichtlich verschleierten und verschlüsselten Kommunikationsbeziehungen aufgezeigt. Darüber hinaus wird die Frage beantwortet, welche Verkehrsdaten für die Zusammensetzung von inhaltlich vollständigen Kommunikationsvorgängen erforderlich sind.

Hinsichtlich der Identifikation der Ursprungs- und Zielorte von Kommunikationsvorgänge werden wissenschaftlich anerkannte und praktisch erprobte Verfahren zur IP-Geolokalisation erläutert. Zusätzlich wird auf Aspekte und Indizien eingegangen, welche Rückschlüsse darauf zulassen, ob ein Kommunikationspartner Deutscher ist.

Abschließend werden im Gutachten Angriffsmöglichkeiten dargestellt, welche den Regelbetrieb des Internets beeinflussen können und anhand von Beispielen veranschaulicht.

Das vorliegende Sachverständigengutachten beantwortet die Fragen des 1. Untersuchungsausschusses der 18. Wahlperiode aus Sicht der technischen Gegebenheiten und Möglichkeiten. Es wird nicht auf die konkrete Umsetzung der technischen Möglichkeiten eingegangen, da dies in der Domäne der einzelnen Dienstanbieter liegt.

Inhaltsverzeichnis

Danksagung

Dieses Gutachten entstand in Zusammenarbeit mit Prof. Dr. Wolfgang Hommel, Dr. Robert Koch, Marcus Knüpfer, Sebastian Seeber und Lars Stiemert. Für ihre großzügige Unterstützung möchte ich mich sehr herzlich bedanken. Ebenfalls herzlich bedanken möchte ich mich bei der Geschäftststelle des Deutschen Forschungsnetzes, bei Herrn Jochem Pattloch und Dr. Christian Grimm, sowie Prof. Dr. Helmut Reiser vom Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften und der Ludwig-Maximilians Universität München für wertvolle Hinweise.

1 Einleitung

Bevor auf die detaillierte Beantwortung der Fragen, die im Beweisbeschluss SV-13 formuliert sind, eingegangen wird, ist festzuhalten, dass im Gutachten nur auf die technischen Gegebenheiten und Möglichkeiten eingegangen wird. Die tatsächliche Umsetzung der technischen Möglichkeiten liegt in der Domäne jedes einzelnen Dienstanbieters (Internet Service Providers) und ist nicht Gegenstand dieses Gutachtens.

Das Internet besteht aus einer stetig wachsenden Anzahl voneinander unabhängiger Netze, den sogenannten Autonomen Systemen (AS), welche miteinander verbunden sind. Es bestehen zwei Möglichkeiten einen Verbund von AS aufzubauen. So können Betreiber von AS bilaterale Abkommen mit anderen Betreibern schließen und AS mittels dedizierter Leitungen direkt miteinander verbinden. Die zweite Möglichkeit ist die Nutzung von Internet Exchange Points (IXP), an welchen eine Vielzahl an AS angebunden ist. In der Regel ist jedes AS direkt oder über ein bzw. mehrerer IXP mit anderen AS verbunden.

Ein AS besteht im Regelfall aus Teilnetzen, die über Router miteinander verbunden sind und unter einer einzigen administrativen Kontrolle bzw. Instanz stehen. Bevor auf weitere Details eingegangen wird, veranschaulicht Abb. 1 den Verbund mehrerer AS mit einem Internet Exchange Point (IXP) sowie die direkte Verbindung zweier AS.

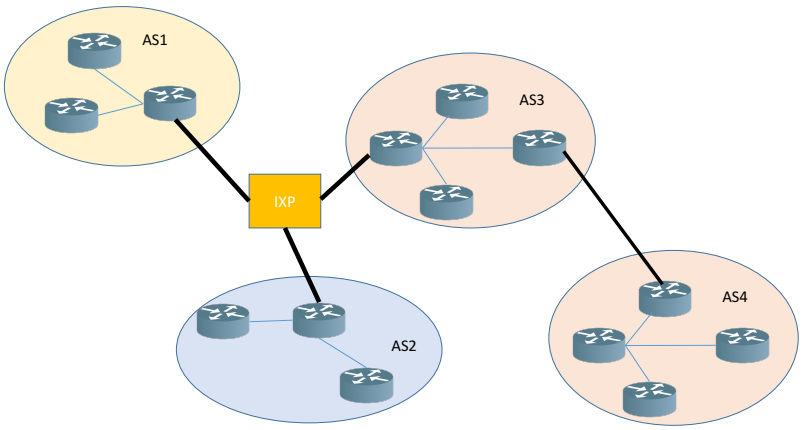


Abbildung 1: Mehrere AS in einem Verbund mit einem IXP

Die AS unterscheiden sich in Größe, räumlicher Ausdehnung und Relevanz. Je nach Größe der eigenen Netzinfrastruktur werden Dienstleister (Internet Service Provider, ISP) in Tier 1, Tier 2 und Tier 3 Kategorien unterschieden. Die Größe des ISP und somit des AS hat Auswirkungen auf das Peering- und Transit-Verhalten des Datenverkehrs, der zwischen den AS ausgetauscht wird. Der Unterscheid zwischen Peering und Transit ist, dass bei letzterem für den Transfer des Datenverkehr über eine Netzinfrastruktur ein Entgelt geleistet werden muss. Beim Peering ist das entsprechend nicht der Fall.

Tier 1 AS sind sehr große AS, die die gesamte Konnektivität im Internet ausschließlich durch das Peering mit anderen großen AS herstellen. Tier 2 AS werden auch als Transit-Provider bezeichnet. Die Konnektivität bei diesen beruht zusätzlich zum Peering auf Transit-Abkommen mit den Tier 1 AS. Die Gruppe der Tier 3 AS bilden kleinere lokale Provider, welche in der Regel keinen Transit-Verkehr für

andere AS anbieten.

Die Verbindung der AS über die IXP kann für Betreiber vorteilhaft sein, da lediglich die Verbindungen zum IXP betrieben werden müssen und somit Kosten gespart werden können. Andererseits können große Providern mit hohem Datenaufkommen den Einsatz von direkten Verbindungen zu anderen AS unter Berücksichtigung technischer und wirtschaftlicher Aspekte bevorzugen.

Insbesondere die Tier 1 AS sind mittels direkter Verbindungen miteinander verbunden, sodass ein relevanter Anteil des Internet-Datenverkehrs ohne Involvierung eines IXP ausgetauscht wird. Studien belegen, dass die Anzahl der direkten Verbindungen zwischen AS weiter steigt [1].

Tier 1 AS sind, nach dem Eintrag in der Peering DB vom Center for Applied Internet Data Analysis (CAIDA) [2], z.B. Level 3 Communications und Cogent Communications [3]. Auch die Deutsche Telekom ist ein Tier 1 Provider, wie anhand der CAIDA-Daten [3] in Abb. 2 visualisiert.

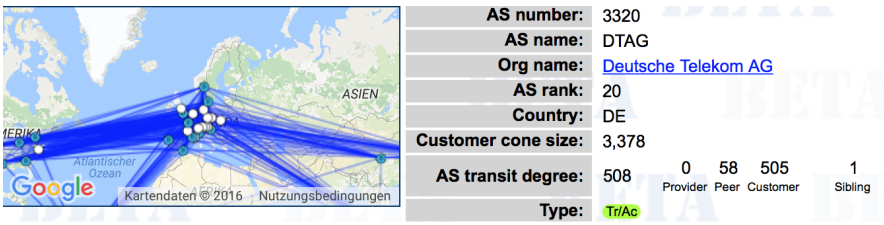


Abbildung 2: Deutsche Telekom, AS3320

Jedes AS ist durch eine eindeutige Nummer, die AS-Nummer (ASN), gekennzeichnet, welche durch die Internet Assigned Numbers Authority (IANA) [4] verwaltet wird. IANA delegiert die Zuteilung der ASN an die Regional Internet Registries (RIR). In Europa ist das das Réseaux IP Européens Network Coordination Centre (RIPE NCC) [5]. Wie in Abb. 2 zu sehen, hat das aufgezeigte AS der Deutschen Telekom die ASN AS3320.

Am Beispiel des Deutschen Forschungsnetzes (DFN) mit der ASN AS680 wird das Szenario konkretisiert [6]. Dabei wird die Topologie auf der physischen Schicht und somit auf der Ebene der Glasfaser (kurz: Fasern) dargestellt. Die Faserplattform des DFN-Wissenschaftsnetzes bestand im Juni 2015 aus ca. 10.250 km bundesweit verlegten Glasfaserpaaren, wie in Abb. 3 visualisiert. Über die Faserplattform werden zwischen den Standorten des Kernnetzes Datenverbindungen mittels sogenannter Dense Wavelength Division Multiplexing (DWDM) Komponenten geschaltet. Die DWDM-Geräte ermöglichen es, über eine Faser eine Vielzahl an parallelen Verbindungen aufzubauen. Die Entscheidung zwischen welchen Standorten des Kernnetzes Verbindungen konfiguriert werden, wird u.a. aus einer kontinuierlichen Beobachtung der Datenströme und der zukünftigen geplanten Entwicklungen auf der Anwendungsebene abgeleitet. Mit der heutigen DWDM-Technik können über eine Faser im DFN-Wissenschaftsnetz maximal 88 Verbindungen zu je 100 Gbit/s geschaltet werden. Die maximale Bandbreite zwischen zwei benachbarten Standorten des Kernnetzes beträgt damit 8.800 Gbit/s oder 8,8 Tbit/s.

Auf Basis der Faserplattform wird die IP-Plattform bzw. die IP-Topologie konfiguriert. Dabei ist die Überwachung der Dienstgüte-Parameter (u.a. Laufzeitmessungen, Schwankungen, Paketverlustrate) auf der IP-Ebene essentiell.

Die Konnektivität zu den europäischen und weltweiten Forschungsnetzen erfolgt über das europäische Netz GÉANT [7], wie in Abb. 4 dargestellt.

Die IP-Plattform des DFN ist an mehrere IXP des kommerziellen Internet angeschlossen. Eine de-

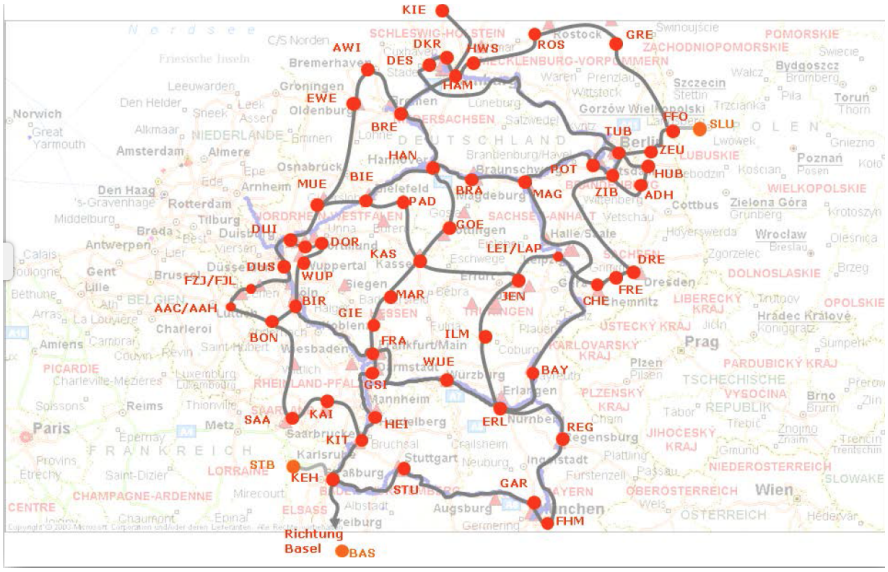


Abbildung 3: Faserplattform des Wissenschaftsnetzes des DFN, Stand Juni 2015

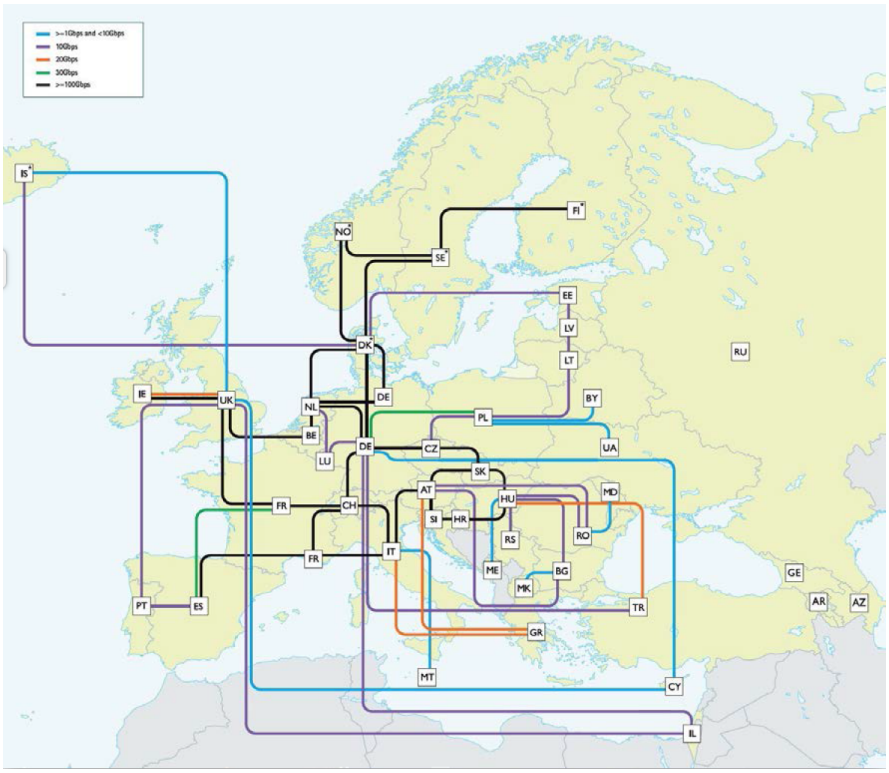


Abbildung 4: Topologie des europäischen Forschungsnetzes

taillierte Auflistung der Peerings und der Transits kann hier nicht gegeben werden, da die Weitergabe dieser Information ein Sicherheitsrisiko für den Dienstanbieter darstellt. Aus den genannten Gründen kann auf die IP-Topologie ebenso nicht näher eingegangen werden.

Durch die Nutzung des RIPE NCC [5] und Eintragung der IP-Adresse des Webservers www.dfn.de (194.95.248.240) werden die öffentlich zugänglichen Informationen des entsprechenden AS-Verbundes (vgl. Abb. 5) jedoch visualisierbar.

Die Intention den Datenverkehr zu klassifizieren und somit aus Sicht der IP-Pakete auf den zu übertragenden Anwendungsmix zu schließen ist fast so alt wie das Internet selbst. Diese Information ist aus unterschiedlichen Aspekten von Relevanz, z.B. um neue Trends in der Internet-Nutzung zu erkennen, die gesamte Topologie und Ressourcen-Nutzung (Router, Switches, Server, usw.) hinsichtlich vorgegebener Zielfunktionen zu optimieren und Anomalien zu entdecken, die beispielsweise auf einen Advanced Persistent Threat hinweisen können. Es gibt eine Vielzahl von wissenschaftlichen Ansätzen für die Klassifizierung von Anwendungen aus dem IP-Datenverkehr. Eine Übersicht verschiedener Ansätze ist z.B. unter [8], [9] zu finden. Jedoch ist u.a. durch die Komplexität und Heterogenität der Netzinfrastruktur sowie Aspekten wie Sicherheit, Leistung und Kosten die Frage nach dem Anwendungsmix im Allgemeinen schwierig zu beantworten. Auf einzelne Ansätze wird in der Beantwortung der Fragen eingegangen.

