# Neutralität und Transparenz von Netzwerken und Anwendungen

Dissertation
zur
Erlangung des akademischen Grades
Doktor-Ingenieur (Dr.-Ing.)
der Fakultät für Informatik und Elektrotechnik
der Universität Rostock

vorgelegt von
Andreas Dähn.

Rostock, den 25. Mai 2014

urn:nbn:de:gbv:28-diss2015-0047-3

## Gutachter

1. Gutachter:
   Prof. Dr. Clemens H. Cap
   Institut für Informatik, Universität Rostock

2. Gutachter:
   Prof. Dr. Hubertus Gersdorf
   Juristische Fakultät, Universität Rostock

3. Gutachter:
   Prof. Dr. Gerhard Schneider
   Institut für Informatik, Universität Freiburg

## Datum der Einreichung

27. Mai 2015

## Datum der Verteidigung

22. Januar 2015

Das ist ein hässliches Gebrechen,
wenn Menschen wie die Bücher sprechen.
Doch reich und fruchtbar sind für jeden
die Bücher, die wie Menschen reden!

*—Oskar Blumenthal (1852-1917)*

# Dank

Wem soll man danken? Dies impliziert immer auch die Frage, wem man nicht danken will. Eine schwierige, weil eine gehässige Frage. Dank nur demjenigen, der durch die Dissertation begleitet hat? Das wäre unfair gegenüber früheren Weggefährten, die vielleicht weit mehr Einfluss hatten, als man es sich eingestehen mag. Was ist mit denjenigen, die eigentlich gar nichts mit dieser Arbeit zu tun hatten – und dadurch vielleicht noch mehr dazu beitrugen als Andere? Alle aufzuzählen scheint gerecht. Doch ist es auch gefährlich, denn wer jetzt versehentlich vergessen, der fühlt sich nicht gedankt – zurecht. Und so scheint die Lösung im Verzicht auf das Konkrete – dem Abstrakten. Mit Ausnahmen, denn sie bestätigen auch diese Regel.

Ich danke Familie und Freunden, außerdem meinem Doktorvater Clemens H. Cap für offenes Ohr und anregende Kritik.

Ich danke (ehemaligen) Mitarbeiterinnen und Mitarbeitern des Lehrstuhls für Informations- und Kommunikationsdienste für Unterstützung und Kritik. Vielleicht sogar mehr für die Kritik, denn sie scheint – obgleich schwerer verdaulich – im Nachhinein weit wichtiger denn Lob. Ebenso danke ich Prof. Gersdorf, seinen Mitarbeitern und allen Anderen „im Haus“, die ich kennen und schätzen lernte.

Ganz besonders danke ich der Landesgraduiertenförderung Mecklenburg-Vorpommern, deren Stipendiat ich zwei Jahre sein durfte.

Ich danke den Kollegen vom Landesfunkhaus in Schwerin und dem Ostseestudio in Rostock, die mich regelmäßig einer Erdung zuführten, indem sie mich aus dem Elfenbeinturm der Wissenschaft holten und zu interessanten Reisen durchs Land mitnahmen.

Ich bedanke mich bei meinen Korrekturlesern Uwe, Petra, Jonas, Martin und Anke.

... und bei allen die ich nun doch vergessen habe und die mich zu dem gemacht haben, der ich nun bin. Danke.

## Zusammenfassung

Neutralität von Netzwerken wird seit einigen Jahren diskutiert: Hierbei geht es klassischerweise um Fragen der differenzierten Behandlung von Datenpaketen beim Transport durch Netzwerke. Beantwortet wird die Frage nach der Netzneutralität üblicherweise binär: Ein Netz wird als neutral oder als nicht neutral klassifiziert; Abstufungen sind nicht vorgesehen. In einer Welt, in der bestimmte Netzdienste ohne Priorisierung schnell an Grenzen stoßen[1] erscheint eine solche Dogmatik zunehmend problematisch oder lässt sich nur durch Ergänzung der Definition um immer weitergehende Nebenbedingungen aufrecht erhalten. In dieser Arbeit wird daher stattdessen die Transparenz von Netzen als Metaebene etabliert: Nicht das Wissen, **ob** eine Verbindung neutral ist, sondern **das Wissen um mögliche Einflüsse** steht im Mittelpunkt. Zentrales Anliegen ist, die Bedeutung dieses Paradigmenwechsels herauszuarbeiten und auf eine solide technische und juristische Grundlage zu stellen.

Die Arbeit liefert Ausgangspunkte für eine spätere Entwicklung von Werkzeugen, mit denen Nutzer (Anwender) in die Lage versetzt werden, möglichst viel darüber zu wissen, welche Einflüsse auf von ihnen verwendete Netzwerkverbindungen (typischerweise der Internetanschluss) und Anwendungen einwirken. In diesem Sinne ist „Transparenz von Netzen und Anwendungen“ in dieser Arbeit zu verstehen: Als dem Nutzer sichtbare Gestaltung, die bislang oft durch Abstraktionsmechanismen unsichtbar gemacht wird. Hierzu werden in dieser Arbeit zwei wesentliche Fragenkomplexe untersucht:

**Erstens** Wie sind bestehende Netzwerke technisch zu ergänzen, um eine überprüfbar transparente Gestaltung von Netzwerkpolicies[2] sicherzustellen? Kann eine solche Gestaltung Netzwerkprovidern vorgeschrieben werden? Wie sind die juristischen Voraussetzungen hierfür in der Bundesrepublik Deutschland und in der Europäischen Union?

**Zweitens** Was bedeutet Transparenz von Anwendungen? Welche Anwendungen sind von mangelnder Transparenz besonders betroffen? Was sind die Auswirkungen mangelnder Transparenz verschiedener Anwendungen auf den Nutzer und für sein Verhalten?

Die Ergebnisse dieser Arbeit sind

**Erstens** die Beschreibung der Schwierigkeit einer konsistenten Definition der Netzneutralität und des damit einhergehenden fundamentalen Nachweisproblems (in diesem Zusammenhang werden bestehende Testverfahren systematisiert und eingeordnet),

**zweitens** die Beschreibung einer standardisierten und überprüfbaren Offenlegung von Netzwerkpolicies als eine Lösung der beschriebenen Probleme einer dogmatischen Klassifikation von Netzen in neutrale und nichtneutral,

[1] z. B. Fernsehen per Internet (IP-TV) oder Voice over IP (VoIP)

[2] auch „Traffic-Engineering-Rules“; Regelsätze, nach denen Netzwerkbetreiber Datenströme lenken, beschleunigen oder verlangsamen

**drittens** die Beschreibung von Transparenz und Intransparenz auf Anwendungsebene als nicht auf einzelne (spezielle) Anwendungen begrenzte Phänomene.

Die beschriebene Transparenz wird zwar bereits teilweise vom Gesetzgeber gefordert; es gibt jedoch keine standardisierte oder prüfbare Spezifikationsform. Eine solche Spezifikationsform wird mit dieser Arbeit vorgelegt.

Auch auf Anwendungsebene gibt es bereits Betrachtungen zur Transparenz – diese beziehen sich jedoch bislang meist nur auf soziale Netzwerke und Suchmaschinen. Dabei ist Intransparenz von Anwendungen insbesondere bei (populären) Webanwendungen eher die Regel als die Ausnahme.

Triebfeder beider Fragen und zugleich Verankerung dieser Arbeit in der realen Welt ist die Annahme, dass die Verfügbarkeit genauerer Informationen über das Verhalten von Netzen und Anwendungen Nutzern bei Entscheidungen wie z. B. der Auswahl eines Internetzugangsanbieters hilft. Diese Annahme ist nicht neu – sie war einer der dominierenden Gedanken bei der Transformation des Telekommunikationsmarktes im Zuge der Auflösung staatlicher Monopole in den 90er Jahren.

Diese Arbeit wird nach einer kurzen Einführung in die Netzwerktechnik den Themenkomplex der Netzneutralität sowie den bestehende Stand der Technik zum Nachweis von Neutralitätsverletzungen aufarbeiten und im Anschluss kritisch betrachten und systematisieren. Dies ist die Grundlage für die im darauffolgenden Kapitel entwickelte Idee der Transparenz von Netzwerken. Der Erörterung der Effekte auf Netzwerkebene schließt sich in logischer Folge die Betrachtung der Auswirkungen von (mangelnder) Transparenz auf Anwendungsebene an.

Diese Arbeit ist konzeptionell gestaltet; sie folgt nicht dem geradezu klassischen und häufig anzutreffenden Schema

> Gegeben Problem $P$, zu dem es die Lösungsansätze $a, b, c$ gibt. Es wurde der Ansatz $a'$ entwickelt, der aber nur unter den Randbedingungen $r_1, .., r_n$ funktionieren wird. Durch Simulation wurde gezeigt, dass $a'$ gegenüber $a$ um $p\%$ bessere Werte der Bewertungsmetriken $x, y, z$ zeigt,

denn diese Arbeit sucht nach dem Rahmen für eine Interpretation von Zahlenwerten und den Grenzen bestehender Definitionen – Betrachtungen, die auf Basis statistischer Auswertung empirischer Experimente kaum möglich sind. Stattdessen steht die Beobachtung und Beschreibung von Phänomenen und deren Einfluss auf den Nutzer im Fokus dieser Arbeit.