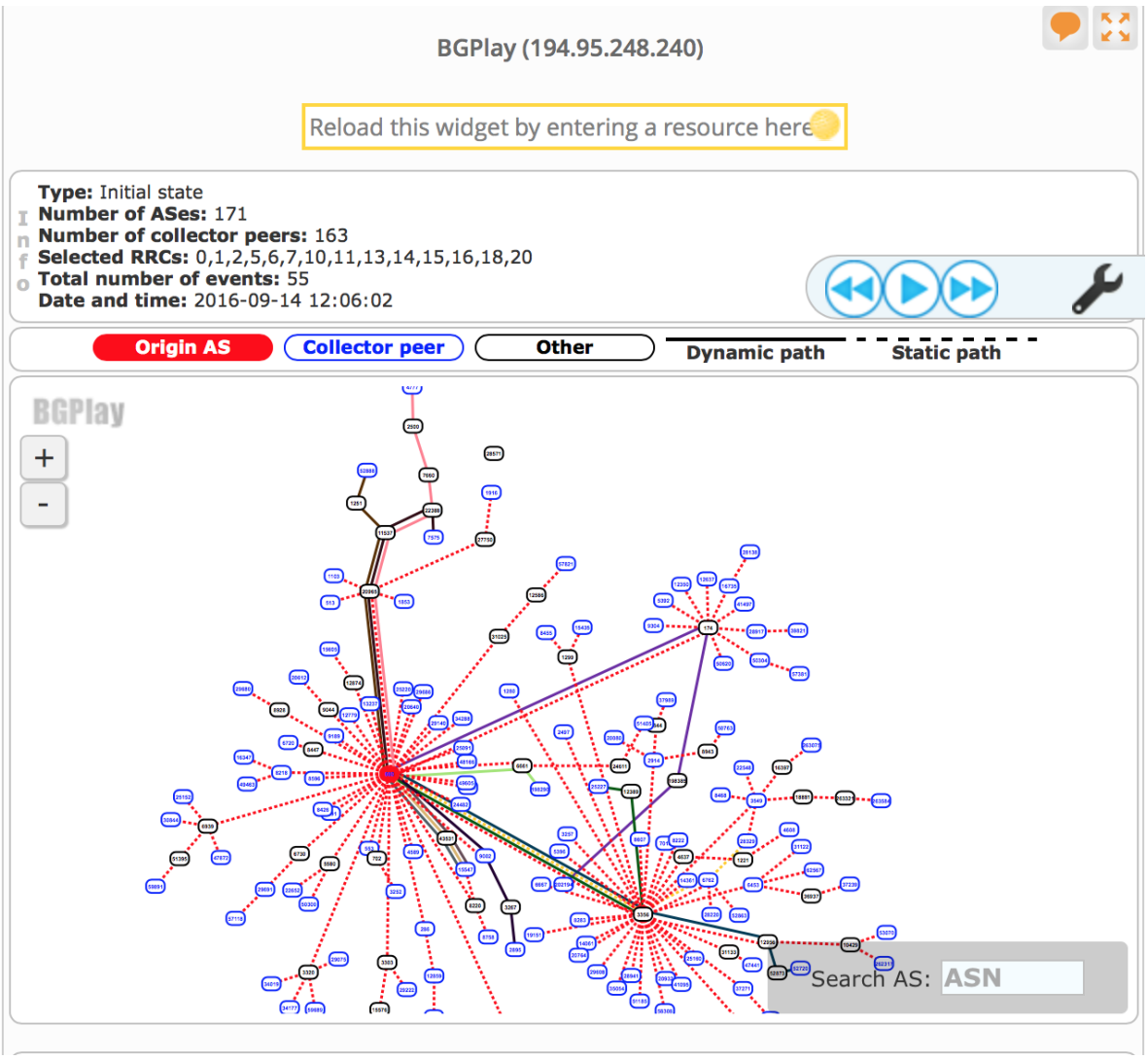


Abbildung 5: DFN (AS680) im AS-Verbund, visualisiert durch RIPE NCC

Das Gutachten nimmt Stellung zu folgenden Fragen:

1. Wie sind die in einem IXP verbundenen AS aufgebaut, aus wie vielen AS aus welchen Ländern besteht ein solcher Verbund in der Regel (am Beispiel DE-CIX) und auf welche Art und Weise erfolgt der Datentransfer (Routing) im Rahmen eines solchen Verbundes?

2. Ob und wie kann der in einem solchen IXP anfallende Netzwerkverkehr nach Inhalten der jeweiligen Kommunikation (z.B. E-Mail, VoIP, Instant Messaging, Webseiten, Video-Streams etc.) differenziert werden und ist es möglich, diese jeweiligen Kommunikationsarten zu quantifizieren (z.B. als Anteil am Gesamtverkehr)?

3. Welche Arten von Verkehrsdaten müssen erfasst werden, um einzelne Datenpakete zu einem inhaltlich vollständigen Kommunikationsvorgang zusammenzusetzen?

4. Wie gestaltet sich das IP-Verkehrsrouting (Transit und Peering), welche Akteure be-stimmen nach welchen Kriterien das Routing (bzw. die Streckenführung) der einzelnen Pakete und welche Parameter werden in welcher Art und Weise von Providern in den Routingtabellen bezogen auf die einzelnen AS erfasst?

5. Ob und wie kann für einzelne Kommunikationsvorgänge Rückschluss auf Ursprungs- und / oder Zielort der übermittelten Kommunikation gezogen werden?

6. Inwiefern kann man sich zur Auswahl konkreter Übertragungstrecken für eine Ausleitung gemäß §27 Abs. 2 TKÜV bzw. zur strategischen Überwachung von Ausland-Ausland-Telekommunikationsverkehr auf Wahrscheinlichkeiten bestimmter regionaler Zuordnung von Kommunikationsvorgängen auf diesen Strecken stützen und wie sind solche Wahrscheinlichkeiten gegebenenfalls zu berechnen oder zu quantifizieren?

7. Welche Charakteristika existieren im Rahmen des paketvermittelten Telekommunikationsverkehrs, die eine nationale Zuordnung von Kommunikationsvorgängen in, von und nach Deutschland innerhalb der AS erlauben und mit welchen Maßnahmen sowie mit welcher Genauigkeit lassen sich solche Zuordnungen treffen?

8. Welche praktisch erprobten und/oder wissenschaftlich anerkannten Methoden zur länder- genauen Geolokalisierung von IP-Adressen bzw. IP-Datenpaketen gibt es (online und offline) und wie zuverlässig sind diese Methoden für eine Zuordnung zum Standort „Deutschland”?

9. Welche Aspekte der den einzelnen Kommunikationsvorgängen zu zuordnenden Verkehrsdaten lassen darüber hinaus eventuell Rückschlüsse darauf zu, ob ein Teilnehmer des jeweiligen Kommunikationsvorgangs Deutscher ist?

10. Welche Möglichkeiten gibt es für Dritte, d.h. Personen die nicht den Betreibern der AS zuzuordnen sind, den Regelbetrieb zu beeinflussen bzw. zu beeinträchtigen und welche Schutzmaßnahmen existieren ggf. gegen solche Manipulationen?

2 Stellungnahme zu den einzelnen Fragen

Im Folgenden wird ausführlich auf die oben genannten Fragen eingegangen.

2.1 Wie sind die in einem IXP verbundenen AS aufgebaut, aus wie vielen AS aus welchen Ländern besteht ein solcher Verbund in der Regel (am Beispiel DE-CIX) und auf welche Art und Weise erfolgt der Datentransfer (Routing) im Rahmen eines solchen Verbundes?

Das primäre Ziel im Internet ist das optimale Routing des Datenverkehrs, mit u.a. geringer Verzögerung, hohem Durchsatz sicherzustellen. IXP wurden eingerichtet, um Betreibern von AS ein quasi lokales Netz anzubieten, über welches sich die Daten schnell und einfach austauschen lassen. Die IXP bieten den AS-Betreibern eine Schicht-2-Infrastruktur (ISO/OSI-Referenzmodell) über die der Datenverkehr weitergeleitet werden kann. Hierbei werden zwischen AS nur Daten ausgetauscht, wenn ein Abkommen zwischen den Betreibern besteht. Diese lassen sich grundsätzlich in zwei Kategorien unterteilen: Transit und Peering. Im Falle von Transit zahlt ein Betreiber (Customer) einem anderen Betreiber (Provider) ein Entgelt für die Weiterleitung des Datenverkehres durch dessen AS. Ein Peering-Agreement bezeichnet eine Vereinbarung zwischen zwei Betreibern, den Datenverkehr des jeweils anderen durch das eigene AS weiterzuleiten, ohne gegenseitigen finanziellen Ausgleich [10].

Vereinzelt betreiben IXP ebenso die Dienste eines sogenannten Route Servers. Diese Plattform bietet eine weitere Möglichkeit der Verbindung von AS. Kunden eines IXP können sich mit diesen Route Servern verbinden (Peering), der BGP-Routen zwischen allen angeschlossenen AS verteilt. Insbesondere für kleine und mittlere AS ist diese Art des multilateralen Peerings aufgrund der geringeren Kosten attraktiv [1, 11].

Der Aufbau eines AS wird schematisch in Abb. 6 dargestellt.

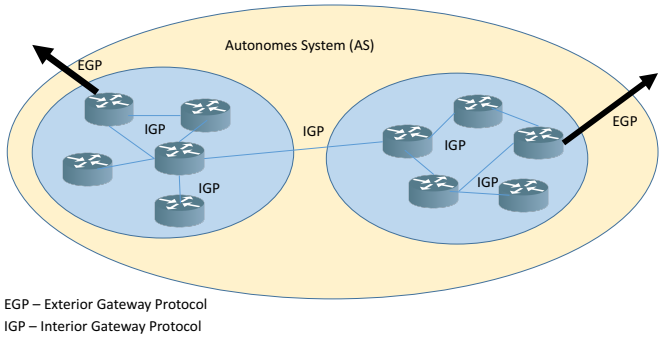


Abbildung 6: Interior und Exterior Routing in einem AS

Der DE-CIX in Frankfurt ist einer der größten IXP weltweit [12]. Mit mehr als 700 angeschlossenen AS aus mehr als 60 Ländern und einem Datenverkehrsaufkommen von 5,1 Tbit/s in der Spitze [13] ist der DE-CIX neben dem London Internet Exchange und dem Amsterdam Internet Exchange führend für den Datenverkehr in Zentral- und Osteuropa. Dabei ist der DE-CIX örtlich nicht nur auf Frankfurt beschränkt, sondern betreibt IXP in insgesamt elf Städten, wie z.B. Hamburg, München,

Madrid, New York und Dubai. Die Herkunft der Betreiber der AS (entsprechend Kunden von DE-CIX) ist nicht auf Europa beschränkt. Da die AS bei der IANA registriert sind, lässt sich die Herkunft der Betreiber bestimmen. Ein AS selbst kann sich dabei geografisch über mehrere Länder erstrecken.

In den einleitenden Worten wurde ein AS als ein Verbund von IP-Teilnetzen unter einer administrativen Instanz (Organisation oder Unternehmen) bezeichnet. In Request for Comments (RFC) 1930 [14] wird diese Definition insofern präzisiert, da ein AS als Gruppe von einen oder mehreren IP-Präfixen definiert wird, die von einem oder mehreren Netzdienstleistern mit einer einzigen und präzise definierten Routing-Policy, geroutet werden. Die Routing-Policy legt fest wie Routing-Informationen zwischen AS ausgetauscht werden.

Es wird zwischen Routing-Protokollen unterschieden, die innerhalb eines AS den Datenverkehr weiterleiten, und Protokollen, die zwischen den AS für das Routing eingesetzt werden. Innerhalb eines AS findet die Wegfindung und -wahl (Routing) mittels eines oder mehrere Interior Gateway Protokolle (IGP) statt. Heutzutage werden zumeist die Protokolle Open Shortest Path First (OSPF) oder Intermediate System to Intermediate System (IS-IS) verwendet. Zwischen den AS wird ein Exterior Gateway Protokoll (EGP) zum Austausch der Erreichbarkeitsinformationen eingesetzt. De facto ist das Border Gateway Protocol (BGP) das einzige derzeit verwendete EGP.

BGP gehört zu der Klasse der Pfad-Vektor-Routingprotokolle. Dies bedeutet, dass BGP-Router, welche als Nachbarn konfiguriert sind, Pfade zu erreichbaren Netzen austauschen. Ein solcher Pfad besteht aus einer Aneinanderreihung von ASN, welche auf dem Weg zum beschriebenen Zielnetz durchlaufen werden müssen. Die Entscheidung, welcher Weg zu einem Zielnetz gewählt wird, hängt neben der Pfadlänge von weiteren Attributen ab. So können lokale Präferenzen, Metriken, das Vorhandensein von redundanten Anbindungen sowie die aktuelle Lastsituation der einzelnen Verbindungen (sogenanntes „Load Balancing") die Routingentscheidung der BGP-Router beeinflussen. Auf diese Problemstellung wird in Abschnitt 2.4 näher eingegangen.

2.2 Ob und wie kann der in einem solchen IXP anfallende Netzwerkverkehr nach Inhalten der jeweiligen Kommunikation (z.B. E-Mail, VoIP, Instant Messaging, Webseiten, Video-Streams etc.) differenziert werden und ist es möglich, diese jeweiligen Kommunikationsarten zu quantifizieren (z.B. als Anteil am Gesamtverkehr)?

Technisch gesehen ist es möglich, einen Großteil des Datenverkehrs in einem IXP nach Anwendung bzw. Kommunikationsart zu differenzieren. Hierfür gibt es mehrere wissenschaftliche Ansätze, welche im Nachfolgenden kurz erläutert werden. Je nach Umgebung und Art des Netzverkehrs funktionieren diese unterschiedlich gut, wobei hier nur auf die für die Fragestellung relevanten eingegangen wird.

Eine Möglichkeit eine Differenzierung zu treffen ist der sogenannte Port-basierte Ansatz. Ports dienen der Adressierung auf der Transportschicht des ISO/OSI-Referenzmodells und sind 16-Bit Zahlen, die zwischen 0 und 65535 liegen. Die Ports zwischen 0 und 1023 werden als Well-Known-Ports bezeichnet, welche von der IANA [4] zentral verwaltet werden. Diese sind für dedizierte Dienste reserviert; andere Anwendungen dürfen diese Ports nicht nutzen. Wenn ein Nutzer beispielsweise eine Webseite mittels des Hypertext Transfer Protocol Secure (HTTPS) von einem Webserver aufrufen möchte, so wird der Server neben der Ziel-IP-Adresse mit der Ziel-Portnummer 443 angesprochen.

Die Ports zwischen 1024 bis 49151 werden als registrierte Ports bezeichnet und werden ebenso von der IANA verwaltet. Auch für die registrierten Ports existiert eine Zuordnung von Portnummern zu dedizierten Diensten und somit Anwendungen. Im Gegensatz zu den Well-Known-Ports ist es grundsätzlich vorgesehen, dass auch andere Anwendungen registrierte Ports nutzen.

Aufgrund der Portnummern können trotzdem Rückschlüsse auf die Art der Kommunikation gezogen werden. Wird beispielsweise festgestellt, dass ein Client einen Server mit dessen IP-Adresse und dem Port 5060 anspricht, so liegt der Schluss nahe, dass es sich hierbei um das Session Initiation Protocols (SIP) [15] handelt und somit der Datenverkehr dem Auf- oder Abbau von VoIP-Verbindungen zuzuordnen ist.

Eine Auflistung aller Well-Known- und registrierten Ports ist in RFC 1700 [16] zu finden, wobei diese Auflistung gemäß RFC 3232 [17] nur noch in einer Online Datenbank der IANA [18] aktualisiert wird. Die Ports 49151 bis 65535 werden als dynamisch bezeichnet und sind frei verfügbar.

Der Port-basierte Ansatz kann einen Anhaltspunkt auf vorhandenen Kommunikationsarten in Datenverkehr geben, liefert aber nicht zwangsläufig zuverlässige Ergebnisse. Zum Umgehen von Kontrollen durch den ISP nutzen Anwendungen zunehmend nicht-standardisierte oder zufällig gewählte Ports. Zur Verschleierung oder zum Umgehen von Beschränkungen durch Firewalls werden auch Well-Known oder registrierte Ports für nicht vorgesehene Zwecke und Dienste verwendet. Dies hat zur Folge, dass der Port-basierte Ansatz falsche Zuordnungen liefert, aber dennoch für eine erste Klassifizierung geeignet ist [19].

Eine weitere Möglichkeit ist das Aufzeichnen und die Analyse des gesamten Netzverkehrs. Auf diese Weise kann neben den Steuerinformationen auch der eigentliche Inhalt der übertragenen Daten analysiert werden. Diese Technologie bezeichnet man als Deep Packet Inspection (DPI) und wird in Netzen insbesondere zur Detektion von Schadsoftware, Spam und anderen unerwünschten Inhalten eingesetzt. Dabei wird in dem analysierten Datenverkehr auf bestimmte Signaturen, d.h. auf bekannte Muster und bestimmte Schlüsselwörter im Datenverkehr, welche charakteristisch für die einzelnen Dienste und Anwendungen sind, geachtet und anhand dessen eine Klassifizierung durchgeführt [19, 20].

Im Allgemeinen bietet diese Methode den besten Einblick in den Netzverkehr, da komplette Pakete aufgezeichnet und analysiert werden. DPI lässt sich theoretisch ebenso auf Netzverkehr in einem IXP anwenden, wobei von einer Big Data Problematik gesprochen werden kann, da es praktisch kaum machbar ist, die Gesamtheit der Daten aufzuzeichnen und zu verarbeiten. Die Analyse des auftretenden Datenverkehrs[1] benötigt aufgrund der Menge an Daten und der notwendigen Verarbeitungsgeschwindigkeit sehr große Speicher- und Rechenkapazität. Herausforderungen für die Verarbeitung und Auswertung der Daten sind ebenso Paket-Fragmentierungen, wiederholt übertragenen Datenpaketen und asymmetrischen Routen.[2] Des Weiteren scheitert DPI aufgrund des Ansatzes bei verschlüsseltem oder verschleiertem Datenverkehr [19, 21].