## Abstract

Over the last couple of years, network neutrality has become an interesting research topic. Typically, questions regarding different handling of data packets during their transport are discussed. The issue of network neutrality usually leads to binary answers: Either a network is neutral – or not: There are no nuances of network neutrality. In a world in which network applications which rely on a certain level of network quality[3] are available to the public, those definitions hardly comply with reality. Either definitions need extra conditions (e.g. why prioritisation of IPTV related traffic is not a neutrality breach) or it has to be accepted that such services and applications can by design not be used with a truly neutral net. Instead of adding up more conditions to fit a neutrality definition with today's networks and applications, this thesis presents the idea of network transparency as a new concept on a meta-level to network neutrality. Not the knowledge, **whether** a network is neutral but the knowledge **which measures are applied to traffic** should be considered. Central element of this thesis is to point out the importance and the ubiquitous aspects of this paradigm shift as well as providing its technical and legal foundation.

This thesis offers starting points for a subsequential developement of tools which can be used by end users to explore what influences their everyday Internet experience in terms of traffic engineering. This is how "transparency of networks and applications" shall be conceived in this thesis: As technical means to make traffic or application engineering visible to users, allowing them to understand influences which are usually hidden by mechanisms such as abstraction. To reach this goal, two research areas have been explored:

**First** Which modifications have to be applied to existing networks in order to enable users to receive and test a given specification of network behavior? How can network providers be obliged to provide such information and testing infrastructure? Is there a legal foundation for such a measure in the federal republic of Germany – or in the European Union?

**Second** What is application transparency? Which applications are more and which are less transparent – and why? How influences a lack of application transparency the users' behavior?

The findings of this thesis are

**First** A precise description of the issues related to a classic binary definition of network neutrality as well as the fundamental problems of neutrality violation detection. In this context, existing approaches are reviewed and classified;

**Second** the description of how to construct a standardised and testable declaration of traffic engineering rulesets as a solution addressing the problems discussed in the context of a binary view on network neutrality issues;


[3] e.g. television over Internet (IPTV) or voice over IP (VoIP)

**Third** the description of application transparency and intransparency as a phenomenon not limited to a handful (specialized) applications.

Described transparency is partially a legal obligation to network providers, but there is no need for a standardised testable declaration of network behavior yet – as suggested within this thesis.

On application layer, there are already discussions on transparency, which unfortunately only relate to social networks and search engines. However, the phenomenon of application intransparency is not restricted to these classes of applications – in fact, most (popular) web applications are rather intransparent.

This thesis is founded on the belief that the availability of detailed (understandable) information will always help customers when making decisions, e.g. when choosing a new Internet service provider. This assumption is not new or unique – it was one of the leading thoughts when the state dropped its monopole on telecommunication services.

The thesis starts with a short introduction to network techniques, followed by a state of the art chapter exploring both definitions of network neutrality and approaches to proof neutrality violations. The following chapter reviews the findings of the previous chapter, systematically classifying the testing approaches. Based on this analysis of existing approaches, chapter 5 develops the idea of network transparency as a concept. Finally (and following the ISO-/OSI network model stack from bottom up) the next chapter explores the effects of application (in-)transparency.

This thesis is conceptual; it does not share the classic scheme

> Given problem $P$, which can be addressed using approaches $a, b, c$. We developed approach $a'$, which is applicable under the conditions $r_1, ..., r_n$. In simulation, $a'$ shows an increase of $p\%$ regarding the metrics $x, y, z$,

since it searches for a new interpretation of measurement values – a task which can hardly be accomplished by interpretation of measurement values. Instead, the observation and description of phenomena relevant to users is focussed in this thesis.

# Inhaltsverzeichnis

# 1. Einführung

Wenn wir die Ziele wollen, wollen wir auch die Mittel.
—*Immanuel Kant*

Ausgangspunkt dieser Arbeit ist die Neutralität von Netzwerken, also ihre Eigenschaft, Datenpakete trotz unterschiedlicher Eigenschaften (wie Absender, Empfänger oder Inhalt) stets gleich zu behandeln. Es gibt jedoch eine ganze Reihe an Beispielen von Netzwerkverhalten, dass als Neutralitätsverstoß gedeutet wird. Hier werden als Einführung zwei alltägliche Vorkommnisse geschildert.

**Einführungsbeispiel 1** (entnommen aus [62])

*Ein Nutzer bucht einen mobilen Internetzugang. Dieser scheint auch gut zu funktionieren. Nach kurzer Zeit fällt jedoch auf, dass es einen bestimmten Dienst (Internettelefonie via „Skype“) gibt, der mit diesem Internetzugang unbrauchbar scheint. Ein Anruf beim Provider führt zur Aussage, dass dies nicht an ihm, sondern am Endgerät liegen müsse. Von Nutzerseite beginnen neugierige Tests, die die folgenden Phänomene aufzeigen:*

- *Eine Verbindung zu* `skype.com` *hat einen extrem geringen Durchsatz (weit unter dem Durchsatz zu anderen Servern),*
- *eine Abhängigkeit dieses geringen Durchsatzes von Endgerät (bei gleichbleibender UMTS-SIM-Karte), Ort oder Zeit lässt sich nicht feststellen,*
- *ein Test mit einem anderen Internetzugangsanbieter vom gleichen geographischen Ort ist ohne Problemfund: Ein „Skype“-Telefonat ist möglich.*

*Erst ein genaues Studium der Vertragsunterlagen bringt die Erkenntnis, dass einige Dienste (darunter auch Voice-over-IP[1]) durch den abgeschlossenen Vertrag nicht abgedeckt sind. Für diese Dienste wird vom Provider ein separat buchbarer Tarif angeboten.*

*Bereits kurz nach der Buchung dieses Tarifs ist das Problem nicht mehr nachweisbar und „Skype“ problemlos nutzbar.*

**Einführungsbeispiel 2** (aus dem Alltag an der Universität Rostock)

*Am Morgen des 29. November 2013 beschrieb mein Kollege Martin Garbe ein interessantes Phänomen: Er versuchte, durch das Universitätsnetz eine bestimmte Datei aus dem Internet herunterzuladen – die Übertragung kam jedoch nie über eine bestimmte Größe*

[1] VoIP

*hinaus. Seinen Browser hatte er schon durch Tests mit anderen Browsern als Ursache ausgeschlossen. Die Verbindung wurde zwar nicht unterbrochen, doch es konnten auch keine weiteren Datenpakete ausgetauscht werden. Von Internetanschlüssen außerhalb der Universität war ein Download problemlos möglich*[2]*. Außerdem schien das Phänomen primär Dateien im MSI-Format*[3] *zu betreffen. Von Neugierde gepackt verbrachten wir den Rest des Tages damit, das Phänomen genauer zu ergründen und im Detail zu dokumentieren [63]. Für eine möglichst kontrollierte Umgebung wurde eine der betroffenen Dateien auf einen Testserver hochgeladen, der entsprechend des in Abb. 1 skizzierten Netzes erreichbar war.*

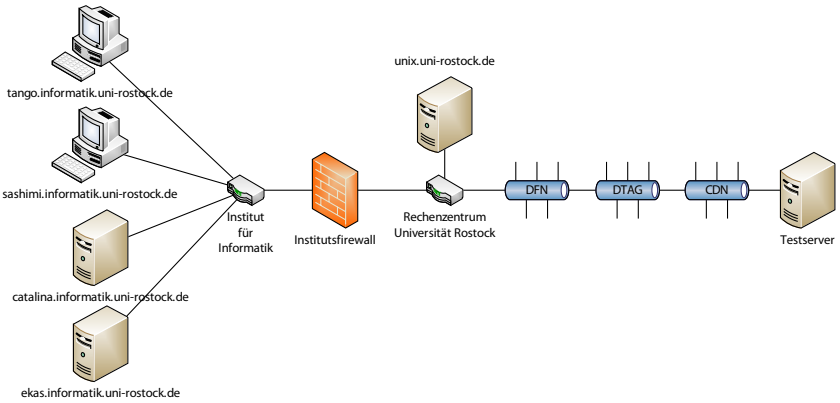


Abbildung 1.1.: Testumgebung: Die beobachteten Datenpakete bewegten sich durch das Instituts- und Rechenzentrumsnetz in das Netzwerk des „Deutschen Forschungsnetzes“ (DFN) um schließlich nach Passage des Netzes der „Deutschen Telekom“ (DTAG) durch das CDN-Netzwerk dem Server zugestellt zu werden.

*Bei* `tango` *und* `sashimi` *handelt es sich um Arbeitsplätze;* `ekas` *und* `catalina` *sind Mitarbeitern und Studenten des Instituts zugängliche unixoide Systeme. Die ersten Beobachtungen erfolgten von den Arbeitsplätzen aus. Weiteres Testen ergab schließlich, dass sich auch die anderen Systeme innerhalb des Institutsnetzes identisch verhielten. Somit richtete sich der initiale Verdacht gegen die Insitutsfirewall*[4]*. Um diese Hypothese zu widerlegen, loggten wir uns per SSH auf* `unix.uni-rostock.de` *(einem vom Universitätsrechenzentrum betriebenen Login-Server, der nicht von der Institutsfirewall beeinflusst wird) ein und versuchten von hier aus, die betroffenen Downloads durchzuführen. Wir mussten unsere Hypothese verwerfen, als der Download auch hier scheiterte. Durch Versuche mit Rechnern, die sich an anderen „Orten“ im Internet befanden, konnten die*

[2] und die Dateien konnten weiter durch eine verschlüsselte Verbindung zu den Arbeitsplätzen übertragen werden.
[3] Installationspakete für das Betriebssystem Microsoft Windows
[4] deren Regeln weder den Nutzern zugänglich noch – soweit bekannt – in jedem Detail der Motivation nach verständlich sind.

*letzten Routingstationen vor dem Testserver als Ursache ausgeschlossen werden. Da wir zunächst keine Möglichkeit hatten, weitere Tests mit Stationen durchzuführen, die zu einer weiteren Teilabdeckung des Routingpfads führen würden, begannen wir, den Auslöser des Stockens des Datenverkehrs zu suchen. Wir stießen auf Byte-Sequenzen, die sich nicht in das oder aus dem Netzwerk der Universität Rostock transferieren ließen. Dies betraf sowohl UDP- wie auch TCP-basierte Übertragungen. Diese Erkenntnis war wesentlich, denn sie erlaubte den Verzicht auf eine dedizierte Testgegenstelle, die ein bestimmtes Protokoll unterstützen muss. Stattdessen kann das Verhalten beliebiger Datenpakete betrachtet werden, also auch von Datenpaketen, die für einen Routen-Trace verwendet werden. Daher wurde diese Sequenz genutzt, um mittels* `hping3` *einzelne Routingstationen analog zu* `traceroute` *zu prüfen*[5]*: In einem Test wurde eine zufällige Bytesequenz als Payload verwendet, in einem zweiten die bekannte Problemsequenz. Es ergab sich eine Differenz: Mit der Problemsequenz als Payload brach der Routingpfad innerhalb der Infrastruktur des Rechenzentrums ab; mit der alternativen Sequenz konnte die angepeilte Zielmaschine problemlos erreicht werden. Mit diesen Experimenten war abgesichert, dass das Phänomen von einer bestimmten Kante des Routinggraphen ausging; wir konnten diese Kante in der Infrastruktur des Rechenzentrums identifizieren.*
*Der Leiter unserer Arbeitsgruppe kontaktierte schließlich das Rechenzentrum unter Verweis auf unsere Dokumentation und erhielt die Information, dass es dort eine normalerweise unsichtbare (gegenüber den Nutzern nicht dokumentierte) und auf einer Deep Packet Inspection basierende Intrusion-Prevention-Technik gäbe, bei deren Signaturupdates es zu Problemen gekommen wäre. Bereits kurze Zeit nach der Kontaktaufnahme war unsere Beobachtung nicht mehr reproduzierbar.*

Auch wenn die Beispiele für sich betrachtet den Eindruck erwecken könnten, dass es sich hierbei nur um singuläre Banalitäten handelte und es vielleicht sogar absurd erscheint, einen „Schluckauf“ im Netzwerk zum Aufhänger einer Dissertation zu stilisieren – dieser Eindruck täuscht. Bei genauerer Betrachtung werden die Tragweiten der Eingriffe und das mit ihnen verbundene (Missbrauchs-) Potential deutlich. Beispiel 1 illustriert die Verbreitung und kommerzielle Nutzung von Netzwerkpolicies, denn was vom Nutzer als „Zusatzpaket“ gebucht wurde, war die Änderung einer einzelnen Klassifikationsregel, die den Datenaustausch mit `skype.com` begrenzte. Beispiel 2 belegt die Verbreitung von Technik, die – trotz bester Absichten – verheerende Folgen haben kann. Es wurden Datenströme herausgefiltert, die eine obskure Bytefolge enthielten. Technisch ist hier kein Unterschied zu einem Datenstrom, der einen Satz in menschlicher Sprache, eine Parole, eine Idee verbreitet – all das ist für Netzwerkinfrastruktur nur Paketinhalt, der bei einer Analyse gefunden werden kann.

Beide Beispiele zeigen dabei einen fundamentalen Mangel auf: Den Mangel an vom Nutzer wirklich nachvollziehbarer Deklaration von Netzwerkpolicies. Im Beispiel 1 wäre

[5] Hierbei werden Pakete mit einem niedrigen Time-To-Live- (TTL-) Wert verschickt. Jeder Router dekrementiert diesen Wert. Erreicht er 0, wird das Paket verworfen und der Router schickt eine ICMP-Nachricht an den Absender, dass das Paket nicht zugestellt werden konnte. Für einen Routen-Trace werden Pakete mit einer TTL von 0,1,... versandt und die Antworten gesammelt. Hierbei können die Pakete einen Payload haben oder nicht; dies sollte das Verfahren als solches nicht beeinflussen.

die Suche nach der Ursache der geringen Performance von „Skype“ schnell und eindeutig beantwortet, würde es eine Liste von Netzwerkpolicies geben, die einen entsprechenden Eintrag enthielte oder einfach eine entsprechende Information in lesbarer Form. Im Beispiel 2 wäre das Filtern von Datenströmen in einer solchen Deklaration zu finden gewesen. Diese Konstellation verdeutlicht, dass die Idee einer Spezifikation mehr als eine akademische Spitzfindigkeit ist.

Doch was, wenn die Filter nicht deklariert gewesen wären? Im Beispiel 2 war es lediglich ein technisches Artefakt, dass einen Test zuließ, mit dessen Hilfe die Ursache des Phänomens lokalisiert werden konnte. Nur weil der Filter beliebige Pakete und mit diesen korrelierte Datenströme erfasste, konnte der blockierte Inhalt als Payload eines Traces verwendet werden. Im Fall von Beispiel 1 wären entsprechende Tests schon deshalb um Größenordnungen aufwändiger, weil „Skype“ ein nicht offengelegtes und verschlüsseltes Protokoll verwendet. Beispiel 2 macht deutlich, warum es mit einer reinen Spezifikation von Netzwerkpolicies nicht getan ist und weshalb diese überprüfbar sein müssen: Weil eine Spezifikation willentlich oder unwillentlich falsch sein kann und dies dem Nutzer nicht unmittelbar auffallen muss.

Beachtung verdient auch die Reaktion des Internetzugangsproviders im Beispiel 1, der die Ursache des Phänomens zunächst der vom Nutzer betreuten Technik zuschrieb. Eine solche Argumentation kann nur deshalb Erfolg haben, weil Nutzer bei vielen der von ihnen verwendeten Dienste kein Wissen über deren innere Vorgänge besitzen. Hier setzt die Frage der Anwendungstransparenz an: Wie können Anwendungen geschaffen werden, die es dem Nutzer erlauben, für ihn sichtbare Effekte klar einzelnen Komponenten zuzuordnen?

Die Beispiele weisen noch eine Gemeinsamkeit auf: Beide illustrieren, wie schnell Fragen der Anwendungs- wie auch der Netzneutralität und -transparenz den Nutzer in unterschiedlichen Formen erreichen. Aber auch, wie schnell sie für ihn auch wieder unsichtbar werden können.

## 1.1. Fragestellung und Abgrenzung

In diesem Abschnitt wird zunächst die Vielfalt der Forschungsfragen ausgeführt, die sich aus dem Thema Netzneutralität entwickeln; schließlich (ab S. 24) werden die in dieser Arbeit beantworteten Fragen konkretisiert und abgegrenzt.

### 1.1.1. Das Themenfeld Netzneutralität

In den letzten Jahren hat sich ein Spannungsfeld aufgebaut, in dem sich die Interessen von Netzwerkbetreibern, Inhalteanbietern und Endkunden gegenüberstehen:

*Der Netzbetreiber im ersten Einführungsbeispiel wird den Aufschlag für die Nutzung von „Skype“ mit der intensiven Netznutzung durch dieses Programm begründen; der Nutzer wird begründen, dass er doch bereits für den Netzzugang zahlen würde und dass er*

*mit dem Aufschlag praktisch doppelt für einen Netzzugang zahlen soll. Er würde – je nach persönlichem Standpunkt – vielleicht ergänzen, dass sich offenbar nur entsprechend zahlungskräftige Kunden per UMTS mittels „Skype“ unterhalten können.*

*Das Universitätsrechenzentrum aus dem zweiten Einführungsbeispiel wird mit der Sicherheit seiner Infrastruktur und damit in letzter Konsequenz mit der Sicherheit seiner Nutzer argumentieren; einige Nutzer würden argumentieren, dass sie ein Netzwerk schätzen, in dem sie herunterladen können, was sie möchten – es muss auch möglich sein, einen Schadcode herunterzuladen, wenn es das Forschungsinteresse erfordert. Es zeigt auch die Konsequenzen mangelnder Offenheit auf: Misstrauen gegen Infrastruktur, deren genaues Verhalten unklar ist – hier gegen die Institutsfirewall. Ein weiterer Konflikt ergibt sich mit dem grundrechtlichen Schutz der Kommunikation: Kein Datenpaket erreicht oder verlässt ganz offenbar ein Ziel im Netzwerk der Universität Rostock, ohne dass sein Inhalt durch das Rechenzentrum analysiert wurde. Ob ein Sicherheitssystem im Netzwerk als Einschränkung von Neutralität empfunden wird, ist eine Frage subjektiver Positionen und kaum objektiv zu beantworten; beide Seiten haben ihre wenig kompatiblen Argumente.*

Kernthema der Netzneutralität ist die Weiterleitung von Datenpaketen im Internet. Durch technische Entwicklungen ist es möglich geworden, einzelnen Datenströme (wie z. B. einen Download oder ein Video) voneinander zu unterscheiden und diesen unterschiedliche Prioritäten zuzuweisen (vgl. Einstiegsbeispiel 1 oder [1] für ein Endnutzer-Szenario). Aus Endkundensicht entspricht dies unterschiedlichen Qualitäten (Geschwindigkeit, Durchsatz, Jitter) des Netzwerks in Abhängigkeit von Verbindungspartner, genutztem Protokoll oder übertragenem Inhalt. Dies kann auch – wie im Einführungsbeispiel 2 – dazu führen, dass ein bestimmter Inhalt[6] nicht durch das Netzwerk übertragen werden kann.