Flow-basierte Verfahren sind eine dritte Art von möglichen Differenzierungsansätzen. Hierbei werden sogenannte Flow-Daten analysiert. Ein Flow ist als eine Menge von IP-Paketen definiert, die einen konkreten Beobachtungspunkt innerhalb eines festgelegten Zeitrahmens passieren und über eine Menge von gemeinsamen Eigenschaften verfügen. Solche Eigenschaften können Steuerinformationen wie Quell- und Ziel-IP-Adressen, Port-Nummern oder auch anwendungsspezifische Informationen sein. Ein Paket wird immer dann einem bestimmten Flow zugeordnet, wenn es alle vorher definierten Eigenschaften besitzt [22]. Mit der Hilfe von Flow-Export Protokollen, wie NetFlow oder IP Flow Information eXport (IPFIX), können diese Flow-Daten für eine Analyse exportiert werden [23].

Zum einen können aus diesen Daten ebenso die Informationen über verwendete Ports extrahiert und somit mittels des Port-basierten Ansatzes der Datenverkehr klassifiziert werden. Zum anderen lassen sich weitere Eigenschaften der verschiedenen Flows zur Klassifizierung nutzen. Beispielsweise können die Anzahl der Pakete während des definierten Zeitfensters oder die durchschnittliche Pa-

[1] Bsp. bei DE-CIX: Derzeit durchschnittlich 3 Tbit/s bzw. 1,2 Petabyte pro Tag [13]

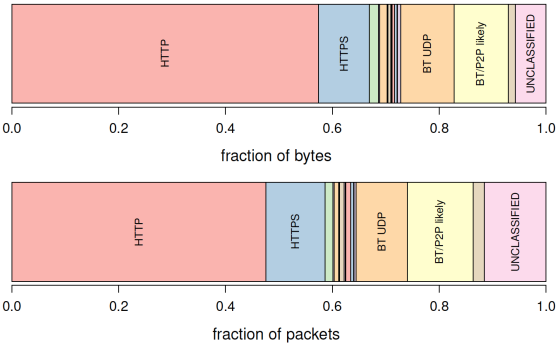
[2] Hinweg und Rückweg zwischen Kommunikationspartnern sind unterschiedlich

ketgröße Rückschlüsse auf die Anwendung zulassen. Die Klassifizierung erfolgt bei diesen Ansätzen anhand statistischer Methoden (z.B. [24], [25] und [26]).

Der Vorteil von Flow-basierten Ansätzen ist die geringe Datenmenge, die gespeichert, verarbeitet und analysiert werden muss. Ferner ist die Möglichkeit gegeben verschlüsselten Datenverkehr zu analysieren, da der eigentliche Inhalt (Payload) nicht betrachtet wird. Dem gegenüber steht die geringere Genauigkeit der Klassifizierung im Vergleich zu DPI [19].

Eine Quantifizierung der Kommunikationsarten als Anteil am Gesamtverkehr in einem IXP ist beispielsweise mittels des Protokolls sFlow [27] möglich. sFlow ist ein Industriestandard, welcher in vielen paketverarbeitenden Geräten implementiert ist und die Funktionalität besitzt, einerseits die Anzahl von Paketen an den Schnittstellen der Geräte zu erfassen und andererseits repräsentativen Datenverkehr zu exportieren. sFlow ist trotz der namentlichen Nähe klar von NetFlow und IPFIX abzugrenzen, da die eben erwähnte Funktionalität nicht zum Umfang von Flow-Export Protokollen gehört [23]. Eine im Jahr 2011 durchgeführte Studie [12] analysiert und quantifiziert anhand von sFlow den Datenverkehr an einem IXP. Diese Studie zeigt beispielsweise, dass mehr als 50% des Datenverkehrs unverschlüsselte Webseiten Aufrufe mittels Hypertext Transfer Protocol (HTTP) waren. In der Analyse wurden ebenso weitere Anwendungen quantifiziert.

Eine detailliertere Differenzierung und Quantifizierung des Datenverkehrs an einem IXP wird in [20] erläutert. Hierbei beziehen sich die Autoren auf Messungen aus den Jahren 2011 bis 2013. In Abb. 7 wird beispielhaft die Klassifizierung des Datenverkehrs an einem IPX dargestellt. Die gezeigten Werte beziehen sich dabei auf eine Stichprobe vom September 2013, bei welcher der Datenverkehr an einem großen Europäischen IXP für eine Woche lang (168 Stunden) ausgewertet wurde. Auch hier bestätigt sich, dass mehr als 50% des Datenverkehrs (gemessen in Paketanzahl und Datenmenge) Webseiten Aufrufe mittels HTTP und HTTPS sind. Ebenso ist zu sehen, dass mehr als 94% der Datenmenge klassifiziert werden konnten.



Protocol	% bytes	% pkts
HTTP	57.39	47.52
HTTPS	9.53	11.08
RTMP	1.72	1.48
NNTP	1.41	0.87
NNTPS	0.63	0.38
SMTP	0.53	0.91
DNS	0.45	0.87
SSH	0.42	0.61
other known	0.68	0.74
BT UDP	10.00	9.57
BT/P2P likely	10.14	12.31
P2P likely	1.32	2.10
unclassified	5.78	11.56

Abbildung 7: Prozentuale Quantifizierung und Klassifizierung des Netzverkehrs nach Protokollen gemessen in Paketen und Bytes (September 2013) [20]

Letztendlich ist eine vollständige Differenzierung des Datenverkehrs in der Regel nicht möglich. Auf Grundlage der beschriebenen Ansätze existiert eine Vielzahl an Methoden und Techniken, welche aus den genannten Gründen jedoch jeweils keine vollständige Klassifizierung ermöglichen. Die Verwendung von verschlüsselter Kommunikation, der Einsatz von Tunneltechnologien und virtuellen privaten Netzen (VPN) und eine Kommunikation mittels proprietärer Protokolle sind beispielhafte Methoden, wie eine Differenzierung erschwert bzw. verhindert werden kann [28]. Eine weitere exemplarische Möglichkeit, Datenverkehr zu verschleiern und somit einer vollständigen Klassifikation entgegenwirkend sind sogenannte Covert Channels, bei welchen gezielt versucht wird, die Existenz

einer stattfindenden Kommunikation zu verstecken. Die Grundidee hierbei ist, dass Informationen und Daten in regulärem Datenverkehr eingebettet werden und somit nicht erkennbar sind [29].

Diese Beispiele sind nur eine Auswahl der existierenden Verschleierungs- und Verschlüsselungsmethoden, welche die Schwierigkeit der Differenzierung verdeutlichen.

2.3 Welche Arten von Verkehrsdaten müssen erfasst werden, um einzelne Datenpakete zu einem inhaltlich vollständigen Kommunikationsvorgang zusammenzusetzen?

Verkehrsdaten umfassen laut §96 Telekommunikationsgesetz (TKG):

1. die Nummer oder Kennung der beteiligten Anschlüsse oder der Endeinrichtung, personenbezogene Berechtigungskennungen, bei Verwendung von Kundenkarten auch die Kartennummer, bei mobilen Anschlüssen auch die Standortdaten,

2. den Beginn und das Ende der jeweiligen Verbindung nach Datum und Uhrzeit und, soweit die Entgelte davon abhängen, die übermittelten Datenmengen,

3. den vom Nutzer in Anspruch genommenen Telekommunikationsdienst,

4. die Endpunkte von festgeschalteten Verbindungen, ihren Beginn und ihr Ende nach Datum und Uhrzeit und, soweit die Entgelte davon abhängen, die übermittelten Datenmengen,

5. sonstige zum Aufbau und zur Aufrechterhaltung der Telekommunikation sowie zur Entgeltabrechnung notwendige Verkehrsdaten.

Diese reichen jedoch *nicht* aus, um einen inhaltlich vollständigen Kommunikationsvorgang zusammenzusetzen. Zwar beinhalten Verkehrsdaten hierfür wichtige und erforderliche Daten wie die IP-Adressen der Kommunikationspartner, jedoch sind für die Rekonstruktion des Inhaltes immer die jeweils transportierten Nutzdaten erforderlich, was das Aufzeichenen des gesamten Kommunikationsvorgangs erfordert.

Die bereits erwähnten Flows, welche oftmals zur Analyse von Verkehrsdaten genutzt werden, und hierfür Quelladresse, Zieladresse, Ausgangsport und Zielport, die Sequenznummern der Pakete sowie ggf. Zeitstempel zur eindeutigen Zuordnung eines Pakets zu einer bestimmten Kommunikation festhalten, reichen für eine inhaltlich vollständige Rekonstruktion nicht aus. Weiterhin muss beachtet werden, dass Flows selbst für das Zusammensetzen eines Kommunikationsvorganges ohne dessen Inhalt ungenügend sein können.

Die wichtigsten Standards für Flows sind NetFlow von Cisco sowie sFlow. Letzteres basiert maßgeblich auf den Arbeiten von Hewlett Packard und der Universität von Genf. Die erste Version von NetFlow wurde 1996 von Cisco entwickelt, die entsprechenden Funktionalitäten gingen in Ciscos Internetwork Operating System (IOS) ein. Seither wurde die ursprüngliche Variante mehrmals erweitert, aktuell ist Version 9 aus dem Jahre 2004 (RFC 3954 [30]), jedoch ist die Version 5 noch am weitesten verbreitet. NetFlow Version 9 ist weiterhin die Basis für IPFIX [31, 22], das von der Internet Engineering Task Force (IETF) als offener Standard entwickelt wird.

Die erste Version von sFlow wurde 2001 verabschiedet (RFC 3176 [32]); seit 2002 ist Hardware mit entsprechender Unterstützung verfügbar, z.B. Switches von Hewlett Packard, D-Link und Allied Telesyn.
Ein ursprünglich wichtiger Unterschied zwischen NetFlow und sFlow liegt im Umfang der Evaluation von Netzpaketen. Während bei NetFlow jedes Paket, das durch den Router läuft, evaluiert wird,

nutzt sFlow eine Abtastrate, so dass nur jedes n-te Paket analysiert wird, um die Last auf der aktiven Netzkomponente geringer zu halten. Typische Abtastraten liegen beispielsweise im Bereich von 100, entsprechend wird bspw. nur jedes 100ste Paket analysiert. Somit müssen hierbei Algorithmen genutzt werden, um eine korrekte, statistische Repräsentation des Datenverkehrs zu erhalten. Bei entsprechend leistungsfähiger Hardware ist es auch möglich die Abtastrate auf eins zu setzen. Mit den Routern Cisco 12000 wurde auch für NetFlow die Nutzung von Abtastraten eingeführt.
Im ungünstigsten Fall kann es bei der Nutzung von Abtastraten vorkommen, dass sehr kurze Kommunikationsvorgänge vereinzelt unberücksichtigt bleiben.

Werden Daten innerhalb von Tunneln transportiert, gestaltet sich die Situation noch schwieriger. Beispielsweise können Daten zwischen Geräten, welche über IPv6-Adressen kommunizieren, auch über dazwischen liegende IPv4-Netze transportiert werden. Dabei werden die eigentlichen Datenpakete inklusive ihrer Headerinformationen durch spezielle Protokolle mit zusätzlichen IPv4-Adressen versehen. Auf diese Weise werden die Daten mittels eines Tunnels durch das IPv4-Netz transportiert. Somit müssen die Nutzdaten des Transportprotokolls zwingend untersucht werden, um an die Metadaten wie die Ziel- und Quelladressen sowie die Ports der eigentlichen Kommunikation zu gelangen. Bei verschlüsseltem Payload kann dies typischerweise nicht gewährleistet werden.

Nochmals erwähnt sei, dass sich anhand reiner Metadaten, wie sie bei der Generierung dieser Verkehrsdaten regelmäßig genutzt werden, *keine* Rekonstruktion des inhaltlichen Kommunikationsvorganges realisieren lässt, da keine Betrachtung bzw. Speicherung der eigentlichen Nutzdaten erfolgt. Dies erfordert zwingend die Speicherung der Nutzdaten durch eine vollständige Aufzeichnung des Kommunikationsvorganges.

Darüber hinaus ist hier nochmals zu betonen, dass bei den zuvor vorgestellten Flow Standards für eine bidirektional stattfindende Kommunikation zwei separate unidirektionale Flows generiert werden. Da Hin- und Rückkommunikation verschiedene Pfade im Netz nehmen können (vgl. Abschnitt 2.4), ist die vollständige Erfassung der Kommunikation bzw. der dazugehörigen Metadaten (Flows) äußerst schwierig.

2.4 Wie gestaltet sich das IP-Verkehrsrouting (Transit und Peering), welche Akteure bestimmen nach welchen Kriterien das Routing (bzw. die Streckenführung) der einzelnen Pakete und welche Parameter werden in welcher Art und Weise von Providern in den Routingtabellen bezogen auf die einzelnen AS erfasst?

Innerhalb eines AS ist der jeweilige Betreiber für das Routing verantwortlich. Hierbei setzen die Betreiber in der Regel ein oder mehrere IGP ein, welche sie entsprechend ihrer Bedürfnisse konfigurieren. Wie in Abschnitt 2.1 aufgezeigt, sind AS heterogen und können nicht allgemeingültig beschrieben werden.

Wie bereits erwähnt, wird zwischen den AS das BGP [33] als Routing-Protokoll eingesetzt. Zum Austausch von Routing-Informationen werden in BGP Nachbarschaften konfiguriert. Benachbarte Router tauschen über Transmission Control Protocol (TCP) Verbindungen die bekannten Informationen aus. Dabei handelt es sich entweder um Netze, die der jeweilige Router von anderen BGP-Nachbarn empfangen hat oder um Netze, welche sich im jeweils eigenen AS befinden. Ein IXP stellt den Betreibern der AS als neutraler Partner eine Infrastruktur zur Verfügung, worüber BGP-Nachbarschaften aufgebaut werden können.

Aufgrund der Peering- und Transit-Beziehungen ist das Internet ein vermaschtes Netz, welches aus vielen AS und den Verbindungen zwischen den AS besteht. Tier 1 Provider innerhalb des Internets kennen aufgrund der vielfältigen Verbindungen durchschnittlich zehn bis zwanzig unterschiedliche Pfade zu einzelnen Zielnetzen. Es ist hierbei üblich, dass die Netzbetreiber zwischen den unterschied-

lichen AS mehr als eine Verbindung unterhalten. Dies bringt zum einen den Vorteil, dass durch die redundante Anbindung der Ausfall einzelner Verbindungen aufgefangen werden kann. Zum anderen kann eine Lastverteilung auf mehrere Leitungen erfolgen. Welche Verbindung letztendlich für den Datenverkehr zu einem bestimmten Zielnetz gewählt wird, beruht auf der Entscheidung des Routing-Algorithmus, welcher in BGP implementiert ist [34].

In BGP sind eine Reihe von Attributen spezifiziert, welche in den Routing-Entscheidungen und -Tabellen Anwendung finden. BGP-Nachbarn tauschen diese für einzelne Zielnetze aus. Ein wichtiges Attribut ist der sogenannte AS-PATH, welcher eine Auflistung von AS-Nummern darstellt, durch welche ein Zielnetz erreichbar ist. Im Regelfall wird bei der Routing-Entscheidung die Route mit dem kürzesten AS-PATH gewählt.

Der NEXT-HOP ist ein weiteres gespeichertes Attribut. Dieser ist jeweils die IP-Adresse des benachbarten Routers, von welchem die Route (AS-PATH) zu einem Zielnetz empfangen wurde. BGP umfasst darüber hinaus weitere Attribute, welche die Routing-Entscheidungen auf den einzelnen Routern beeinflussen. So kann auf einem Router eine LOCAL PREFERENCE vergeben werden, die einzelnen Pfaden Vorrang gewährt, obwohl ggf. kürzerer Pfade existieren.

Durch den gerade beschriebenen Präferenzwert kann ein BGP-Router beeinflussen, welche der empfangenen Routen zu einem Zielnetz gewählt wird. Durch das MULTI-EXIT-DISCRIMINATOR (MED) Attribut kann ein AS, welches ein Zielnetz über mehrere Wege an das Nachbar-AS bekannt gibt, beeinflussen, welcher Weg präferiert werden soll. Darüber hinaus wird je Pfad eine ORIGIN-ID erfasst, die Auskunft darüber gibt, ob ein Zielnetz mittels eines IGP oder eines EGP dem BGP-Prozess bekannt wurde [33].