Aus dieser Entwicklung resultieren konfliktäre Ziele: Netzwerkbetreiber interessieren sich für den Einsatz von Differenzierungstechniken, um die Beförderung von Daten mit hoher Qualität höher vergüten zu lassen (vgl. Einführungsbeispiel 1) – wahlweise nach Übereinkunft mit dem Endkunden (vgl. Einführungsbeispiel 1; [64]) oder dem Inhalteanbieter (vgl. [65]). Analoges Vorgehen ist bezüglich Diensten denkbar, die ein besonderes Datenvolumen verursachen (vgl. [66]). Viele Inhalteanbieter haben kein Interesse an einer solchen Abgabe; sie profitierten vom bisherigen Vorgehen “Best Effort“ (bei dem für jedes Datenpaket einzeln das jeweils optimale Vorgehen nur in Abhängigkeit von der Zieladresse gesucht wird) und begründen den anfänglichen Erfolg vieler heutiger Großunternehmen mit ebendieser Gleichbehandlung aller Datenpakete, da jedes kleine Garagenunternehmen im Internet die gleichen Chancen wie ein Global Player habe (vgl. [2], Rn. 35). Einige rücken jedoch von dieser Linie ab und zeigen sich offen für Vereinbarungen mit Netzbetreibern (vgl. [67]). Andere Inhalteanbieter wünschen sich indes ein Eingreifen von Zugangsnetzbetreibern bei Urheberrechtsverletzungen (vgl. [68]). Exponierte Vertreter der Nutzer befürchten öffentlich den Schritt zum „Zwei-Klassen-Internet“: Der Klasse

[6] von dem z. B. angenommen wird, dass er schädlich oder Teil eines Angriffs ist

derer, die sich nur ein „Basis-Internet“ leisten können und die Klasse derjenigen, die ein Internet höherer Performance nutzen können. Daneben ist die Bevorzugung eigener Dienste im Fall vertikal integrierter Unternehmen[7] vorstellbar – genauso wie entsprechende strategische Vorgehensweisen zur Unterdrückung unliebsamer Inhalte mit Hilfe einer entsprechenden Infrastruktur. Unregelmäßig wird die Forderung nach einer gesetzlichen Festschreibung einer Gleichbehandlung aller Datenpakete erhoben (z. B. in [69]) und auf andere Staaten verwiesen (für einen Überblick vgl. [2, 3]).

Dieses Spannungsfeld wirft vielfältige Forschungsfragen in unterschiedlichen Disziplinen auf. Wirtschaftswissenschaften und Rechtswissenschaften können hier einen Aufgabenbereich des Wettbewerbsrechts sehen: Sollten für Anbieter mit erheblichem Marktanteil spezielle Regeln gelten, die ein Ausnutzen dieser Vormachtstellung – beispielsweise zur Stärkung anderer Unternehmenszweige oder der Unterdrückung aufkommender Konkurrenten – untersagen? Der Internetzugangsanbieter aus dem Einführungsbeispiel 1 bietet selbst ebenfalls einen Telefoniedienst an. Sollte oder kann ihm das Diskriminieren von Verbindungen zu einem konkurrierenden Anbieter im Internet untersagt werden?
Die Ethik fragt, ob eine Datenstromdifferenzierung gerecht ist (oder gerechtfertigt) und stellt die Frage der Inhalteneutralität (also der Frage, welche Ideen durch ein bestimmtes Medium übertragen werden können oder dürfen) auf eine andere, philosophische Ebene: Ist es richtig, bestimmte Inhalte zu blockieren oder niedrig zu priorisieren? Gibt es einen qualitativen Unterschied zwischen beiden Eingriffen? Im Einführungsbeispiel 1 war „Skype“ ja grundsätzlich erreichbar – wenn auch nur mit sehr geringem Durchsatz; im Beispiel 2 hingegen wurde ein bestimmter Inhalt komplett ausgefiltert. Im Beispiel 1 wird der Zugriff nur für bestimmte (datenintensive) Formen der Nutzung blockiert, im Beispiel 2 für jede. Ist es ethisch vertretbar, Infrastruktur und Algorithmen zu entwerfen, die möglichst effizient Inhalte suchen um sie zu unterdrücken? Ist der Einsatz derartiger Technik zum Schutz vor aufrührerischen Gedanken vertretbar? Zum Schutz vor Malware? Und wo ist die Grenze des ethisch einwandfreien Traffic Engineerings?
Die Soziologie beschäftigt die Frage, welche Folgen diese Entwicklung für die Gesellschaft haben kann: Welche gesellschaftliche Rolle spielt der Zugriff auf das Internet, welche der Zugriff auf bestimmte Inhalte – und wie wirkte es sich aus, stünden bestimmte Inhalte (z. B. ausführbare Dateien) oder Dienste (z. B. VoIP) nur nach weiteren Zahlungen zusätzlich zu den Kosten für den Internetzugang an sich zur Verfügung? Welche Rückwirkungen hat das von Einigen vorgezeichnete Schreckensbild des „2-Klassen-Internet“ auf die ganz reale Gesellschaft? Folgt eine Spaltung der Gesellschaft in diejenigen, die sich aus den offen und ohne weitere Zahlungen im Internet verfügbaren Medien unterrichten und diejenigen, denen darüber hinaus weitere Dienste nutzen können? Was lehrt die Geschichte der Verbreitung anderer Medien und ihrer Siegeszüge durch die Gesellschaftsschichten?
Die Frage nach einem Verbot der Differenzierung stellt sich für Wettbewerbs- und Grundrechtswissenschaften. Grund- und Medienrecht verlangen nach Schutz und Diskriminie-

[7] Unternehmen, die z. B. neben dem Internetzugang auch eigene Inhalte anbieten oder anderweitig mit im Internet verfügbaren Diensten konkurrieren (z. B. beim Angebot von Telefoniediensten).

rungsfreiheit missliebiger Inhalte. Doch warum sollte ein Internetzugangsanbieter nicht nach seinen eigenen Moralvorstellungen – oder denen einer Glaubensgemeinschaft, einer gesellschaftlichen Gruppierung oder der Regierung – die Durchleitung bestimmter Inhalte durch sein Netz unterbinden? Warum sollte er seine wirtschaftlichen Interessen verraten und die Angebote konkurrierender Unternehmen nicht aus dem Bereich des Abrufbaren verschwinden lassen?
Die Kommunikationswissenschaft beobachtet die Diskussion um die „Freiheit des Internets“ mit all ihren Argumenten und Scheinargumenten. Obgleich sich alle Beiträge vorgeblich der Thematik widmen, lassen sie sich beispielsweise in Kommunikationsbeiträge, die eigentlich der Vernetzung dienen, solche die das bestehende System zu sichern versuchen und solche, die von persönlichen Differenzen getrieben sind differenzieren [4]. Welche Interessengruppe vertritt welche Meinung und weshalb?
Die Informatik schließlich sucht nach Möglichkeiten, eine stattfindende Differenzierung von Datenströmen nachzuweisen – oder eine solche Differenzierung unmöglich bzw. wenigstens sichtbar zu machen. Andere Zweige der Informatik suchen gleichzeitig nach immer ausgefeilteren Algorithmen zur Identifikation bestimmter Daten – selbst wenn diese verschlüsselt übertragen werden. Wie kann auch bei kompletter Verschlüsselung erkannt werden, dass ein Datenstrom ein Video oder ein Bild, ein Webstream von Radio Moskau oder MDR Figaro ist?

Verwandte und ebenso interdisziplinäre Fragen stellen sich auch angesichts der rasanten Entwicklung der Verarbeitung und Analyse von Daten und der Anwendung dieser Erkenntnisse in populären Webanwendungen: Ihr Verhalten ist weitestgehend undokumentiert. Dies gilt insbesondere für Suchmaschinen (vgl. [5, 6]), die in ihrer Gatekeeper-Funktion eine besondere Rolle für das Internet einnehmen[8]. Wegen ihrer Rolle für die vom Internet geprägte Gesellschaft sind sie mittlerweile Forschungsobjekt ebenso vieler Wissenschaftsgebiete geworden: Rechtswissenschaftler diskutieren das Recht juristischer Personen, in bestimmten Kontexten gefunden und in anderen Kontexten nicht gefunden zu werden, Wettbewerbswissenschaftler betrachten den von einem Quasi-Monopolisten beherrschten Suchmaschinenmarkt und weiterreichende Auswirkungen auf Kaufentscheidungen (vgl. [7, 8, 9]). Was bedeutet es rechtlich, wenn eine Suchmaschine bereits nach Eingabe eines einzelnen Buchstabens als Vervollständigung Suchterme liefert, die zu Seiten führen, auf denen ein Unternehmen oder eine Person bloßgestellt wird? Muss ein Suchmaschinenanbieter in seine Algorithmen eingreifen, wenn sie Ergebnisse liefern, die beleidigenden oder verleumderischen Charakter annehmen?
Informatiker versuchen, in Suchmaschinen verwendete Algorithmen[9] zu verstehen und suchen nach Antworten auf die Frage, wovon die Ergebnisse einer Suche – neben dem Suchbegriff – noch abhängig sind (vgl. [10]). Kommunikationswissenschaftler beobachten die alltäglichen kleinen und großen Manipulationen, mit denen bestimmte Themen befördert und andere gebremst werden. Ethiker fragen, ob diese Individualisierung von Inhalten richtig ist; schließlich handelt es sich hierbei um ein neues altes Phänomen,

[8] und die als Ersatz eines Kataloges dienen
[9] die zuvor von anderen Informatikern entwickelt und umgesetzt wurden

wenn sich der Inhalt eines Mediums oder sich die Antwort eines technischen Systems in scheinbarer Abhängigkeit vom Nutzer verändert – ein Verhalten das ein Buch nicht zeigt, wohl aber die Überlieferung durch mündliche Weitergabe. Schließlich stellt sich die Frage, wohin die Rolle der Suchmaschinen als Gatekeeper zur im Internet verfügbaren Information führt, denn bereits jetzt hängt das Ergebnis beispielsweise einer „Google"-Suche offensichtlich von deutlich mehr als nur von den eingegebenen Suchtermen ab, wie eine Suche nach dem Term „1. FC" aus Rostock bzw. Frankfurt am Main zeigt (vgl. Abb. 6.3 auf S. 130). Alle zusammen fragen, auf welcher Datenbasis eine Anwendung nach welchen Kriterien welche Ergebnisse aus welchen Gründen anzeigt.

Offensichtlich ist das Themengebiet Netzneutralität Ausgangspunkt vieler Fragestellungen, die sich in den Forschungsbereichen der Informatik, Wirtschafts- und Rechtswissenschaft sowie der Soziologie, Philosophie und Kommunikationswissenschaften bewegen. Entsprechend wichtig ist die folgende Abgrenzung, welche Themen im Rahmen dieser Arbeit bearbeitet werden – und welche nicht.

### 1.1.2. Thematische Abgrenzung dieser Arbeit

In dieser Arbeit werden die folgenden Themen detailliert behandelt:

- Ausgehend von der Frage, was eigentlich Netzneutralität ist, werden zunächst unterschiedliche Definitionen untersucht und systematisch unter Betrachtung der technischen Grundlagen eingeordnet. Denn obgleich der Begriff der Netzneutralität als etabliert angesehen werden kann ([11], [1], [12]), ist eine regelmäßige Diskussion einzelner technikabhängiger Aspekte der Definition notwendig, um den fortschreitenden Stand der Technik abzubilden: Mit der Entwicklung des Internetzugangs vom reinen Netzzugang hin zu einem gemeinsamen Medium für Netzzugang, Telefonie, Radio und Fernsehen (Next Generation Network (NGN)) sind naturgemäß auch unterschiedliche Vorstellungen von Netzneutralität verbunden.
- Bestehende Werkzeuge zum Nachweis von Neutralitätsverletzungen ([13, 14, 15, 16, 17]) werden systematisiert und die von diesen verwendeten Algorithmen untersucht und hinsichtlich ihrer Stärken und Schwächen betrachtet.
- Basierend auf den vorhergehenden Abschnitten wird der Begriff der Transparenz von Netzwerken und Anwendungen als Metaebene zur Neutralität eingeführt. Der Begriff umfasst sowohl die Spezifikation von Netzwerkpolicies wie auch die Beschreibung eines Testverfahrens bezüglich der Einhaltung deklarierter Policies.
- Als weiterer Begriff wird die Anwendungstransparenz eingeführt und ein dreistufiges Verfahren zur Erstellung transparenter Anwendungen vorgeschlagen sowie hinsichtlich seiner Umsetzbarkeit bewertet.

In dieser Arbeit gibt es keine politische, moralische oder ethische Einschätzung der einen oder anderen Position zur Netzneutralität. Genauso wenig gibt es Betrachtungen wettbewerbsrechtlicher oder wirtschaftlicher Aspekte (wie z. B. in [18]), oder einen

(detaillierten) Vergleich verschiedener Regelungen in unterschiedlichen Ländern [3, 2]. Fragen der Suchmaschinenneutralität oder der Neutralität verschiedener Dienste in sozialen Netzwerken werden am Rande gestreift (ausführlicher in [5]), stehen aber ebenfalls nicht im Mittelpunkt dieser Arbeit. Ausgehend von der aktuellen Positionierung der für eine Regelung zuständigen Bundesnetzagentur [19] wird auf einen Vorschlag für eine gesetzliche Regelung der Netzneutralität verzichtet[10]. Auch die sozialen Auswirkungen der Neutralität (oder Nicht-Neutralität) von Netzwerken werden nicht näher betrachtet, ebenso wird auf eine Betrachtung der Rolle der Netzneutralität für den Meinungspluralismus (ausführlicher in [20]) verzichtet.

Bereits die Vielfalt der Fragestellungen die sich aus dem Themenkomplex Netzneutralität ergeben, legt einen interdisziplinären Ansatz nahe. Diese Arbeit verfolgt einen solchen Ansatz, der sich der Informatik und der Rechtswissenschaft verschrieben hat. Das Hauptaugenmerk wird dabei auf dem Informatik-Anteil liegen, die Fragestellungen der Rechtswissenschaft werden jedoch stets als Rahmen und Prüfstein vorhanden sein.

Insbesondere seien Leser aus dem Fachgebiet der Rechtswissenschaften um Nachsicht bezüglich Zitierweise, Aufbau und Stil gebeten. Sie erfolgt im Interesse eines einheitlichen Vorgehens im gesamten Dokument weitestgehend entsprechend den Gepflogenheiten der Ingenieurwissenschaften. Es wurde versucht, eine „sanfte“ Annäherung zu finden.

## 1.2. Aufbau

Im folgenden Teil der Arbeit wird eine kurze Einführung in die zu Grunde liegende Netzwerktechnik gegeben. Diese Einführung beschränkt sich auf die für das weitere Verständnis dieser Arbeit wesentlichen Grundlagen. In Kapitel 3 wird der bestehende Forschungsstand zur Netzneutralität wiedergegeben; das darauffolgende Kapitel systematisiert und bewertet den Forschungsstand. Hierbei liegt ein besonderer Fokus auf den Techniken zum Nachweis von Brüchen der Netzneutralität, da diese im Folgekapitel als Grundlage eines Verfahren zum Nachweis der Einhaltung von Netztransparenz dienen werden. Neben diesen technischen Aspekten der Netzneutralität wird auch der rechtliche Rahmen für Vorgaben zur Netzneutralität in der Bundesrepublik Deutschland und in der Europäischen Union betrachtet. In einem kurzen Abriss wird die öffentliche Debatte um die Netzneutralität in der Bundesrepublik dargestellt. In Kapitel 5 wird die überprüfbar transparente Gestaltung von Netzen im Detail ausgeführt. Kernidee ist die Information des Benutzers, indem Netzwerkbetreiber zur Publikation von Regelsätzen verpflichtet werden, die die Auswirkungen ihrer Netzwerkpolicies zusammenfassen. Ferner wird unter Rückgriff auf einen aus Kap. 4.2 bekannten Ansatz eine Methodik zur Überprüfung der angegebenen Netzwerkregeln skizziert. Kapitel 6 generalisiert den Gedanken der Transparenz von Netzwerken schließlich auf die Transparenz von Anwendungen. Dabei wird – getrieben von alltäglichen Beobachtungen mehrerer Jahre – das Phänomen der Anwendungstransparenz beschrieben und es werden mögliche Motive für eine intransparente

[10] Diese wäre z. B. in Form eines Vorschlags einer Verordnung nach § 41a TKG denkbar.

Gestaltung von Anwendungen gesucht. Das Augenmerk liegt hierbei besonders auf Webanwendungen, da sich bei diesen die Programmlogik besonders einfach einer Analyse durch den interessierten Nutzer entziehen kann. Mehr noch: Sie kann vom Anbieter jederzeit verändert werden. Schließlich werden Vorschläge zur transparenteren Gestaltung von Anwendungen beschrieben. Abschließend werden die wesentlichen Ergebnisse dieser Arbeit zusammengefasst und bewertet.

# 2. Einführung in Computernetzwerke

> Das Internet ist ein großer Müllhaufen, in dem man allerdings auch kleine Schätze und Perlen finden kann.
> —*Joseph Weizenbaum*

An dieser Stelle werden die für das weitere Verständnis dieser Arbeit vorausgesetzten Begriffe aus dem Netzwerkbereich eingeführt. Weitergehende Informationen hierzu finden sich insbesondere in Tannenbaum [21].