Im RFC 1997 [35] wurde zusätzlich noch die Möglichkeit eingeführt, Routen mit einem COMMUNITY Attribut zu markieren. Eine COMMUNITY ist dabei als eine Gruppe von Zielnetzen definiert, welche über gemeinsame Eigenschaften verfügen [35]. In jedem AS besteht die Möglichkeit, die Zielnetze zu einer COMMUNITY zuzuordnen, wobei standardmäßig alle Zielnetze der globalen Internet-Community zugeordnet sind. Beispielsweise kann ein Weg zu einem Zielnetz mit einer COMMUNITY-Markierung versehen werden, welche die geografische Lage des Gateways ausdrückt, über welches dieser Weg gelernt wurde. Auf diese Weise kann im Anschluss der Datenverkehr bei Verwendung dieser Route in diesem AS so gefiltert werden, dass beispielsweise eine Weiterleitung auf eine Region beschränkt wird.

Im Rahmen von Peering- und Transit-Vereinbarungen besteht ebenso die Möglichkeit, dass AS-Betreiber mit Hilfe der COMMUNITY-Markierungen die Wegewahl entsprechend beeinflussen. Hierbei sind die Möglichkeiten vielfältig und beruhen auf bilateralen Abkommen der Betreiber, wobei häufig die Option der geografischen Beschränkung enthalten ist [34].

Eine letzte Möglichkeit zur Beeinflussung des Routing-Prozesses ist die Nutzung von Filtern. Durch auf den BGP-Routern implementierte Filter ist es möglich, Informationen über spezifische Zielnetze von bestimmten Nachbarn zu ignorieren. Auf diese Weise finden diese Informationen im eigentlichen Routing-Prozess keine Beachtung. Gleichermaßen besteht die Möglichkeit mittels Filter zu verhindern, dass Pfade zu spezifischen Zielnetzen an definierte Nachbarn weitergegeben werden [36].

In Listing 1 wird am Beispiel die gespeicherte Information auf dem BGP-Router rs1 am DE-CIX Frankfurt für das Subnetz der Universität der Bundeswehr München (137.193.0.0/16) dargestellt [37]. Es ist zu sehen, dass zu diesem Zielnetz zwei Wege existieren (erster Weg via NEXT-HOP 80.81.192.222, zweiter Weg via NEXT-HOP 80.81.193.222). Der AS-PATH besteht jeweils nur aus dem AS680, welches unter Verwaltung des DFN steht. Die LOCAL PREFERENCE ist auf diesem Router für beide Wege gleich konfiguriert und hat den Wert 100. Die Entscheidung des Routers, dass der erste Weg der präferierte ist, beruht auf dem Wert des MED. Dieser ist bei dem ersten Weg mit

845 niedriger als bei dem zweiten Weg (1092). Dementsprechend bevorzugt das AS680 in diesem Fall den Weg über den Router mit der IP-Adresse 80.81.192.222, was der in Listing 1 dargestellte Router in die Routing-Entscheidung einfließen lässt.

```
> show ip bgp 137.193.0.0
* * * Note: the first route is the BEST * * *
137.193.0.0/16      via 80.81.192.222 on bond0 [R192_222 2016-08-24 02:04:56] * (100) [AS680i]
Type: BGP unicast univ
BGP.origin: IGP
BGP.as_path: 680
BGP.next_hop: 80.81.192.222
BGP.med: 845
BGP.local_pref: 100
BGP.community: (0,6695) (6695,8075) (6695,16509) (6695,20825) (6695,25152)
via 80.81.193.222 on bond0 [R193_222 2016-09-06 16:08:07] (100) [AS680i]
Type: BGP unicast univ
BGP.origin: IGP
BGP.as_path: 680
BGP.next_hop: 80.81.193.222
BGP.med: 1092
BGP.local_pref: 100
BGP.community: (0,6695) (6695,16276)
```

Listing 1: BGP Prefix Info für UniBwM Subnetz 137.193.0.0 auf dem DE-CIX BGP-Router rs1 [37]

Auf den Routern werden weiterhin Daten zu den einzelnen Nachbarn erfasst. Beispielsweise ist im Listing 2 die Zusammenfassung des BGP-Routers EE1 am CERN in der Schweiz [38] dargestellt. Im oberen Bereich befinden sich Informationen zu diesem Router selbst und Eigenschaften des BGP-Prozesses sowie Informationen über gelernte Zielnetze. Weiterhin befinden sich in der unteren Hälfte der Zusammenfassung Informationen über alle BGP-Nachbarn des Routers in Tabellenform. Dabei wird deutlich, dass die Router sich die IP-Adresse des jeweiligen Nachbarn, die benachbarte ASN, den Status sowie die Dauer der aktuell bestehenden Nachbarschaft speichern. Ferner wird die Information über die Anzahl der Zielnetze, die empfangen, gefiltert sowie gesendet wurden, vorgehalten.

```
BGP4 Summary
  Router ID: 192.65.184.1    Local AS Number: 513
  Confederation Identifier: not configured
  Confederation Peers:
  Cluster ID: 513
  Maximum Number of IP ECMP Paths Supported for Load Sharing: 4
  Number of Neighbors Configured: 24, UP: 23
  Number of Routes Installed: 1150037, Uses 98903182 bytes
  Number of Routes Advertising to All Neighbors: 2371741 (606137 entries), Uses 29094576 bytes
  Number of Attribute Entries Installed: 363557, Uses 32720130 bytes
  Neighbor Address  AS#          State   Time          Rt:Accepted Filtered Sent     ToSend
  62.40.100.9       20965        ESTAB   109d 0h58m    9           0        28       0
  62.40.124.157     20965        ESTAB   109d 0h58m    16267       0        28       0
  83.97.88.33       21320        ESTAB   109d 0h58m    326797      0        28       0
  192.16.155.2      59624        ESTAB   13d20h59m     23          0        28       0
  192.16.155.18     2697         ESTAB   3d 3h 5m      3           0        28       0
  192.16.155.22     24167        ESTAB   2d12h 5m      8           0        28       0
  192.16.155.30     17579        ACTIV   29d23h 9m     0           0        0        28
  192.16.155.66     43115        ESTAB   16h 0m46s     1           0        28       0
  192.65.184.2      513          ESTAB   0h52m11s      160385      0        559753   0
  192.65.184.3      513          ESTAB   73d 9h39m     6536        0        599490   0
  192.65.184.4      513          ESTAB   79d 0h59m     0           0        606025   0
  192.65.184.24     513          ESTAB   14d15h35m     1           0        606025   0
  192.65.184.138    32361        ESTAB   73d10h 4m     5           0        28       0
  192.65.184.150    2603         ESTAB   109d 0h58m    980         0        28       0
  192.65.184.210    559          ESTAB   73d10h 4m     13524       0        28       0
  192.65.184.214    293          ESTAB   109d 0h58m    130         0        28       0
  192.65.184.218    559          ESTAB   73d10h 4m     13524       0        28       0
  192.65.184.221    559          ESTAB   73d10h 4m     13524       0        28       0
  192.65.196.3      513          ESTAB   109d 0h58m    11          0        1        0
  192.65.196.4      513          ESTAB   4d 0h45m      9           0        1        0
  192.65.196.5      513          ESTAB   109d 0h58m    9           0        1        0
  192.65.196.6      513          ESTAB   4d 0h45m      11          0        1        0
  193.51.191.214    2200         ESTAB   106d10h32m    148         0        28       0
  193.159.166.221   3320         ESTAB   73d10h 4m     598120      0        28       0
```

Listing 2: BGP Zusammenfassung des BGP-Routers EE1 am CERN [38]

Letztendlich sind die Betreiber der AS die Akteure, welche das Routing bestimmen. Die Kriterien, nach welchen die Routen gewählt werden, sind die beschriebenen BGP-Attribute. Jeder Betreiber hat

dabei die Option das Routing zu beeinflussen. Anhand von ausgewählten und definierten Policies kann der Routing-Prozess des BGP verändert werden. Routing-Policies in BGP werden durch das Setzen von lokalen Präferenzen (LOCAL PREFERENCE), das gezielte Ausfiltern von Pfaden zu Zielnetzen und die Verwendung von COMMUNITY-Markierungen implementiert und umgesetzt [36].

2.5 Ob und wie kann für einzelne Kommunikationsvorgänge Rückschluss auf Ursprungs- und / oder Zielort der übermittelten Kommunikation gezogen werden?

Kommunikationsvorgänge zwischen jeweils zwei Endpunkten A und B erfordern, dass jedem Endpunkt eine IP-Adresse IP_A bzw. IP_B zugewiesen wird. Kommunikationsinhalte im Sinne der Nutzdaten der jeweils eingesetzten Anwendung, die eine beliebige Größe (Datenvolumen, z.B. mehrere Megabytes oder Gigabytes) annehmen können, werden in IP-Pakete aufgeteilt, deren Einzelgröße durch die eingesetzte Übertragungstechnik begrenzt wird (z.B. 1.500 Bytes). Damit die Transitsysteme (Router) auf dem Transportweg zwischen A und B wissen, an welches Ziel das IP-Paket geleitet werden soll und an welchen Absender ggf. Fehlerinformationen (z.B. bei Netzausfällen oder Überlast) gesandt werden sollen, enthält jedes einzelne IP-Paket die Angabe der Quell- und der Ziel-IP-Adresse. Für alle IP-Pakete eines (technischen) Kommunikationsvorgangs bleiben die Quell- bzw. Ziel-IP-Adressen unter Berücksichtigung der Kommunikationsrichtung ($A \rightarrow B$ bzw. $B \rightarrow A$) gleich.

Rückschlüsse auf Ursprungs- und Zielorte können folglich anhand der Quell- und Ziel-IP-Adressen wie folgt gezogen werden:

- Jede IP-Adresse ist einem Autonomen System (AS, siehe Einleitung und Abschnitt 2.1) zugewiesen, das wiederum einer administrativen Instanz im Sinne einer Organisation oder eines Unternehmens zugeordnet ist. Die AS-Registrierungsdaten lassen somit einen groben Rückschluss auf den jeweiligen Ort zu. So kann das bereits als Beispiel genutzte Deutsche Forschungsnetz (DFN) eindeutig dem Land Deutschland, aber offensichtlich keiner konkreten einzelnen Region, Stadt oder Anschrift zugeordnet werden, da der DFN-Verein eine über ganz Deutschland verteilte Kommunikationsinfrastruktur betreibt.

- Jede IP-Adresse kann einem der vergebenen IP-Adressbereiche, die oft kleiner sind als ganze AS, zugeordnet werden. Über das Internet-Protokoll whois (vgl. Abschnitt 2.8) kann beispielsweise abgefragt werden, dass die IP-Adresse 46.243.122.50, unter welcher der Webserver www.bundestag.de zu erreichen ist, dem IP-Adressbereich 46.243.122.0/24 zugeordnet ist, der auf die Babiel GmbH in Düsseldorf registriert ist.

- Da die beiden genannten Möglichkeiten offenkundig nur ggf. unzureichend präzise Auskünfte liefern und insbesondere die Registrierungsanschriften nicht mit den Betriebsorten der Kommunikationsendpunkte übereinstimmen müssen, wird verstärkt auf IP-Geolokalisations-Datenbanken zurückgegriffen. Eine Abfrage derartiger Datenbanken anhand einer IP-Adresse liefert eine mögliche geografische Position. Abschnitt 2.8 erläutert das Verfahren und seine Qualitäts- und Zuverlässigkeitsaspekte vertiefend.

- Die Weiterleitung von IP-Paketen über Transitsysteme erfolgt auf Basis des Austausches von Routing-Informationen (siehe Abschnitt 2.1). Diese können auch von außen ausgewertet werden, um den Weg eines IP-Pakets über die einzelnen Transitsysteme nachzuvollziehen. Sofern für Transitsysteme, die sich in unmittelbarer logischer Nachbarschaft zur Quell- bzw. Zieladresse eines IP-Pakets (first/last hop) befinden – analog zum Briefkasten, in den ein Brief

am Absendeort eingeworfen wird, bzw. das zuständige Zustellpostamt auf Empfängerseite – zuverlässige Ortsinformationen vorliegen, lässt sich darauf rückschließen, dass die Quell- bzw. Ziel-IP-Adresse mit hoher Wahrscheinlichkeit im jeweiligen „Einzugsbereich“, der sich durch die Verkabelungs- oder Funkreichweite und die damit physikalisch bedingten Übertragungslatenzen ergibt, liegt.

Neben diesen inhärent verfügbaren IP-Paket-Metadaten können Inhalte der übertragenen Nutzdaten ausgewertet werden, um auf den Standort der Endpunkte zu schließen, beispielsweise in Webserver-Anfrage eingebettete GPS-Koordinaten des vom Anwender genutzten mobilen Endgerätes. Die hierfür in Frage kommenden Daten und ihre Zuverlässigkeit, u.a. im Hinblick auf Fälschungssicherheit, sind ggf. separat zu untersuchen.

Bei allen genannten Ansätzen ist zu berücksichtigen, dass sie lediglich technische Kommunikationsvorgänge zwischen zwei IP-Endpunkten abdecken, wie sie z.B. bei der direkten Kommunikation zwischen zwei Endgeräten auftreten. Praktisch ergeben sich daraus signifikante Einschränkungen:

- Reale Kommunikationsvorgänge z.B. zwischen zwei natürlichen Personen können über beliebig viele Zwischenstationen im Internet laufen. Wird beispielsweise eine Nachricht einer Person aus den USA an eine Person in Großbritannien über einen (z.B. bei einem Hosting-Provider angemieteten) Internet-Server (hier: IP-Proxy) in Deutschland verschickt, so liegen auf technischer Ebene zwei getrennte Kommunikationsvorgänge ($USA \rightarrow DE$ und $DE \rightarrow GB$) vor, die ohne geeignete Korrelation nicht als ein Kommunikationsvorgang ($USA \rightarrow GB$) identifiziert werden können. Durch den Einsatz von virtuellen privaten Netzen (VPN) oder Anonymisierungsdiensten (z.B. Onion-Routing wie TOR) können die realen Ursprungs- und Zielorte verschleiert werden. In der Praxis ist der Einsatz von Verschleierungsmaßnahmen derzeit mit Einschränkungen (z.B. zeitliche Verzögerungen und reduzierte Datenübertragungsrate) verbunden; theoretische Modelle ermöglichen jedoch eine zuverlässige Verschleierung auch bei einer globalen Überwachung des gesamten Internet-Datenverkehrs [39, 40].

- Mehrere voneinander unabhängige natürliche Personen können gleichzeitig dieselbe IP-Adresse für Kommunikationsvorgänge verwenden. Ein triviales Beispiel ist ein Hotel, das allen seinen Gästen WLAN-/Internet-Zugang über einen einzigen DSL-Anschluss (und damit eine einzige IP-Adresse) anbietet. Eine Differenzierung einzelner Personen erfordert dann die Auswertung weiterer Daten, z.B. die in Anfragen an Webserver eingebetteten Informationen über das verwendete Endgerät (z.B. Betriebssystem, Browser-Software, Bildschirmauflösung). Die Zuverlässigkeit verringert sich dabei mit zunehmender Personenanzahl signifikant.

Zusammenfassend ist eine Identifikation der Ursprungs- und Zielorte nur grob granular und bei nicht eingesetzten Verschleierungsmaßnahmen möglich. Aus der Ortsinformation (z.B. Hotel X im Ort Z) kann allgemein nicht auf weitere Eigenschaften der Kommunikationspartner (z.B. Nationalität) geschlossen werden.

2.6 Inwiefern kann man sich zur Auswahl konkreter Übertragungstrecken für eine Ausleitung gemäß §27 Abs. 2 TKÜV bzw. zur strategischen Überwachung von Ausland-Ausland-Telekommunikationsverkehren auf Wahrscheinlichkeiten bestimmter regionaler Zuordnung von Kommunikationsvorgängen auf diesen Strecken stützen und wie sind solche Wahrscheinlichkeiten gegebenenfalls zu berechnen oder zu quantifizieren?

Bei jeglicher Vorgehensweise zur Auswahl konkreter Übertragungsstrecken für eine Ausleitung bzw. zur strategischen Überwachung von Ausland-Ausland-Telekommunikationsverkehren sind folgende Aspekte zu berücksichtigen:

- Übertragungsstrecken sind i.A. nicht für den Ausland-Ausland-Telekommunikationsverkehr dediziert, d.h. jede gewählte Übertragungsstrecke kann auch Inland-Ausland- und ggf. Inland-Inland-Telekommunikationsverkehr transportieren.
- Übertragungsstrecken sind nicht statisch. Abgesehen von Router-internen Optimierungen wird prinzipiell für jedes IP-Paket neu entschieden, über welchen Weg es zum Ziel geleitet werden soll. Dies bedeutet einerseits, dass ggf. nicht alle zu einem Kommunikationsvorgang gehörenden IP-Pakete über denselben Pfad übertragen werden, auf dem eine oder mehrere Übertragungsstrecken ausgeleitet werden. Andererseits können sich die Anteile an Ausland-Ausland- bzw. Inland-Ausland- und Inland-Inland-Telekommunikationsverkehr jederzeit ändern, sofern dies nicht durch andere Maßnahmen im In- und Ausland verhindert wird.
- Wie in Abschnitt 2.5 beschrieben können Verschleierungsmaßnahmen eingesetzt werden, die einen Ausland-Ausland-Kommunikationsvorgang z.B. in einen Ausland-Inland- bzw. einen Inland-Ausland-Kommunikationsvorgang umwandeln.

Die Auswahl konkreter Übertragungsstrecken impliziert, dass eine flächendeckende Ausleitung z.B. aus naheliegenden rechtlichen und ökonomischen Gründen nicht in Frage kommt und somit das Verhältnis aus erzieltem Nutzen zur Anzahl ausgeleiteter Übertragungsstrecken zu maximieren ist. Die folgenden Ausführungen fokussieren auf die strategische Überwachung von Ausland-Ausland-Telekommunikationsverkehr, gelten aber analog für ggf. im Inland zu betrachtende Ursprungs- oder Zielorte von Kommunikationsvorgängen.

Das nachfolgend vorgestellte Auswahlverfahren zielt auf die Maximierung des Volumens an ausgeleitetem Ausland-Ausland-Telekommunikationsverkehr (Quantität) unter effizienter Ressourcennutzung ab. Bei einer konkreten Umsetzung einer Auswahl sollten möglichst zusätzliche qualitative Informationen berücksichtigt werden, da die Weiterverarbeitung von ausgeleitetem Telekommunikationsverkehr i.A. wesentlich stärker von der Qualität als der Quantität der zu verarbeitenden Daten profitiert.

Ausland-Ausland-Telekommunikationsverkehr ist dadurch charakterisiert, dass er zwei „Grenzübergänge“ („Einreise“ und „Ausreise“) zu vollziehen hat. Durch eine hinreichende Überwachung der Auslandsübergänge ist eine Überwachung des rein inländischen Telekommunikationsverkehrs im Kontext der Frage somit obsolet. Im Zusammenhang mit den genannten Verschleierungsmaßnahmen ist hingegen der Ausland-Inland-Telekommunikationsverkehr, welchem entsprechender Inland-Ausland-Telekommunikationsverkehr zugeordnet werden kann, relevant und ggf. von besonderer Qualität. Hierzu ist zwingend eine Übertragungsstrecken übergreifende Korrelation erforderlich.

Wie in den Abschnitten 2.5 und 2.8 dargelegt, ist anhand von Verfahren wie der IP-Geolokalisation eine hinreichend zuverlässige Entscheidung, ob sich die Quell- oder Ziel-IP-Adresse eines IP-Pakets

im Ausland befindet, vollautomatisiert möglich, sofern keine Verschleierungsmaßnahmen ergriffen wurden. Beispielsweise muss durch manuelle, stichprobenartige Kontrollen die Qualität und somit die Zuverlässigkeit der dafür eingesetzten Datenbestände kontinuierlich sichergestellt werden. Die Wahrscheinlichkeiten bestimmter regionaler Zuordnung sind dabei ausschließlich von der Qualität der eingesetzten Datenbestände abhängig.

Somit beschränkt sich die Auswahl auf Übertragungsstrecken, die entweder (1) ein möglichst hohes absolutes Volumen oder (2) einen möglichst hohen relativen Anteil an Telekommunikationsverkehr transportieren, dessen Ursprungs- und/oder Zielort dem Ausland zugeordnet werden kann. Hierbei ist zu berücksichtigen, dass auch Inland-Inland-Telekommunikationsverkehr ggf. „Umwege“ über ausländische Transitstationen nehmen kann, sodass der entsprechende Telekommunikationsverkehr bei der Überwachung der ausgewählten Übertragungsstrecken ausgefiltert werden muss.

Die Quantifizierung des Volumens bzw. des Anteils an Auslands-Auslands-Datenverkehr erfordert eine rudimentäre statistische Auswertung des Netzverkehrs an denjenigen Transitstationen, die Übertragungsstrecken vom bzw. ins Ausland bedienen. Hierfür eignet sich beispielsweise die in Abschnitt 2.2 beschriebene Flow-Analyse-Technik, bei der die eingesetzten Telekommunikationskomponenten u.a. Informationen über die Quell- und Ziel-IP-Adresse von IP-Paketen sowie das übertragene Datenvolumen liefern, ohne dass hierfür der Datenverkehr der entsprechenden Übertragungsstrecke ausgeleitet werden muss:

- Das absolute Volumen an relevantem Telekommunikationsverkehr ergibt sich aus der Summe der Flow-Volumina, von denen Quell- und/oder Ziel-IP-Adresse dem Ausland zugeordnet werden.
- Der relative Anteil ergibt sich durch die Division des ermittelten absoluten Volumens an relevantem Telekommunikationsverkehr zum Volumen des gesamten Datenverkehrs.

Weitere Ausführungen zur Geolokalisierung von IP-Adressen und deren Zuverlässigkeit sind Gegenstand von Abschnitt 2.8.

2.7 Welche Charakteristika existieren im Rahmen des paketvermittelten Telekommunikationsverkehrs, die eine nationale Zuordnung von Kommunikationsvorgängen in, von und nach Deutschland innerhalb der AS erlauben und mit welchen Maßnahmen sowie mit welcher Genauigkeit lassen sich solche Zuordnungen treffen?

Im Kontext des Wortlauts der Fragestellung ist zu bedenken, dass AS (vgl. Einleitung und Abschnitt 2.1) allgemein nicht die beste Einheit sind, anhand derer eine nationale Zuordnung von Kommunikationsvorgängen vorgenommen werden sollte: Wie bereits erläutert stellen AS zunächst lediglich Ansammlungen von IP-Netzen unter einer gemeinsamen Verwaltung dar. Dies lässt Konstellationen zu, in denen eine multinationale Organisation ein AS z.B. mit der Anschrift ihres Hauptsitzes registriert, innerhalb des AS jedoch mehrere IP-Netze betreibt, die in mehreren Ländern verortet und beispielsweise nicht über das Internet, sondern anderweitig betriebene Übertragungsstrecken AS-intern verbunden sind. Die nationale Zuordnung des gesamten AS zu einem Land wäre in diesem Fall irreführend. Eine Mehrfachzuordnung zu allen in Frage kommenden Ländern wäre hingegen unpräzise. Jedoch können Kommunikationsvorgänge auftreten, bei denen IP-Pakete von Absendern zu Empfängern im selben Land aus rein technisch-ökonomischen Gründen über Transitsysteme im Ausland transportiert werden. Dieses kann ggf. mit einer intuitiven nationalen Zuordnung des Kommunikationsvorgangs schwierig zu vereinen sein.

Somit empfiehlt sich eine Betrachtung auf zwei anderen Ebenen:

1. Auf physischer Ebene, d.h. durch Betrachtung der Verkabelungsinfrastruktur wie z.B. Lichtwellenleiter-Kabeltrassen, lässt sich für einen in Deutschland mündenden Anschluss entscheiden, ob die Gegenstelle ebenfalls in Deutschland oder im Ausland liegt. Damit ist zwar keine nationale Zuordnung einzelner Kommunikationsvorgänge möglich, transnationale Kommunikationsvorgänge müssen aber offensichtlich zwingend über transnationale Übertragungsstrecken abgewickelt werden. Somit kann eine Fokussierung auf Übertragungsstrecken erfolgen, die für eine weitere Auswertung der zuordnungsrelevanten Charakteristika in Frage kommen.

2. Die durch das Routing innerhalb der und zwischen den AS vermittelten IP-Pakete sind durch ihre Quell- und Ziel-IP-Adressen charakterisiert. Diese IP-Adressen ermöglichen eine in der Regel zuverlässige geografische Zuordnung auf Landesebene (vgl. Abschnitten 2.5 und 2.8). Zur Zuordnung konkreter Kommunikationsvorgänge ist deshalb zu empfehlen, die am Kommunikationsvorgang beteiligten IP-Adressen und nicht nur die AS der Kommunikationsendpunkte zu betrachten.

Komplementär zu dieser Betrachtung verbleibt die noch unbeantwortete Teilfragestellung, wie eine nationale Zuordnung bei Datenverkehr innerhalb eines AS erfolgen kann. Auch in diesem Fall lässt sich zunächst wieder auf den Mechanismus der IP-Geolokalisation verweisen, sofern die beiden Charakteristika „Quell-“ und „Ziel-IP-Adresse“ konkreter Kommunikationsvorgänge verfügbar sind, beispielsweise durch die Ausleitung von Übertragungsstrecken oder die Bereitstellung von Flow-Daten (vgl. Abschnitt 2.2). Von der Qualität der zur IP-Geolokalisation genutzten Datenbestände hängt die Präzision des Verfahrens direkt ab. Da der Datenverkehr innerhalb eines AS jedoch nach individuellen Vorgaben (Routing-Einstellungen), der für das AS zuständigen Verwaltung erfolgt, und den anderen AS gegenüber i.A. nicht offengelegt wird bzw. von außen zugänglich ist, existiert keine im Sinne ihrer universellen Anwendbarkeit pauschale definierte Maßnahme. Bei AS, bei denen eine multinationale Ausdehnung anzunehmen ist, die eine präzise Unterscheidung der beteiligten Länder erforderlich macht, aber durch IP-Geolokalisationsdatenbestände noch nicht hinreichend beschrieben ist, sind deshalb ebenso individuelle Maßnahmen zu planen und umzusetzen, die zu einer ausreichend detaillierten Offenlegung der AS-internen Routing-Einstellungen und bedienten Standorte führen.

Insgesamt ergibt sich der Bedarf, als Voraussetzung für die nationale Zuordnung von Kommunikationsvorgängen zunächst die transnationalen Übertragungsstrecken zu identifizieren, über welche die erforderlichen Charakteristika z.B. für die IP-Geolokalisation zugänglich gemacht werden können. Auf organisatorischer Ebene können diese Informationen von inländischen AS- und IXP-Betreibern offengelegt werden, da i.A. die genauen Orte von Gegenstellen, zu denen kabelgebundene Verbindungen eingerichtet werden, bekannt und dokumentiert sind (z.B. in Peering-Abkommen) oder zumindest, z. B. anhand des Verlaufs von Kabeltrassen über Ländergrenzen, nachverfolgt werden können. Zudem können die öffentlichen Internet-Routing-Informationen (siehe Abschnitt 2.1) ausgewertet werden, die darüber Auskunft geben, welche AS im Sinne der Paketvermittlung unmittelbar benachbart sind. Für jedes der so identifizierten Nachbar-AS kann die nationale Zuordnung wie oben beschrieben durchgeführt werden, um zu entscheiden, ob die über den durch die beiden AS definierten Routing-Pfadabschnitt abgewickelten Kommunikationsvorgänge einen transnationalen Übergang beinhalten.

2.8 Welche praktisch erprobten und/oder wissenschaftlich anerkannten Methoden zur ländergenauen Geolokalisierung von IP-Adressen bzw. IP-Datenpaketen gibt es (online und offline) und wie zuverlässig sind diese Methoden für eine Zuordnung zum Standort „Deutschland"?

IP-Geolokalisation bezeichnet die Zuordnung einer logischen Adresse, beispielsweise der IP-Adresse eines Hosts (Endgerätes), zu einer physikalischen respektive geografischen Position [41]. Im Forschungsgebiet der Geolokalisation von IP-Adressen haben sich verschiedene Ansätze herausgebildet, welche sich gemäß Endo et. al [42] grundsätzlich in zwei wesentliche Kategorien einordnen lassen [43]:

- Verfahren basierend auf aktiven Messungen;
- Passive bzw. semantische Methoden.

Grundsätzlich gehen alle Einteilungen, unter anderem Dahnert [44] und Eriksson [45], ebenfalls in diesen beiden Kategorien auf. Als eine mögliche dritte Kategorie zählen sogenannte Hybrid Ansätze, welche aus einer Kombination von aktiven und passiven Methoden bestehen und dadurch versuchen die Schwächen der jeweiligen Verfahren zu überwinden. Allerdings basieren alle derzeit bekannten hybriden Verfahren hauptsächlich auf aktiven Messungen und lassen sich somit ebenfalls hier einordnen [42].

Verfahren basierend auf aktiven Messungen

Zur Ermittlung der möglichen geografischen Position eines Hosts werden aktive Messungen herangezogen. Das heißt, es erfolgt eine direkte Interaktion mit dem Zielsystem, wobei typischerweise Verzögerungswerte, beispielsweise Round Trip Times (RTT) [41] oder mittels HTTP [46], bestimmt werden. Diese werden über Vergleiche mit Messungen bekannter Server-Standorte in Relation gesetzt, umso auf die Position des Zielhosts zu schließen [47]. All diese Verfahren nutzen für diese Messungen sogenannte Landmarks. Landmarks sind Hosts mit bekanntem geografischen Standort, die entweder passiv auf Anfragen reagieren oder aktiv für ausgehende Messungen zur Verfügung stehen [41].

Die Korrelation zwischen Latenzwerten und geografischen Distanzen ist ein fundamentaler Bestandteil all dieser Verfahren. Entgegen konventioneller Meinungen, dass eine solche Korrelation nicht vorhanden ist [48], bestätigen Ziviani et al. [47] eben deren Existenz. Jedoch ist diese zu schwach ausgeprägt, um sie in einem mathematischen Model zu formulieren [49].

Nach Eriksson et al. [50] besteht der Prozess zur Latenz-basierenden IP-Geolokalisation eines Zielsystems t aus folgenden grundlegenden Schritten, welche je nach Ansatz angepasst bzw. abgewandelt werden:

1. Aufbau einer Menge an Landmarks $\mathcal{L}$ mit n unterschiedlichen Landmarks $L_{1..n} \in \mathcal{L}$.
2. Bestimmung der Latenz[3] r zwischen Zielsystem t und jedem Landmark L_i aus der Menge $\mathcal{L}$.
3. Berechnung der geografischen Distanz d auf Basis der Korrelation zwischen gemessener Latenz r und geografischen Strecken.
4. Lokalisation des Ziels t basierend auf den bekannten Standorten der Landmarks $L_i \in \mathcal{L}$ sowie den berechneten Distanzen d.

[3] Zumeist basierend auf mehrfachen Latenzmessungen für die einfache Strecke und der Auswahl der minimalen Latenz aus der Menge aller Messungen [50, 41]

Eine Auswahl gängiger Verfahren, die zudem die Grundlage für eine Vielzahl aktueller Ansätze bilden, wird im Folgenden kurz erläutert:

- *Shortest Ping* [51] ist ein Verfahren basierend auf Latenzmessungen (hier: RTT), wobei als geografischer Standort das Landmark angenommen wird, was in Bezug auf die Latenz am nächsten zum zu lokalisierendem Ziel ist.
- *GeoPing* [41] ist eines der ersten Verfahren basierend auf aktiven Messungen. Es werden Latenzmessungen zwischen verschiedenen Landmarks sowie von diesen zum Ziel durchgeführt, um anhand der so gewonnen Daten sogenannte Latenzvektoren aufzustellen. Als physikalischer Standort wird hierbei das Landmark mit dem ähnlichsten Muster in Bezug auf das Ziel angenommen.

Verfahren basierend auf aktiven Messungen können weiterhin hinsichtlich von Randbedingungen (Constraints) und der Nutzung der zugrunde liegenden Netztopologie unterschieden werden:

- *Constraint-Based Geolocation (CBG)* [52] bestimmt die geografische Position des Zielsystems anhand von Multilateration [53, 54] und Randbedingungen in Bezug auf die Distanzen. Somit wird ein kontinuierlicher Lösungsraum aufgespannt.
- *Topology-Based Geolocation (TBG)* [51] basiert auf CBG, nutzt aber zusätzlich Informationen über die Topologie des zugrunde liegenden Netzes. Hierbei wird versucht Router auf dem Pfad zum Zielhost zu lokalisieren.
- *Octant* [55] ist ein Framework zur IP-Geolokalisation und ist, ebenso wie die hierauf basierenden Ansätze, Stand der Technik in diesem Forschungsgebiet. Durch das modulare Design können Areale beispielsweise durch demographische Daten weiter eingeschränkt und somit die Genauigkeit der Lokalisation erhöht werden [46, 56]. Octant ist ein typischer Vertreter der hybriden Ansätze.

Weitere Beispiele für hybride Verfahren sind Spotter [54], HawkEyes [44] und POSIT [45].

Passive bzw. semantische Methoden

Semantische Verfahren basieren auf der Extraktion standortbezogener Informationen zu einer gegebenen Adresse mit Hilfe von beispielsweise umfangreichen Datenanalysen. Diese Ansätze sind passiv und basieren zumeist auf Abfragen öffentlich zugänglicher Datenbestände, d.h. es findet keine Interaktion mit dem Zielsystem statt. Somit werden Abfragen in nahezu Echtzeit ermöglicht wodurch sie wiederum für große Datenmengen geeignet sind.