Werden Daten im Internet übertragen, so geschieht dies in Form einzelner Pakete. Dabei enthält jedes Datenpaket eine (numerische) Absender- wie auch eine Zieladresse. Da diese Adressen in numerischer Form schlecht zu merken sind, gibt es mit DNS einen Dienst, der menschenlesbare Namen in Adressen umwandelt. Wenn sich Absender und Empfänger eines Datenpakets nicht innerhalb eines lokalen Netzes befinden (was bei Internetverbindungen die Regel ist), müssen Datenpakete weitergeleitet (geroutet) werden. Hierbei können weitere Netze durchquert werden. Ohne Kap. 3 vorgreifen zu wollen: Fragen der Netzneutralität stellen sich in diesem Zusammenhang insbesondere bezüglich der Weiterleitung von Paketen und der Prioritäten mit der einzelne Pakete behandelt (weitergeleitet) werden. Beim Nachweis von Neutralitätsverletzungen werden entsprechend die Eigenschaften der Paketweiterleitung untersucht. Fragen des Datenschutzes stellen sich besonders bezüglich der Adressierung und der Nutzlast eines Datenpaketes.

In diesem Kapitel wird zunächst die Adressierung einzelner Maschinen im Netzwerk anhand von Hardware- und IP-Adresse erläutert; hieran schließen sich Erläuterungen des Domain-Name-Systems wie des Routings an, gefolgt von einer Einführung der Kennzahlen zur Bewertung der Qualitäten eines Netzwerkes. Schließlich werden in Vorbereitung auf das Thema Netzneutralität einzelne Ansätze zu Erkennung von Datenströmen in Netzwerken diskutiert. Den Abschluss bildet ein Überblick über rechtliche Aspekte, die im Kontext von Neutralitäts- oder Transparenzprüfungen von Interesse sind.

## 2.1. Adressierung

In diesem Abschnitt wird die grundlegende Adressierung von Geräten im Netzwerk kurz dargelegt. IP-Adressen identifizieren Computer in verteilten Netzwerken wie dem Internet. Obwohl IP-Adressen weder technisch an einzelne Maschinen noch einzelne Maschinen an exakt eine IP-Adresse gebunden sind, kann im Allgemeinen davon ausgegangen

werden, dass eine IP-Adresse einen einzelnen Knoten im Netzwerk identifiziert[1]. Bei IP-Adressen handelt es sich um eine rein logische Adressierung; sie wird klassischerweise vom Betriebssystem und nicht von der Hardware einer Maschine behandelt[2].

Zum gegenwärtigen Zeitpunkt verwenden die meisten mit dem Internet verbundenen Rechner noch IP-Adressen mit einer Länge von 32 Bit (IPv4). Seit langem ist eine Umstellung auf Adressen mit einer Länge von 128 Bit (IPv6) geplant; diese ist bislang jedoch noch nicht flächendeckend erfolgt. In dieser Arbeit wird daher weiterhin von IPv4-Adressen ausgegangen, soweit dies relevant ist.

IP-Adressen können statisch oder dynamisch vergeben werden. Im Falle einer statischen Zuweisung ist eine IP-Adresse einem bestimmten Gerät fest zugeordnet und wird typischerweise händisch festgelegt. Dynamische IP-Adressen werden im Unterschied hierzu heute fast ausschließlich mithilfe des DHCP (Dynamic Host Configuration Protocol) vergeben und von einem mit einem Netzwerk verbundenen Gerät nach Herstellen einer physikalischen Verbindung erfragt. Bei DSL-Internetzugängen ist es üblich, nach 24 Stunden eine Zuweisung einer anderen IP-Adresse vorzunehmen; hierfür gibt es keine (vom Netzwerk ausgehende) Notwendigkeit. In der weiteren Nutzung bestehen zwischen statischen und dynamischen IP-Adressen keine Unterschiede.

## 2.2. Namensauflösung

Computer identifizieren sich gegenseitig anhand der erläuterten Adressen; im Internet zumeist anhand von IPv4-Adressen. Menschliche Anwender bevorzugen (aussprechbare oder merkbare) Namen. Diese werden mithilfe des Domain Name Systems (DNS) in IP-Adressen umgewandelt („aufgelöst“). DNS-Server sind als zentral organisiertes, hierarchisches System aufgebaut. Im Kontext der Netzneutralität ist insbesondere die Manipulationsanfälligkeit des DNS von Bedeutung: Durch Fälschung oder Manipulation einer DNS-Antwort kann ein Nutzer fehlgeleitet werden; wird eine Antwort auf eine DNS-Anfrage hingegen unterdrückt[3], so wird der Domainname von Anwenderprogrammen typischerweise als nicht vorhanden betrachtet[4]. Wird die Antwort gefälscht oder manipuliert, wird der Nutzer einen anderen Rechner kontaktieren, als ursprünglich beabsichtigt.

Für eine einzelne IP-Adresse können mehrere Domainnamen definiert sein. Es gibt eine Umkehrung, mit deren Hilfe eine IP-Adresse zu (genau) einem Domainnamen aufgelöst werden kann (allerdings muss nicht für jede IP-Adresse ein Domainname hinterlegt sein; außerdem werden hierbei nicht alle auf eine Adresse verweisenden Namen enthalten sein).

[1] Abweichungen hiervon finden sich insbesondere im Serverbereich.
[2] wiewohl es aus Performancegründen Ausnahmen gibt
[3] z. B. herausgefiltert
[4] Oder es gibt eine entsprechende Fehlerbehandlung seitens der Anwendung, welche bei fehlender Antwort einen Timeout meldet. Fraglich bleibt, wie eine Anwendung im Anschluss hieran weiter verfährt.

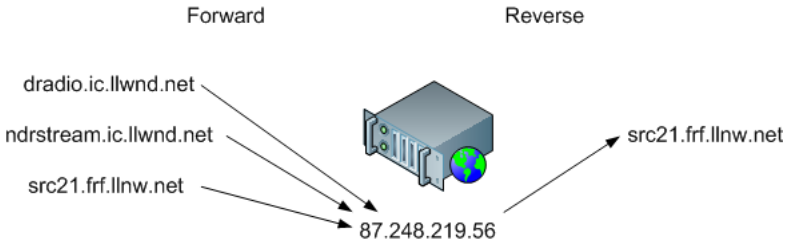


Abbildung 2.1.: DNS: Forward- und Reverse-Auflösungen am Beispiel eines Streaming-Media-Anbieters (Stand 2011).

Dieser sogenannte „reverse lookup“ kann zu für Endanwender nicht hilfreichen Domainnamen führen; häufig wird ein für die Techniker des jeweiligen Betreiberunternehmens hilfreicher Name verwendet, der Rückschlüsse auf die physikalische oder virtuelle Maschine erlaubt (s. Abb. 2.1). Das Domain Name System ist hierarchisch organisiert: Um `dradio.ic.llwnd.net` in eine Adresse aufzulösen, wird zunächst ein für die Top-Level-Domain `net` zuständiger DNS-Server nach Daten zu `llwnd` angefragt. Diese Antwort wird u.a. den Verweis auf einen DNS-Server enthalten, der untergeordnete Domains von `llwnd` auflösen kann – entsprechend kann rekursiv schließlich die IP-Adresse von demjenigen DNS-Server erfragt werden, der der Domain `ic.llwnd.net` zugeordnet ist.

Die Manipulationsanfälligkeit von DNS ist insbesondere deshalb bedeutsam, weil sich auch Absicherungsmechanismen wie SSL (z. B. in https) in Zertifikaten auf die Domainnamen beziehen. Somit wäre mit einer gefälschten DNS-Antwort und einem gefälschten Zertifikat die perfekte Vorspiegelung einer fremden Identität möglich. Allerdings ist davon auszugehen, dass ein Missbrauch auch bei einem nicht-validen Zertifikat möglich wäre, da ein wesentlicher Teil der Nutzer ein fehlerhaftes Zertifikat nicht ernsthaft als Warnsignal betrachtet, wie beispielsweise die Beobachtungen im Rahmen der Diplomarbeit von Rene Lindhorst 2007 [22] zeigten.

## 2.3. Routing

Als Routing wird die Paketweiterleitung in Abhängigkeit von der IP-Adresse des Ziels bezeichnet. Das Internet entsteht durch Routing: Erst durch die zielgerichtete Paketweiterleitung wird es möglich, dass auch die in den Einzelnetzen angeschlossenen Rechner von jedem anderen Rechner aus erreichbar sind und die lokalen Netze an Bedeutung verlieren. Dabei werden die Netze einzelner Betreiber als „autonome Systeme“ (AS) bezeichnet, da sie unter eigenständiger Verantwortung und Verwaltung stehen.

Ein einzelner Router ist eine Maschine mit Verbindung zu mindestens zwei Netzwerken die Datenpakete entgegennimmt, deren Empfänger sich nicht im lokalen Netz befinden.

Der Router entscheidet anhand der Zieladresse des Datenpakets, an welchen anderen Rechner (und ggf. über welche Verbindung) diese Pakete weiterzuleiten sind. Hierzu ist es notwendig, dass der Router Informationen darüber besitzt, wie ans Internet angeschlossene Rechner erreicht werden können. Diese Informationen werden entweder über geeignete Protokolle weitergegeben oder statisch von administrativer Seite vorgegeben.