Typische grundlegende Beispiele dieser Kategorie werden im Folgenden kurz erläutert:

- *Datenbestände der Regional Internet Registries und Domain Registrare*
  Der gesamte IP-Adressraum wird hierarchisch von fünf RIR verwaltet, welche wiederum ihrerseits Adressbereiche an verschiedene Local Internet Registries (LIR), National Internet Registries (NIR) und ISP delegieren (vgl. Abb. 2.8). Die grundlegende Zuweisung erfolgt zentral über die IANA anhand des geschätzten geografischen Einsatzbereiches der IP-Adressen.

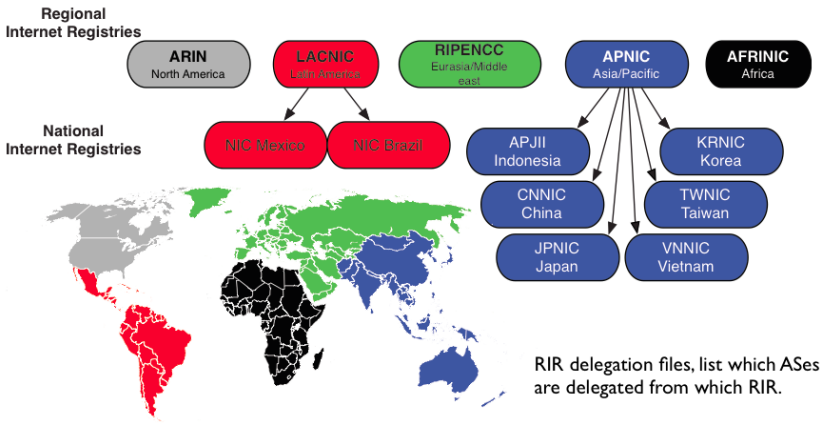


Abbildung 8: Hierarchische Struktur und geografische Zuständigkeit der RIR [57]

Die Datenbestände der einzelnen RIR sowie auch Domain Registrare, wie beispielsweise die DENIC, können mittels whois [58] abgefragt werden. Anhand dieser Daten können Rückschlüsse auf Zuständigkeiten und geografische Standorte gezogen werden (vgl. Listings 3 und 4).

```
query@whois:~$ whois 137.193.6.6
...
inetnum:        137.193.0.0 - 137.193.255.255
netname:        UNIBWMNET
descr:          Werner-Heisenberg-Weg 39, D-85579 Neubiberg
...
country:        DE
...
person:         Ludwig Bayer
address:        Universitaet der Bundeswehr Muenchen
address:        Rechenzentrum
address:        Werner-Heisenberg-Weg 39
address:        85579 Neubiberg
address:        Germany
phone:          +49 89 6004 3219
...
```

Listing 3: whois Abfrage RIR (gekürzt)

```
query@whois:~$ whois unibw.de
% Copyright (c) 2010 by DENIC
...
Domain: unibw.de
Nserver: dns01.rz.unibw-muenchen.de
Nserver: dns03.rz.unibw-muenchen.de
Nserver: ws-kar1.win-ip.dfn.de
...
Name: Stefan Schwarz
Address: Universitaet der Bundeswehr
Address: Rechenzentrum
Address: Werner-Heisenberg-Weg 39
PostalCode: 85579
City: Neubiberg
CountryCode: DE
Phone: +49 89 6004 3200
...
Email: stefan.schwarz@unibw-muenchen.de
...
```

Listing 4: whois Abfrage DNS (gekürzt)

- *Fully Qualified Domain Names (FQDN)*
  Bei dieser Methode wird versucht die geografische Position des Zielsystems mittels Analyse

des FQDN dieses Hosts zu bestimmen. Zusätzlich werden Systeme, die in unmittelbarer topologischer Verbindung zum Zielsystem stehen, ebenfalls untersucht [59, 41]. Dies ist vor allem dann von Vorteil, falls das Ziel keinen FQDN besitzt oder eine Auflösung des gesamten Pfades zum Ziel nicht möglich ist [51]. Grundlage für dieses Vorgehen sind Best Practices von Netzbetreibern einzelne Knoten mit Hinweisen auf Funktion und Standort zu versehen [41]. Hierfür werden häufig international standardisierte Länder-, Städte- oder auch International Air Transport Association (IATA) Codes verwendet [60, 41].

```
trace@fqdn:~$ tracepath vangogh.cs.berkeley.edu
...
 2:  winruta.RZ.UniBw-Muenchen.de                          1.798ms asymm  3
 3:  cr-garl-te0-7-0-2.x-win.dfn.de                        2.995ms
 4:  cr-fra2-hundredgige0-0-0-3.x-win.dfn.de              11.678ms
 5:  dfn.mx1.fra.de.geant.net                             11.445ms
 6:  internet2-gw.mx1.fra.de.geant.net                    98.574ms
...
15:  soda-10g-edge.EECS.Berkeley.EDU                      188.044ms
16:  soda-288-aggrt-229.EECS.Berkeley.EDU                 381.129ms
17:  vangogh.CS.Berkeley.EDU                              178.810ms reached
```

Listing 5: Analyse FQDN mittels Tracerouting (gekürzt)

- *Geoservice Anbieter*
  Die Nutzung von Geoservice-Anbietern bzw. Geodatenbanken wie MaxMind [61] hat sich im Verlauf der letzten Jahre zur vorherrschenden Methode der Lokalisation von IP-Adressen entwickelt [54, 62]. Das Angebot reicht von freien und kommerziellen Datenbanken bis hin zu Abfrage[4] basierenden Bezahlmodellen. Die Zuordnung einer IP-Adresse zu einer geografischen Position erfolgt über den Abgleich der angefragten Adresse beziehungsweise des Domain-Namens mit den Datenbeständen der jeweiligen Anbieter, wobei Umfang und Genauigkeit variieren [62, 63, 64, 65, 57].

- *Domain Name System (DNS)*
  Neben der bereits dargestellten Abfrage einer Domain mittels whois, besteht zudem die Möglichkeit DNS-Server direkt nach geografischen Informationen abzufragen. RFC 1876 [66] beschreibt einen experimentellen Ansatz (Sub)Netz- oder auch Host-basierend unter anderen Längen- und Breitengrade als sogenannten Resource Record im DNS zu hinterlegen. Ein DNS Resource Record beschreibt den Typ eines DNS-Eintrages. Beispielsweise steht A für einen Hosteintrag IP Version 4 sowie AAAA für IP Version 6 (vgl. Listing 6).

```
query@dns:~$ dig www.bundestag.de
...
;; QUESTION SECTION:
;www.bundestag.de.                IN      A

;; ANSWER SECTION:
www.bundestag.de.        1809    IN      A       46.243.122.50
...
;; ADDITIONAL SECTION:
anycast1.irondns.net.    71451   IN      A       195.253.64.5
anycast1.irondns.net.    71451   IN      AAAA    2a01:5b0:4::5
...
```

Listing 6: DNS Resource Record Abfrage mittels dig

[4] Beispielsweise mittels Application Programming Interfaces (API) oder Webschnittstelle

Listing 7 zeigt die Abfrage eines solchen LOC Resource Records, welcher neben Breiten- bzw. Längengrad (32 53 N 117 14 W) auch Angaben zur Höhenlage (107.00m) und Präzision (30m 10m 10m) enthält.

```
query@dns:~$ dig -t LOC www.caida.org
...
;; QUESTION SECTION:
;www.caida.org.                  IN      LOC

;; ANSWER SECTION:
www.caida.org.   600    IN   CNAME      ciderv6.caida.org.
ciderv6.caida.org.   600  IN   LOC   32  53  N  117  14  W  107.00m  30m  10m  10m
...
```

Listing 7: Abfrage DNS LOC Resource Record für einen bestimmten Host mittels dig

- *Clustering*
  Clustering Verfahren wie beispielsweise GeoCluster [41] teilen den IP-Adressraum in Blöcke auf und versuchen basierend auf der Annahme, dass alle IP-Adressen eines Clusters sich geografisch im selben Areal lokalisieren lassen, jedem dieser Cluster eine geografischen Region zuzuordnen. Um einen entsprechenden Datenbestand aufzubauen, können grundsätzlich alle in diesem Kapitel aufgeführten Verfahren genutzt werden. Zusätzlich werden BGP-Routingtabellen und Information zu den AS herangezogen [67, 42, 41, 59].

- *IPv6 Extension Header*
  Im Bereich IPv6 wurde 2013 eine Idee zu einem neuen Extension Header IP-LOC vorgestellt [68]. Dieser Header ermöglicht das Vorhalten von GPS Koordinaten mit Nutzer spezifischem Genauigkeitsgrad.

Bewertung gängiger Verfahren

In Tabelle 1 werden gängige Verfahren, unabhängig ihrer Verbreitung, zur IP-Geolokalisation gegenüber gestellt. Um die Übersichtlichkeit zu gewährleisten, erfolgt die Einteilung in aktive, passive sowie hybride Verfahren. Die Spalten beschreiben neben der Bezeichnung des Verfahrens, inwiefern Landmarks benötigt werden, welche IP-Version unterstützt wird, online oder offline Nutzbarkeit sowie ob der Lösungsraum stetig oder diskret ist. Die letzte Spalte beschreibt inwiefern das Verfahren grundsätzlich kommerziell oder frei verfügbar, d.h. nicht kostenpflichtig ist. Alle Angaben beziehen sich auf das ursprüngliche Verfahren und dessen grundlegende Nutzbarkeit sowie die Technik. Je nach Erweiterung lassen sich diverse Verfahren kombinieren und erweitern, um so mögliche Schwächen zu reduzieren. Dies ist in der folgenden Tabelle 1 nicht berücksichtigt.

Wie in Tabelle 1 dargestellt, sind bis auf den IPv6 IP-LOC Extension Header alle Verfahren für IPv4 und IPv6 geeignet. Allgemein kann festgehalten werden, dass für die Nutzung im Rahmen von großen Datenmengen nur passive Methoden als nutzbar erachtet werden können. Die Lokalisation von IP-Adressen mittels aktiver Messungen ist grundsätzlich genauer, allerdings auch ressourcenintensiver und daher nur für einzelne IP-Adressen in Erwägung zu ziehen. Hinzu kommt, dass der Aufbau und das Betreiben der Messinfrastruktur inklusive einer ausreichenden Menge an Landmarks mit vermehrten personellen und finanziellen Ressourcenbedarf verbunden ist. Gemäß CAIDA [69] konnten vereinzelt aktive Lokalisierungen bis auf 600 Meter genau durchgeführt werden [56]. Diese Ergebnisse basieren allerdings auf Annahmen, z.B. dass Firmen ihre Webserver In-House hosten. Diese Annahme ist in Zeiten von Outsourcing und Cloud Computing vor allem in Europa nicht haltbar. Außerdem sind solche Ergebnisse zumeist in homogenen (Forschungs-)Netzen erzielt und somit nicht ohne Weiteres für Szenarien in der realen Welt übertragbar. Die meisten Forschungsarbeiten gehen

Tabelle 1: Übersicht gängiger IP-Geolokalisationsverfahren (basierend auf [65])

VERFAHREN	LANDMARKS BENÖTIGT	PASSIV/ AKTIV/HYBRID	LÖSUNGSRAUM	IP Version	OFFLINE/ ONLINE	KOMMERZIELL
Geoservices	nein	passiv	diskret	4 & 6	online/offline	Anbieter abhängig
Whois (RIR/DNS)	nein	passiv	diskret	4 & 6	online	nein
Analysis von FQDN	nein	passiv	diskret/stetig	4 & 6	online	nein
DNS LOC RR	nein	passiv	diskret	4 & 6	online	nein
Clustering	nein	passiv	diskret	4 & 6	offline	nein
IPv6 IP-LOC	nein	aktiv/passiv	diskret	6	online	nein
Reine Latenz Messung (RTT)	ja	aktiv/passiv	diskret/stetig	4 & 6	online	nein
CBG/TBG/Octant u.ä.	ja	hybrid	stetig	4 & 6	online	nein

zudem davon aus, dass ihre im amerikanischen Raum erzielten Ergebnisse auch für Europa gelten, beachten hierbei nicht die komplexere Netzinfrastruktur. Auch wird nur sporadisch auf die Auswahl- und Platzierungsproblematik der benötigten Landmarks eingegangen, welche maßgeblich für die Genauigkeit der aktiven Verfahren verantwortlich sind [70].

In diesem Zusammenhang sei darauf hingewiesen, dass verbreitet die Annahme getroffen wird, dass in moderner Glasfaser-Infrastrukturen die Ausbreitungsgeschwindigkeit des Signals bei $\frac{2}{3}$ [71] bzw. $\frac{4}{9}$ [56] der Lichtgeschwindigkeit c liegt. Wenn die Ergebnisse von Hillmann et al. [70] betrachtet werden, dann sind diese Werte zu optimistisch. Für den europäischen Raum sind 20% bis 25% der Lichtgeschwindigkeit c realistisch.

Aufgrund der komplexen Netzinfrastruktur in Europa haben die deterministischen und stochastischen Anteile der Gesamtlatenz, verursacht durch beispielsweise größere Anzahl an Routern auf dem Pfad zum Ziel sowie der sogenannten "Letzten Meile", erhöhten Einfluss im Vergleich zu den USA. Tabelle 2 vergleicht einen aktuellen Ansatz [65] mit gängigen Verfahren und dem Fokus auf Europa, unter Berücksichtigung der oben erwähnten Ausbreitungsgeschwindigkeit des Lichts in Lichtwellenleitern. Zur Lokalisation einer IP-Adresse auf Landesebene sind Geodatenbanken mit 96% bis 98% Genauigkeit [63, 64, 57] ausreichend und zudem nutzbar hinsichtlich großer Abfragemengen. Gemäß der Beobachtung von Koch et al. [65] ist auch die Nutzung von whois zur Abfrage der RIR-Datenbestände vergleichbar mit Geoservice-Anbietern wie MaxMind. Die Genauigkeit von Verfahren wie DNS LOC Resource Records [66] und IP LOC Extension Header [68] ist grundsätzlich abhängig vom jeweiligen Operator. Da solche Informationen u.a. sicherheitsrelevant sein können, sind diese Ansätze zudem nicht weit verbreitet [72, 56, 68].

Tabelle 2: Vergleich von [65] mit aktiven und passiven (kursiv) IP-Geolokalisationsverfahren mit Fokus auf Europa

VERFAHREN	Dragoon [65]	*whois*	*MaxMind*	Spotter	CBG	ACBG/TBG
DURCHSCHNITTLICHE ABWEICHUNG IN KM	134	410	463	754	768	770

Alle der oben aufgeführten Verfahren lassen sich hinsichtlich Genauigkeit und Aussagekraft direkt oder indirekt negativ beeinflussen. Sobald bei aktiven Messungen Anonymisierungstechniken

wie Proxy-Server oder VPN eingesetzt werden, lässt sich im Regelfall[5] nur der Proxy-Server bzw. der VPN-Endpunkt lokalisieren. Hinzu kommt, dass in der Theorie auch eine Manipulation der aktiven Messungen durchführbar ist, wodurch beispielsweise ein anderes Land als Ausgangspunkt eines Kommunikationsvorganges vorgetäuscht werden kann [73, 74, 75].

2.9 Welche Aspekte der den einzelnen Kommunikationsvorgängen zuzuordnenden Verkehrsdaten lassen darüber hinaus eventuell Rückschlüsse darauf zu, ob ein Teilnehmer des jeweiligen Kommunikationsvorgangs Deutscher ist?

Betrachtet man zunächst rein die Metadaten eines Kommunikationsvorganges, kann (sofern keine Manipulation oder Verschleierung vorliegt) durch IP-Geolokalisation der Quelladresse mit einer hohen Wahrscheinlichkeit ermittelt werden, ob sich diese in Deutschland befindet. Weiterhin lässt sich durch Analyse der aufgerufenen Webseiten anhand der DNS-Anfragen, also der Namensauflösung beim Aufruf einer Webseite, sowie durch die Untersuchung der auf dem Server der jeweiligen Zieladresse angebotenen Inhalte, ein Rückschluss über die Sprache und Herkunft des Nutzers ziehen. Dies kann jedoch nur ein grober Anhalt sein, da viele Webauftritte verschiedene Sprachen anbieten und die genutzte Sprache nicht ohne Analyse der Datenpakete identifiziert werden kann. Somit kann eine Auswertung rein auf Metadaten immer nur Indizien liefern, ob ein Teilnehmer des Kommunikationsvorganges Deutscher ist.

Einfacher stellt sich die Situation bei einer Analyse der eigentlichen Datenpakete dar, falls diese in unverschlüsselter Form vorliegen. Hier kann die genutzte Sprache innerhalb einer Konversation zusammen mit den bereits erwähnten IP-Geolokalisationsdaten Hinweise darauf liefern, ob ein Kommunikationsteilnehmer Deutscher ist. Die Genauigkeit einer Auswertung basierend auf diesen Daten bedarf jedoch einer wissenschaftlichen Untersuchung.

Abhängig von der genutzten Anwendung, stehen gegebenenfalls noch weitere Indizien zur Verfügung, beispielsweise ein genutzter Mailprovider sowie die in diesem Kontext von einem Kommunikationsteilnehmer verwendete Email-Adresse, falls diese z.B. aus Vor- und Nachnamen besteht. Auch Telefonie- oder Chatdienste können bei Analyse der Datenpakete Hinweise durch beispielsweise die genutzte Sprache liefern. Hierzu muss jedoch gesagt werden, dass insbesondere nach den Veröffentlichungen im Rahmen der Snowden-Affäre mittlerweile zahlreiche Dienste verschlüsselt werden, eine Analyse der Datenpakete somit nicht mehr trivial möglich ist. Andere beliebte Anwendungen, wie beispielsweise Skype, verwenden Verschlüsselungsverfahren und können typischerweise nicht auf ihre Inhalte hin untersucht werden, wenn nicht Einzelfälle wie zum Beispiel der Anruf eines Skype-Kontaktes von einem Mobiltelefon oder einem nicht IP-basierten Festnetztelefon vorliegen und dies zur Untersuchung im Teil des öffentlichen Telefonnetzes genutzt werden kann.

Durch die steigende und automatische Nutzung von Verschlüsselung bei vielen Diensten, ist ein direkter Zugriff auf die Datenpakete nur noch in einer sinkenden Zahl von Anwendungen und Diensten möglich. Eine Evaluation rein auf der Basis von Metadaten kann jedoch nur schwache Indizien bezüglich der Nationalität eines Kommunikationsteilnehmers liefern.

[5] Mittels Implementierungs-, Konfigurations- oder Designfehlern lässt sich der Lokalisationsvorgang unter Umständen weiter vorantreiben

2.10 Welche Möglichkeiten gibt es für Dritte, d.h. Personen die nicht den Betreibern der AS zuzuordnen sind, den Regelbetrieb zu beeinflussen bzw. zu beeinträchtigen und welche Schutzmaßnahmen existieren ggf. gegen solche Manipulationen?

Der Regelbetrieb kann grundsätzlich auf verschiedene Weisen beeinflusst werden, u.a. durch den Angriff auf das genutzte Routing-Protokoll (BGP), durch die Ausnutzung von Schwachstellen in aktiven Komponenten der Internet-Infrastruktur (insbesondere Router) oder durch die Einschleusung und Ausnutzung von beispielsweise Hardware-basierten Hintertüren (Backdoors).

Beeinflussung und Missbrauch des Routingprotokolls BGP

Eine Möglichkeit der Beeinflussung des Regelbetriebs ist die Manipulation von BGP-Routen. An den Grenzen der AS werden die über ein AS zu erreichenden IP-Netze per BGP bekannt gegeben. Oftmals ist es hierbei der Fall, dass verschiedene Wege zum gewünschten IP-Netz führen. Die Entscheidung der Router für die letztendliche Weiterleitung basiert auf den Routingtabellen, die per BGP aktualisiert werden. Eine Schwachstelle hierbei ist, dass bei der Übermittlung von BGP-Nachrichten, also der Bekanntgabe von Erreichbarkeitsinformationen, im Allgemeinen keine Autorisierung erfolgt. Ein prominentes Beispiel für die Manipulation von Routingeinträgen ist die Umleitung des YouTube-Datenverkehrs zu Pakistan Telecom im Jahre 2008 [76]. Im Zuge einer Sperrung zu Inhalten von YouTube, die durch die Regierung Pakistans erlassen wurde, handelte Pakistan Telecom wie folgt: Es erfolgte eine Bekanntgabe des IP-Netzes von YouTube bei Pakistan Telecom via BGP, sodass innerhalb kürzester Zeit eine Vielzahl von Routern via BGP mit den fehlerhaften Routing-Informationen versorgt wurden und der Datenverkehr zu YouTube in der Folge nicht mehr korrekt weitergeleitet wurde. Die Wiederherstellung des korrekten Routings dauerte ca. 2 Stunden. Eine solche, fälschliche Bekanntgabe von fremden IP-Netzen im eigenen AS wird auch BGP Prefix Hijacking genannt. Ein weiteres Beispiel eines Angriffs auf den Internetverkehr eines Landes ist eine Aktion der Iranian Cyber Army mit dem Ziel Dänemark im Jahre 2011, bei dem der Datenfluss nach Dänemark über den Iran umgeleitet wurde.

Die Durchführung eines Denial of Service (DoS) Angriffs ist eine weitere Möglichkeit der Beeinflussung. Im Allgemeinen spricht man von einem DoS sobald ein Dienst nicht mehr nutzbar ist. Dies kann zum Beispiel durch die Überlastung eines Servers mit Anfragen herbeigeführt werden, indem eine sehr hohe Zahl von Anfragen von vielen verschiedene Systemen parallel an einen Dienst gerichtet wird. Hier spricht man entsprechend von einem Distributed Denial of Service (DDoS). Es gibt eine Vielzahl von Möglichkeiten, ein Netz mithilfe eines (D)DoS Angriffs zu beeinflussen oder zu beeinträchtigen. In Bezug auf die Weiterleitung von Daten über AS hinaus ist ein DoS Angriff zu nennen, der wiederum auf BGP abzielt. Hierbei sendet der Angreifer eine große Anzahl an neuen IP-Netzen, die über ein AS erreicht werden können, wodurch zum einen sehr viele BGP-Nachrichten generiert werden, die durch den Router zu verarbeiten sind. Zum anderen können im Router selbst die Routing-Tabellen so aufgebläht werden, dass eine effizienten Weiterleitung von IP-Paketen nur eingeschränkt oder gar nicht mehr möglich ist. (D)DoS Angriffe sind oftmals sehr auffällig und schnell zu detektieren, können jedoch erhebliche Auswirkungen haben, wie die DDoS-Angriffe auf Estland im April 2007 gezeigt haben, die drastische Auswirkungen auf beispielsweise die Arbeitsfähigkeit von Banken hatten.

Weitere Angriffsverfahren, wie sie auch aus anderen Bereichen des Internet bekannt sind, lassen sich ebenfalls auf BGP anwenden. Beispielsweise könnten unverschlüsselte Routing-Informationen auch mittels eines Man-in-the-Middle Angriffes (MITM) manipuliert werden. Detailliertere Analysen der Schwachstellen von BGP finden sich sowohl in RFC 4272, BGP Security Vulnerabilities Analy-

sis [77] als auch beispielsweise unter [78].

Um den Schutz von BGP zu verbessern, können verschiedene Verfahren und Erweiterungen genutzt werden, welche beispielsweise Authentisierungs- und Validierungsverfahren einführen, um Manipulationen zu erkennen und zu verhindern. Eine Schutzmöglichkeit wurde zum Beispiel durch das Secure Border Gateway Protocol (S-BGP) vorgestellt, welches eine sichere und skalierbare Architektur für ein Authentisierungssystem für BGP beschreibt. In der Praxis ist es jedoch oftmals schwierig und sehr langwierig, entsprechende Protokollergänzungen oder -änderungen, welche IXP-übergreifend implementiert werden müssen zu motivieren und umzusetzen.

Auch eine Verschlüsselung des Datenverkehrs, beispielsweise mittels IPSec, sowohl für IPv4 als auch IPv6, lässt sich nutzen, um die Angriffsschwelle auf BGP deutlich zu erhöhen und beispielsweise MITM-Angriffe zu verhindern.

Manipulation durch DNS-basierte Angriffe

Das Domain Name System, welches für die Auflösung von Domainnamen in die zugehörigen IP-Adressen verantwortlich ist, stellt ebenfalls eine Möglichkeit dar, den Regelbetrieb zu beeinflussen. Dadurch, dass das Protokoll nahezu überall eingesetzt wird und auch oftmals nicht im Rahmen von technischen Filtermaßnahmen wie Firewalls gefiltert wird, können sich Optionen für Angreifer eröffnen. Zwei grundlegende Arten von DNS-basierten Angriffen sind zum einen MITM-Angriffe wie bspw. das sogenannte *Cache Poisoning* [79]. Bei dieser Angriffsart werden gefälschte DNS-Adressen an den Resolver geschickt, damit sie von diesem in den DNS-Cache aufgenommen werden. Erfolgt danach eine Anfrage, die an den manipulierten DNS-Resolver geleitet wird, liefert dieser die im Cache befindlichen Daten zur Namensauflösung aus. Ein Nutzer kann somit auf einen falschen Server umgeleitet werden. Zum anderen kann ein Angriff auf die DNS-Struktur erfolgen, indem der verbindungslose DNS-Datenverkehr (UDP) umgeleitet und manipuliert wird.

Weitere Angriffe auf DNS ermöglichen beispielsweise die Ausnutzung offener, rekursiver DNS-Server zur Durchführung von DDoS-Angriffen [80] oder Domain Phishing [81]. DNS Fast Fluxing [82] Verfahren können zum Verbergen des Standortes eines Servers genutzt werden. Diese stehen hier jedoch nicht im Schwerpunkt der Fragestellung nach Manipulationsmöglichkeiten außerhalb des Betreibers des AS und werden somit nicht vertieft.

Manipulation durch Ausnutzung von Konfigurationsfehlern oder Softwareschwachstellen in Routern

Konfigurations- und Programmierfehler kommen regelmäßig vor, auch in aktiven Komponenten im Internet. Diese können bei Bekanntsein durch einen Angreifer ausgenutzt werden, um beispielsweise Routing-Informationen zu manipulieren. Entsprechende Schwachstellen müssen durch eine tiefgehende Analyse des Zielsystems gefunden werden, können jedoch auch bei diversen Firmen erworben oder im Darknet gekauft werden. Ein besonderes Beispiel ist die jüngste Veröffentlichung von Schadcode, in diesem Kontext als „Cyber-Waffen" bezeichnet, welcher durch einen Hacking-Angriff auf die sogenannte Equation Group erbeutet wurde [83]. Dieser Gruppe wurde nachgewiesen Schadsoftware, welche im Rahmen von Operationen der NSA eingesetzt wurde, entwickelt zu haben. Die veröffentlichten Daten enthielten mehrere Zeroday-Exploits u.a. für Cisco Router, also Schadcode für Schwachstellen, die dem Hersteller des jeweiligen Produkts selbst noch nicht bekannt sind und somit einen erheblichen Wert darstellen.

Eingriffe durch Ausnutzung von Hardware-Backdoors

Eine weitere Manipulationsmöglichkeit bietet das Vorhandensein von Hardware-Backdoors. Diese ermöglichen den Zugriff auf ein System, beispielsweise einen Router im Internet, um diesen ohne das Wissen des Besitzers zu überwachen oder fernzusteuern. Hardware-Backdoors weisen eine besondere Gefährdung auf, da sie auf niedrigster technischer Ebene implementiert sind und durch vorhandene Sicherheitssysteme und Schutzmaßnahmen typischerweise nicht detektiert werden können. Im Rahmen der Snowden-Affäre wurden Details zu verschiedenen Hardware-Backdoors bekannt, die für zahlreiche Systeme aller wichtigen Hersteller produziert wurden, beispielsweise GODSURGE, DEITYBOUNCE, IRONCHEF oder FLUXBABBITT [84]. Diese wurden auch in Deutschland im Rahmen der sogenannten Interdiction-Operations, bei denen Hardware auf dem Transportweg abgefangen und die Backdoors eingebaut wurden, in Systeme installiert. Große Telekommunikationsprovider wie die Deutsche Telekom haben hierauf reagiert und bessere Schutzmaßnahmen für die Übersendung von Hardware mit den entsprechenden Herstellen umgesetzt. Die letztendlich derzeit durch Hardware-Backdoors vorliegende Gefährdung lässt sich jedoch nur unzureichend abschätzen.

Während Angriffe auf BGP bereits auf Basis des erforderlichen Wissens durchgeführt werden können, benötigt die Positionierung und Ausnutzung von Hardware-Backdoors erhebliche Ressourcen und Fähigkeiten, sodass entsprechende Angriffe derzeit nur durch wenige Akteure durchgeführt werden können. Eine Detektion ist jedoch erheblich komplexer und Teil derzeit laufender Forschungen.

3 Schlussbemerkungen

Das vorliegende Sachverständigengutachten beantwortet die Fragen des 1. Untersuchungsausschusses der 18. Wahlperiode aus Sicht der technischen Gegebenheiten und Möglichkeiten. Es stellt die technischen Rahmenbedingungen der paketvermittelten Datenübertragung bezogen auf autonome Systeme und Internet Exchange Points dar.

Dabei wird zunächst am Beispiel des DE-CIX in Frankfurt auf den Aufbau und die Funktionalität von IXP eingegangen. Es wird gezeigt, dass eine Vielzahl an AS, deren Betreiber aus vielen unterschiedlichen Ländern stammen, an einem solchen IXP angeschlossen ist. Gleichzeitig wird dargestellt, dass ein steigender Anteil des Internet-Datenverkehrs über direkte Verbindungen zwischen AS, ohne Involvierung der IXP, ausgetauscht wird.

Darüber hinaus wird die Funktionalität des Routing-Protokolls BGP erläutert, welches de facto das einzige verwendete Protokoll zur Wegefindung und -wahl zwischen den AS ist. Durch die Definition und Umsetzung von Policies ist es den Betreibern der AS möglich das BGP-Routing zu beeinflussen.

Es werden Möglichkeiten aufgezeigt, die eine Klassifizierung des Datenverkehrs an einem IXP erlauben. Hierbei werden der Port- und der Flow-basierte Ansatz sowie die Deep Packet Inspection erläutert. Insbesondere werden technische Gegebenheiten betrachtet und die Problemstellung hinsichtlich verschleierten und verschlüsselten Kommunikationsbeziehungen aufgezeigt. Ferner wird die Frage beantwortet, welche Verkehrsdaten für die Zusammensetzung von inhaltlich vollständigen Kommunikationsvorgängen notwendig sind. Betont wird dabei, dass Flow-basierte Ansätze zwar Erkenntnisse hierzu liefern können, aber nicht für die Zusammensetzung inhaltlich vollständiger Kommunikationen geeignet sind. Dafür ist die die Aufzeichnung und Auswertung des gesamten Datenverkehrs zwischen Kommunikationspartner erforderlich.

Eine Identifikation der Ursprungs- und Zielorte der Kommunikationsvorgänge kann in der Regel nur grob granular und bei nicht eingesetzten Verschleierungsmaßnahmen erfolgen. Auf Grundlage von Verfahren wie IP-Geolokalisation können hinreichend zuverlässige Entscheidungen getroffen werden, ob sich die Quell- oder Ziel-IP-Adresse eines Datenpakets im Ausland befindet. Wissenschaftlich anerkannte und praktisch erprobte Verfahren zur IP-Geolokalisation werden hinsichtlich des Ansatzes, der Funktionsweise und der Genauigkeit erläutert. Es wird ebenso festgehalten, dass aus geografischen Informationen nicht grundsätzlich weitere Eigenschaften der Kommunikationspartner abgeleitet werden können. Ferner wird auf Aspekte und Indizien eingegangen, welche Rückschlüsse darauf zulassen, ob ein Kommunikationspartner Deutscher ist. Zusätzlich werden die Problemstellungen des Einsatzes von Verschlüsselungstechnologien aufgezeigt. Darüber hinaus wird festgehalten, dass Rückschlüsse hinsichtlich der Sprache, des Namens und der aufgerufenen Webseiten im Internet immer nur Indizien liefern können und die Genauigkeit dieser Datenauswertung weiterer Untersuchungen bedarf.

Im letzten Teil des Gutachtens werden Angriffsmöglichkeiten dargestellt, welche den Regelbetrieb der AS und der IXP im Internet beeinflussen können und anhand aktueller Beispiele veranschaulicht. Vor allen Dingen werden Angriffsmöglichkeiten bezüglich BGP und DNS sowie mögliche Schutzmaßnahmen erläutert. Des Weiteren wird auf Schwachstellen in der Implementierung, dem Design und der Konfiguration der Routersoftware bzw. der Protokolle sowie auf Hardware-Backdoors eingegangen. Dabei wird herausgestellt, dass hierfür explizites Wissen bzw. erhebliche Ressourcen notwendig sind.

Literatur

[1] BITAG, “Interconnection and Traffic Exchange on the Internet,” Broadband Internet Technical Advisory Group, Tech. Rep., Nov. 2014.

[2] CAIDA, “Center for Applied Internet Data Analysis (CAIDA),” http://www.caida.org/home/ (16.09.2016).