Ein einzelner Internetanschluss hat – insbesondere im Bereich privater Nutzer – eine einzelne IP-Adresse. Um diesen Anschluss mit mehreren Endgeräten nutzen zu können, ist es üblich, Datenpakete weiterzuleiten und ihnen dabei eine andere Absenderadresse zu erteilen (Network Address Translation, NAT). Dies geschieht insbesondere durch für den privaten Internetzugang spezialisierte Router: Der Router erhält vom Internetzugangsanbieter die „öffentliche“[5] IP-Adresse; die Rechner im privaten Netzwerk des Nutzers hingegen erhalten IP-Adressen aus einem hierfür reservierten Adressbereich, der nicht anderweitig genutzt wird. Bei einem Internetzugriff werden die Datenpakete von der Gateway mit der öffentlichen IP-Adresse als Absender weitergeleitet. Die entfernte Maschine sendet ihre Antwort an die öffentliche Adresse der absendenden Gateway. Die Antwort wird von der Gateway schließlich an den ursprünglichen Kommunikationspartner adressiert und diesem zugestellt. Durch diese Technik ist es möglich, dass sich mehrere Rechner in einem privaten Netzwerk einen Internetzugang teilen können.

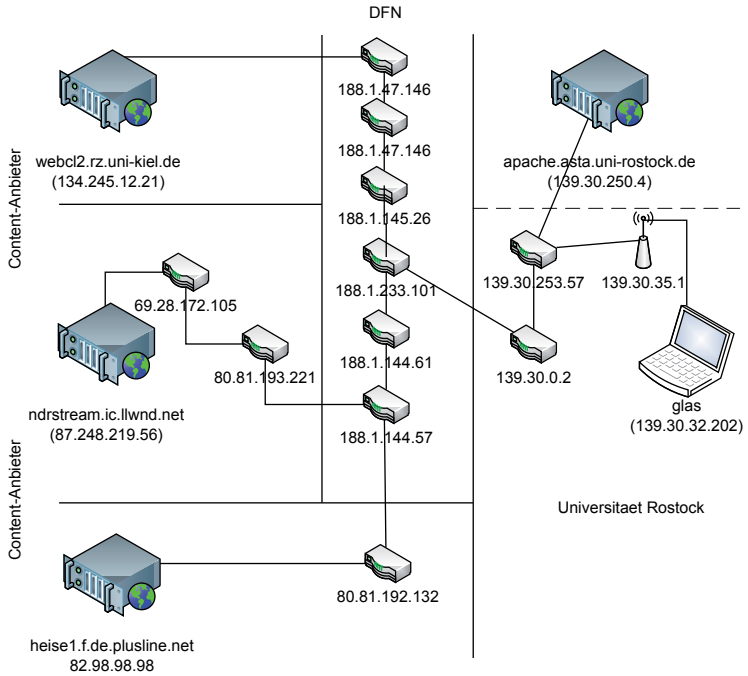


Abbildung 2.2.: Beispiel für Routingpfade, entnommen aus [23].

[5] d.h. weltweit eindeutige und weltweit erreichbare

Auf dem Weg vom Sender zum Empfänger passieren Datenpakete üblicherweise mehrere Router; hieraus ergibt sich ein Routingpfad (siehe Abb. 2.2).

Zu beachten ist hierbei, dass jeder der beteiligten Router zur Manipulation transportierter Datenpakete in der Lage ist; eingesetzte Prüfsummen dienen lediglich der Erkennung von Übertragungsstörungen, nicht der Sicherung der Integrität der Datenübertragung. Einige Manipulationen (wie etwa NAT, s.o.) sind absolut erwünscht und notwendig; andere Manipulationen hingegen sind es nicht. Aus Sicht des Endpunktes einer Kommunikation entsteht hier ein Dilemma: Ohne weitere Informationen ist nicht unterscheidbar, ob ein Datenpaket in der vorliegenden Form tatsächlich von der Gegenstelle stammt, ob es zwar grundsätzlich von der Gegenstelle stammt aber manipuliert wurde oder aber ob es von einer Station des Routing-Pfades injiziert wurde und gar nicht von der angenommenen Gegenstelle stammt.

Beispielhaft sollen hier drei Szenarien der Manipulation übertragener Datenpakete im Detail untersucht werden: Ein Fall von gewünschtem Adressspoofing, ein Fall eines Verbindungsabbruchs und ein Fall der Manipulation übertragener Daten. Alle Beispiele wurden bereits in [23] bearbeitet.

**Beispiel 1: Adressspoofing** Aus Sicht eines Nutzers zeigte das drahtlose Campusnetzwerk der Universität Rostock lange Zeit[6] folgendes Verhalten: Bevor sich der Nutzer nicht mit seinem Rechenzentrumsaccount gegenüber einer bestimmten Webseite des „BlueSocket"-Systems authentifiziert hatte, konnten keine Datenpakete mit anderen Stationen ausgetauscht werden. HTTP-Verbindungen zu anderen Zielen wurden automatisch auf diese BlueSocket-Login-Seite umgelenkt. Die Netzwerktopologie entspricht der Skizze in Abb. 2.2.

Im Hintergrund ist hierfür ein HTTP-Redirect zuständig – allerdings könnte dieser ohne Manipulationen nur von der eigentlich angefragten Seite kommen. Dies ist jedoch aus mehreren Gründen höchst unwahrscheinlich – beginnend damit, dass z. B. der Server von dem `heise.de` angeboten wird nicht wissen kann, wann ein Nutzer auf eine (welche?) Login-Seite umgeleitet werden soll.

Denkbar ist ebenfalls eine Manipulation von DNS-Informationen, die an den Client ausgeliefert werden (bis eine Authentifikation erfolgt ist), denn auch auf diesem Wege ließen sich Anfragen umleiten. Der Browser würde versuchen, `heise.de` in eine IP-Adresse aufzulösen – doch der DNS-Server würde anstelle der tatsächlich von `heise.de` verwendeten IP-Adresse zunächst die der „BlueSocket"-Login-Seite zurückliefern. Dieser Ansatz wäre denkbar, hätte jedoch das Problem, dass nicht garantiert ist, dass für jeden Aufruf erneut eine DNS-Auflösung vorgenommen wird – die bewusst falsch ausgelieferte Adresse könnte zwischengespeichert werden und bei weiteren Verbindungsversuchen zu Irritationen führen. Daneben würde diese Lösung für alle Nutzer fehlschlagen, die einen fest eingestellten anderen DNS-Server verwenden.

Tatsächlich wurde vom „BlueSocket"-System Addressspoofing eingesetzt: Bevor einer IP-Adresse eines im WLAN befindlichen Geräts ein authentifizierter Nutzer zugeord-

[6] während der Nutzung der „BlueSocket„-Authentifikation

net werden konnte, wurden von diesem abgesandte HTTP-Anfragen stets mit Redirects beantwortet.

Auf Ebene des HTTP-Protokolls wurde eine Anfrage wie etwa

```
GET / HTTP/1.1
Host: heise.de
Accept: */*
```

an den zu `heise.de` gehörenden Server mit der IP-Adresse 193.99.144.80 an Port 80 gestellt. Der Browser des Nutzers erhielt die Antwort

```
HTTP/1.1 302 Found
Date: Fri, 04 Mar 2011 13:23:17 GMT
Server: Apache
Location: http://blue4.wlan.uni-rostock.de/login.pl?action=
which_interface&destination=http://heise.de/%3f
Connection: close
Transfer-Encoding: chunked
Content-Type: text/html; charset=iso-8859-1

118
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
<HTML><HEAD>
<TITLE>302 Found</TITLE>
</HEAD><BODY>
<H1>Found</H1>
The document has moved <A HREF="http://blue4.wlan.uni-rostock.de/login.pl?
action=which_interface&amp;destination=http://heise.de/%3f">here</A>.<P>
</BODY></HTML>

0
```

Ganz offensichtlich wurde die Anfrage nach der Ressource / auf dem Server `heise.de` durch ebendiesen Server mit einem Redirect zu `http://blue4.wlan.uni-rostock.de/[...]`, beantwortet.