[3] CAIDA, “CAIDA AS Rank,” as-rank.caida.org (15.09.2016).

[4] IANA, “Internet Assigned Numbers Authority (IANA),” www.iana.org (15.09.2016).

[5] RIPE, “RIPE NCC,” www.stat.ripe.net (15.09.2016).

[6] DFN, “Deutsches Forschungsnetz - DFN,” www.dfn.de (15.09.2016).

[7] GÉANT, “GÉANT,” www.geant.org (15.09.2016).

[8] A. Callado, C. Kamienski, G. Szabo, B. P. Gero, J. Kelner, S. Fernandes, and D. Sadok, “A Survey on Internet Traffic Identification,” *IEEE Communications Surveys & Tutorials*, vol. 11, no. 3, pp. 37–52, 2009.

[9] A. Dainotti, A. Pescape, and K. Claffy, “Issues and Future Directions in Traffic Classification,” *IEEE Network*, pp. 35–40, 2012.

[10] A. S. Tanenbaum and D. J. Wetherall, *Computer Networks*, 5th ed. Pearson Education Limited, 2011.

[11] P. Richter, G. Smaragdakis, A. Feldmann, N. Chatzis, J. Boettger, and W. Willinger, “Peering at peerings: On the role of ixp route servers,” *Proceedings of the 2014 Conference on Internet Measurement Conference*, pp. 31–44, 2014.

[12] B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Walter., “Anatomy of a Large European IXP,” *Proceedings of the ACM SIGCOMM*, 2012.

[13] DE-CIX, “DE-CIX German Internet Exchange,” www.de-cix.net/about/quick-facts/ (15.09.2016).

[14] J. Hawkinson and T. Bates, “Guidelines for creation, selection, and registration of an Autonomous System (AS),” RFC 1930 (Best Current Practice), Internet Engineering Task Force, Mar. 1996, updated by RFCs 6996, 7300. [Online]. Available: http://www.ietf.org/rfc/rfc1930.txt

[15] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, “SIP: Session Initiation Protocol,” RFC 3261 (Proposed Standard), Internet Engineering Task Force, Jun. 2002, updated by RFCs 3265, 3853, 4320, 4916, 5393, 5621, 5626, 5630, 5922, 5954, 6026, 6141, 6665, 6878, 7462, 7463. [Online]. Available: http://www.ietf.org/rfc/rfc3261.txt

[16] J. Reynolds and J. Postel, “Assigned Numbers,” RFC 1700 (Historic), Internet Engineering Task Force, Oct. 1994, obsoleted by RFC 3232. [Online]. Available: http://www.ietf.org/rfc/rfc1700.txt

[17] J. Reynolds, "Assigned Numbers: RFC 1700 is Replaced by an On-line Database," RFC 3232 (Informational), Internet Engineering Task Force, Jan. 2002. [Online]. Available: http://www.ietf.org/rfc/rfc3232.txt

[18] IANA, "Service Name and Transport Protocol Port Number Registry," www.iana.org/assignments/service-names-port-numbers/ (15.09.2016).

[19] S. Valenti, D. Rossi, A. Dainotti, A. Pescapè, A. Finamore, and M. Mellia, "Reviewing traffic classification," *Data Traffic Monitoring and Analysis*, pp. 123–147, 2013.

[20] P. Richter, N. Chatzis, G. Smaragdakis, A. Feldmann, and W. Willinger, "Distilling the Internet's Application Mix from Packet-Sampled Traffic," *International Conference on Passive and Active Network Measurement*, pp. 179–192, 2015.

[21] M. Golling, R. Hofstede, and R. Koch, "Towards Multi-layered Intrusion Detection in High-Speed Backbone Networks," *Proceedings of the 6th International Conference on Cyber Conflict (CyCon)*, pp. 1–17, 2014.

[22] B. Claise, B. Trammell, and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information," RFC 7011 (INTERNET STANDARD), Internet Engineering Task Force, Sep. 2013. [Online]. Available: http://www.ietf.org/rfc/rfc7011.txt

[23] R. Hofstede, P. Celenda, B. Trammell, I. Drago, R. Sadre, A. Sperotto, and A. Pras, "Flow Monitoring Explained: From Packet Capture to Data Analysis with NetFlow and IPFIX," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2037 – 2064, 2014.

[24] A. McGregor, M. Hall, P. Lorier, and J. Brunskill, "Flow Clustering using Machine Learning Techniques," *International Workshop on Passive and Active Network Measurement*, pp. 205–214, 2004.

[25] A. W. Moore and D. Zuev, "Internet Traffic Classification using Bayesian Analysis Techniques," *ACM SIGMETRICS Performance Evaluation Review*, vol. 33, no. 1, pp. 50–60, 2005.

[26] M. Roughan, S. Sen, O. Spatscheck, and N. Duffield, "Class-of-service mapping for QoS: A statistical Signature-based Approach to IP Traffic Classification," *Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement*, pp. 135–148, 2004.

[27] P. Phaal and M. Lavine, "sFlow Version 5," http://www.sflow.org/sflow_version_5.txt (16.09.2016), 2004.

[28] A. Callado, C. Kamienski, G. Szabó, B. P. Gero, J. Kelner, S. Fernandes, and D. Sadok, "A Survey on Internet Traffic Identification," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 3, pp. 37–52, 2009.

[29] S. Zander, G. Armitage, and P. Branch, "A Survey of Covert Channels and Countermeasures in Computer Network Protocols," *IEEE Communications Surveys & Tutorials*, vol. 9, no. 3, pp. 44–57, 2007.

[30] B. Claise, "Cisco Systems NetFlow Services Export Version 9," RFC 3954 (Informational), Internet Engineering Task Force, Oct. 2004. [Online]. Available: http://www.ietf.org/rfc/rfc3954.txt

[31] J. Quittek, T. Zseby, B. Claise, and S. Zander, "Requirements for IP Flow Information Export (IPFIX)," RFC 3917 (Informational), Internet Engineering Task Force, Oct. 2004. [Online]. Available: http://www.ietf.org/rfc/rfc3917.txt

[32] P. Phaal, S. Panchen, and N. McKee, "InMon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks," RFC 3176 (Informational), Internet Engineering Task Force, Sep. 2001. [Online]. Available: http://www.ietf.org/rfc/rfc3176.txt

[33] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)," RFC 4271 (Draft Standard), Internet Engineering Task Force, Jan. 2006, updated by RFCs 6286, 6608, 6793, 7606, 7607, 7705. [Online]. Available: http://www.ietf.org/rfc/rfc4271.txt

[34] K. Obermann and M. Horneffer, *Datennetztechnologien für Next Generation Networks: Ethernet, IP, MPLS und andere*, ser. SpringerLink: Bücher. Springer Fachmedien Wiesbaden, 2013.

[35] R. Chandra, P. Traina, and T. Li, "BGP Communities Attribute," RFC 1997 (Proposed Standard), Internet Engineering Task Force, Aug. 1996, updated by RFC 7606. [Online]. Available: http://www.ietf.org/rfc/rfc1997.txt

[36] M. Caesar and J. Rexford, "BGP Routing Policies in ISP Networks," *IEEE Network*, vol. 19, no. 6, pp. 5–11, 2005.

[37] DE-CIX, "DE-CIX Frankfurt - BIRD Looking Glass," https://lg.de-cix.net (23.09.2016).

[38] CERN, "CERN Looking Glass," http://lg.cern.ch/ (23.09.2016).

[39] O. Berthold, H. Federrath, and S. Köpsell, "Web MIXes: A system for anonymous and unobservable Internet access," *Designing Privacy Enhancing Technologies, Proceedings of the Workshop on Design Issues in Anonymity and Unobservability*, pp. 115–129, 2009.

[40] D. Chaum, F. Javani, A. Sherman, D. Das, A. Kate, A. Krasnova, and J. de Ruiter, "cMix: Anonymization by High-Performance Scalable Mixing," *Proceedings of ACM CCS 2016*, 2016.

[41] V. N. Padmanabhan and L. Subramanian, "An Investigation of Geographic Mapping Techniques for Internet Hosts," *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 4, pp. 173–185, Aug. 2001.

[42] P. T. Endo and D. F. H. Sadok, "Whois Based Geolocation: A Strategy to Geolocate Internet Hosts," *Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications*, pp. 408–413, 2010.

[43] R. Koch, M. Golling, and G. Dreo Rodosek, "Geolocation and Verification of IP-Addresses with Specific Focus on IPv6," *5th International Symposium on Cyberspace Safety and Security (CSS 2013)*, pp. 1–20, 2013.

[44] A. Dahnert, "HawkEyes: An advanced IP Geolocation approach: IP Geolocation using semantic and measurement based techniques," *Second Worldwide Cybersecurity Summit (WCS)*, June 2011.

[45] B. Eriksson, P. Barford, B. Maggs, and R. Nowak, "Posit: An Adaptive Framework for Lightweight IP Geolocation," BU/CS, Tech. Rep., July 2011.

[46] D. Li, J. Cheny, C. Guo, Y. Liu, J. Zhangy, Z. Zhang, and Y. Zhang, "IP-Geolocation Mapping for Involving Moderately-Connected Internet Regions," Microsoft Research, Tech. Rep., 2009.

[47] A. Ziviani, S. Fdida, J. F. de Rezende, and O. C. M. B. Duarte, "Improving the Accuracy of Measurement-based Geographic Location of Internet Hosts," *Computer Networks and ISDN Systems*, vol. 47, no. 4, pp. 503–523, Mar. 2005.

[48] G. Ballintijn, M. van Steen, and A. S. Tanenbaum, "Characterizing Internet Performance to Support Wide-area Application Development," *Operating Systems Review*, vol. 34, pp. 41–47, 2000.

[49] A. Ziviani, S. Fdida, J. de Rezende, and O. Duarte, "Similarity Models for Internet host location," *The 11th IEEE International Conference on Networks (ICON)*, pp. 81–86, Sept 2003.

[50] B. Eriksson and M. Crovella, "Understanding Geolocation Accuracy using Network Geometry," *The 32nd Annual IEEE International Conference on Computer Communications (INFOCOM)*, 2013.

[51] E. Katz-Bassett, J. P. John, A. Krishnamurthy, D. Wetherall, T. Anderson, and Y. Chawathe, "Towards IP Geolocation Using Delay and Topology Measurements," *Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement*, pp. 71–84, 2006.

[52] B. Gueye, A. Ziviani, M. Crovella, and S. Fdida, "Constraint-based Geolocation of Internet Hosts," *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, pp. 1219–1232, Dec. 2006.

[53] B. Gueye, S. Uhlig, A. Ziviani, and S. Fdida, "Leveraging buffering delay estimation for geolocation of Internet hosts," *Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems*, pp. 319–330, 2006.

[54] S. Laki, P. Matray, P. Haga, T. Sebok, I. Csabai, and G. Vattay, "Spotter: A model based active Geolocation Service," *IEEE INFOCOM*, pp. 3173–3181, April 2011.

[55] B. Wong, I. Stoyanov, and E. G. Sirer, "Octant: A Comprehensive Framework for the Geolocalization of Internet Hosts," *Proceedings of the 4th USENIX Conference on Networked Systems Design and Implementation*, pp. 23–23, 2007.

[56] Y. Wang, D. Burgener, M. Flores, A. Kuzmanovic, and C. Huang, "Towards Street-level Client-independent IP Geolocation," *Proceedings of the USENIX Conference on Networked Systems Design and Implementation*, pp. 365–379, 2011.

[57] B. Huffaker, M. Fomenkov, and kc claffy, "Geocompare: a comparison of public and commercial geolocation databases," Network Mapping and Measurement Conference (NMMC), Tech. Rep., May 2011.

[58] L. Daigle, "WHOIS Protocol Specification," RFC 3912 (Draft Standard), Internet Engineering Task Force, Sep. 2004. [Online]. Available: http://www.ietf.org/rfc/rfc3912.txt

[59] Inja Youn, Brian L. Mark and Dana Richards, "Statistical Geolocation of Internet Hosts," International Conference on Computer Communications and Networks, Tech. Rep., 2009.

[60] L. Subramanian, V. N. Padmanabhan and R. H. Katz, “Geographic Properties of Internet Routing,” *USENIX Annual Technical Conference*, June 2002.

[61] MaxMind, Inc., “MaxMind Geolocation Service,” www.maxmind.com (15.09.2016).

[62] I. Poese, M. A. Kaafar, B. Donnet, B. Gueye, and S. Uhlig, “IP Geolocation Databases: Unreliable?” Deutsche Telekom Lab./TU Berlin, Germany, Tech. Rep., March 2011.

[63] Sebastian Zander, “How Accurate is IP Geolocation Based on IP Allocation Data?” Centre for Advanced Internet Architectures (CAIA), Tech. Rep., May 2012.

[64] Y. Yuval Shavitt and N. Zilberman, “A Study of Geolocation Databases,” School of Electrical Engineering, Tech. Rep., July 2010.

[65] R. Koch, M. Golling, L. Stiemert, and G. Dreo Rodosek, “Using Geolocation for the Strategic Preincident Preparation of an IT Forensics Analysis,” *IEEE Systems Journal*, 2015.

[66] C. Davis, P. Vixie, T. Goodwin, and I. Dickinson, “A Means for Expressing Location Information in the Domain Name System,” RFC 1876 (Experimental), Internet Engineering Task Force, Jan. 1996. [Online]. Available: http://www.ietf.org/rfc/rfc1876.txt

[67] Kaushik Srinivasan and Krishna Venkatasubramanian, “Geography of the Web - Design and Analysis of Algorithm,” CSE 450/598, Arizona State University, Tech. Rep., 2003.

[68] Ammar J. Salih, “Enhancing Location Based IP Services,” IETF Internet-Draft, May 2013.

[69] CAIDA, “Internet Protocol Address (IP) Geolocation Bibliography,” http://www.caida.org/projects/cybersecurity/geolocation/bib/ (07.09.2016).

[70] P. Hillmann, L. Stiemert, G. Dreo Rodosek, and O. Rose, “Dragoon: Advanced Modelling of IP Geolocation by use of Latency Measurements,” *Proceedings of the 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, 2015.

[71] R. Percacci and Vespignani, “Scale-free behavior of the internet global performance,” *The European Physical Journal*, 2003.

[72] K. Schneider and S. Venters, “PPP for Data Compression in Data Circuit-Terminating Equipment (DCE),” RFC 1976 (Informational), Internet Engineering Task Force, Aug. 1996. [Online]. Available: http://www.ietf.org/rfc/rfc1976.txt

[73] J. A. Muir and P. C. van Oorschot, “Internet Geolocation: Evasion and Counterevasion,” *ACM Computing Surveys*, vol. 42, no. 1, pp. 4:1–4:23, Dec. 2009.

[74] A. M. Abdou, A. Matrawy, and P. C. van Oorschot, “On the Evasion of Delay-Based IP Geolocation.”

[75] A. Abdou, A. Matrawy, and P. van Oorschot, “CPV: Delay-based Location Verification for the Internet,” *IEEE Transactions on Dependable and Secure Computing*, 2015.

[76] Monika Ermert, “Routing-Kleinkrieg Ursache für YouTube-Ausfall,” https://heise.de/-205345 (15.09.2016), 2008.

[77] S. Murphy, “BGP Security Vulnerabilities Analysis,” RFC 4272 (Informational), Internet Engineering Task Force, Jan. 2006. [Online]. Available: http://www.ietf.org/rfc/rfc4272.txt

[78] K. R. Butler, T. R. Farley, P. McDaniel, and J. Rexford, “A Survey of BGP Security Issues and Solutions,” *Proceedings of the IEEE*, vol. 98, no. 1, pp. 100–122, 2010.

[79] Steve Friedl, “An Illustrated Guide to the Kaminsky DNS Vulnerability,” http://unixwiz.net/techtips/iguide-kaminsky-dns-vuln.html (15.09.2016).

[80] ICANN Security and Stability , “SSAC Advisory SAC008 DNS Distributed Denial of Service (DDoS) Attacks ,” https://www.icann.org/en/system/files/files/dns-ddos-advisory-31mar06-en.pdf (15.09.2016).

[81] John, “DNS - Based Phishing Attack in Public Hotspots,” https://www.exploit-db.com/docs/20875.pdf (15.09.2016).

[82] ICANN Security and Stability, “SAC 025 SSAC Advisory on Fast Flux Hosting and DNS,” https://www.icann.org/en/system/files/files/sac-025-en.pdf (15.09.2016).

[83] Constanze Kurz, “‚Equation Group’: Hacker als Hacking-Opfer?” https://netzpolitik.org/2016/equation-group-hacker-als-hacking-opfer/ (15.09.2016).

[84] Pierluigi Paganini, “A Close Look at the NSA Monitor Catalog – Server Hacking,” http://resources.infosecinstitute.com/close-look-nsa-monitor-catalog-server-hacking/ (15.09.2016).