Dieser Redirect führte zur Login-Seite des „BlueSocket“-Systems. Eine detaillierte Betrachtung der ausgetauschten Daten zeigt folgende Artefakte (gezeigt werden die Datenpakete der Anfrage und der zweite Teil des Headers sowie der Body der Antwort):

```
14:32:35.503903 IP 139.30.32.197.59424 > 193.99.144.80.http: Flags [P.],
ack 1, win 8326, options [nop,nop,TS val 1091444 ecr 2443535463], length 47
E..c.U@.@.:... ..c.P. .P......(... ........
...t..dgGET / HTTP/1.1
Host: heise.de
Accept: */*
```

```
[...]

14:32:35.519615 IP 193.99.144.80.http > 139.30.32.197.59424: Flags [FP.],
seq 65:568, ack 48, win 1448, options [nop,nop,TS val 2443535486 ecr
1091444], length 503
E..+L.@.@....c.P.. ..P. ..)%..............
..d~...t Apache
Location: http://blue4.wlan.uni-rostock.de/login.pl?action=
which_interface&destination=http://heise.de/%3f
Connection: close
Transfer-Encoding: chunked
Content-Type: text/html; charset=iso-8859-1

118
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
<HTML><HEAD>
<TITLE>302 Found</TITLE>
</HEAD><BODY>
<H1>Found</H1>
The document has moved <A HREF="http://blue4.wlan.uni-rostock.de/login.pl?
action=which_interface&amp;destination=http://heise.de/%3f">here</A>.<P>
</BODY></HTML>

0
```

Scheinbar werden Datenpakete zwischen den Hosts 139.30.32.197 auf Port 59424 und 193.99.144.80 auf Port 80 (http) im Rahmen einer bestehenden TCP-Verbindung ausgetauscht. Dieses Verhalten ist nur innerhalb des Campus-WLANs der Universität Rostock zu beobachten gewesen; eine Verbindung von `apache.asta.uni-rostock.de` zeigte dieses Verhalten nicht. All diese Beobachtungen legen den Schluss nahe, dass die Verbindung nicht mit dem Server von `heise.de` sondern mit einer Komponente des WLAN-Systems erfolgte, die sich zumindest für diese Beantwortung die IP-Adresse eines anderen Servers zu eigen machte.

Was in diesem Szenario wie eine elegante und nutzerfreundliche Lösung der Frage, wie man Nutzer zur Authentifikationsseite leiten kann, ohne dass diese sich eine URL merken müssen, wirkt, hat eine Schattenseite. Denn es gibt keine Instanz, die dem Nutzer garantiert, dass diese Technologie alleinig zu seinem Schutz und Komfort eingesetzt wird; es gibt keine Information darüber und es gibt keine Garantie, dass ein solches Vorgehen auf die Nutzung vor einem Login beschränkt ist. Ein bösartiger Betreiber könnte auch andere Umleitungen vornehmen – angedacht sei nur die Umleitung von einer Bankseite auf eine Phishing-Seite.

**Beispiel 2: Inhaltsabhängiger Verbindungsabbruch** Ein Nutzer verwendet die chinesische Suchmaschine `baidu.cn`. Diese zeigt ein seltsames Verhalten: Wann immer

der Nutzer sich Ergebnisse zu bestimmten Suchbegriffen anzeigen lassen möchte, meldet der Browser ein Netzwerkproblem.

Zunächst die Ausgabe von `tcpdump` im Fall des erwarteten Verhaltens (bei einer Suche mit dem Suchbegriff „germany“):

```
14:44:28.966833 IP wlan033084.uni-rostock.de.11479 > 220.181.111.147.http:
 Flags [P.], ack 1, win 65535, length 236
E....3@.@.....!T..o.,..P.J;..BO.P.......GET /s?wd=germany HTTP/1.1
Host: www.baidu.com
User-Agent: ELinks/0.11.7 (textmode; FreeBSD 8.0-BETA2 i386; 197x67-2)
Referer: http://www.baidu.com/
Accept: */*
Accept-Encoding: gzip
Accept-Language: en
Connection: Keep-Alive


14:44:29.273641 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [.], ack 237, win 6432, length 0
E..(.|@.,.....o...!T.P,..BO..J<.P.. Sb..
14:44:29.309478 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [P.], ack 237, win 6432, length 369
E....~@.,..%..o...!T.P,..BO..J<.P.. .x..HTTP/1.1 200 OK
Date: Mon, 04 Apr 2011 12:44:28 GMT
Server: BWS/1.0
Content-Length: 10697
Content-Type: text/html;charset=gbk
Cache-Control: private
Content-Encoding: gzip
Set-Cookie: BAIDUID=704188E60C2467232B6E99245E32E26C:FG=1; expires=Mon,
 04-Apr-41 12:44:28 GMT; path=/; domain=.baidu.com
P3P: CP=" OTI DSP COR IVA OUR IND COM "
Connection: Keep-Alive
```

Als Reaktion auf die Anfrage beginnt die Gegenstelle mit der Auslieferung einer HTML-Seite als Antwort. Wird jedoch als Suchbegriff „falun gong“ verwendet, ist ein anderes Verhalten zu beobachten:

```
14:45:11.197146 IP wlan033084.uni-rostock.de.11479 > 220.181.111.147.http:
Flags [P.], ack 2789374741, win 65535, length 241
E.....@.@.....!T..o.,..P.J<..B{.P...sZ..GET /s?wd=falun%20gong HTTP/1.1
Host: www.baidu.com
User-Agent: ELinks/0.11.7 (textmode; FreeBSD 8.0-BETA2 i386; 197x67-2)
Referer: http://www.baidu.com/
```

```
Accept: */*
Accept-Encoding: gzip
Accept-Language: en
Connection: Keep-Alive


14:45:11.502060 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [.], ack 241, win 7504, length 0
E..(..@.,..z..o...!T.P,..B{..J<.P..P#...
14:45:11.502466 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [R.], seq 1, ack 241, win 2390, length 0
E..(..@.|.yh..o...!T.P,..B{..J<.P.      V6...
14:45:11.503090 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [R.], seq 1461, ack 241, win 2391, length 0
E..(..@.}.xk..o...!T.P,..B...J<.P.      W1H..
14:45:11.503432 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [R.], seq 4381, ack 241, win 2392, length 0
E..(..@.~.u...o...!T.P,..B.1.J<.P.      X%...
```

Zu beachten sind die gesetzten TCP-Flags der Antwort: Die Verbindung wird mittels gesetztem `RST`-Bit im TCP-Header zurückgesetzt. Die Reaktion des Browsers ist typischerweise eine entsprechende Meldung an den Nutzer.
Ob dieses Verhalten von einer Routing-Station oder aber von der Maschine mit der IP-Adresse 220.181.111.147 stammt, ist aus der Position eines Nutzers heraus nicht klar zu trennen.

**Beispiel 3: Manipulation übertragener Daten** Beim Versuch, aus dem Wohnheimnetz des Studentenwerks Rostock eine verschlüsselte Verbindung mit einem SMTP-Server aufzubauen, meldet der Mail-Client, dass Verschlüsselung per `STARTTLS` vom Server nicht unterstützt werden würde. Der gleiche Mail-Client kann erfolgreich eine verschlüsselte Verbindung zu ebendiesem Server aufbauen, wenn er sich aus einem anderen Netzwerk verbindet.

Bei Betrachtung der ausgetauschten Daten auf Ebene des SMTP-Protokolls wird deutlich, dass der Server je nach Netzwerk ein unterschiedliches Verhalten zeigt. Zunächst ein Auszug aus dem Datenaustausch, wie er sich abspielt, wenn eine verschlüsselte Verbindung z. B. von einem privaten DLS-Anschluss oder einem Rechner im WLAN des Universitätsnetzes aus hergestellt werden soll:

```
220 mail.gmx.net GMX Mailservices ESMTP {mp003}
EHLO tralalal
250-mail.gmx.net GMX Mailservices
250-8BITMIME
250-ENHANCEDSTATUSCODES
```

```
250-SIZE
250-AUTH=LOGIN PLAIN
250-AUTH LOGIN PLAIN
250 STARTTLS
```

Zum Vergleich hier der gleiche Datenaustausch aus dem Netzwerk des Studentenwohnheims:

```
220 ****************************************
EHLO tralala
250-mail.gmx.net GMX Mailservices
250-8BITMIME
250-ENHANCEDSTATUSCODES
250-SIZE
250-AUTH=LOGIN PLAIN
250-AUTH LOGIN PLAIN
250 XXXXXXXA
```

Eine Betrachtung auf weiter unterliegenden Schichten bringt keine weiteren Erkenntnisse.

Scheinbar werden bestimmte Teile der Daten vom SMTP-Server zum Client durch andere Zeichen ersetzt; u.a. auch die Angabe über die Unterstützung von `STARTTLS`. Als Reaktion hierauf meldet der Mail-Client dem Nutzer, dass eine entsprechende Verschlüsselung nicht vom Server unterstützt würde, obwohl diese vom Server unterstützt werden. Im Ergebnis werden viele Nutzer auf die Verschlüsselung verzichten und ihre E-Mails ungeschützt abrufen.

Solchen Problemen der Manipulation übertragener Daten kann bedingt[7] mit kryptographischen Mitteln begegnet werden: Eine Verschlüsselung des Paketinhaltes stellt idealerweise neben der Vertraulichkeit übermittelter Daten auch Integrität sicher. Allerdings gibt es für Routingstationen weiterhin Mittel des Eingreifens: Auch Ende-zu-Ende Verschlüsselung verlässt sich im Hintergrund weiterhin auf bekannte Protokolle wie TCP – und deren Header liegen im Datenpaket weiterhin unverschlüsselt vor, können also manipuliert werden (z. B. können die Felder für einen Verbindungsabbruch gesetzt werden, was dann bei der empfangenden Station den Eindruck erwecken würde, das Gegenüber habe die Verbindung zurückgesetzt). Ferner kann jederzeit auch ein Paket mit unsinnigen Daten eingespeist werden, welches die Kommunikation stört.

Die Ursache dieser Probleme ist historisch: Zur Zeit der Entwicklung heute dominierender Protokolle (wie IP, TCP, UDP) konnten alle Verbindungen als gutartig angenommen werden, da sie sich wahlweise in militärischer oder universitärer Obhut befanden (und von böswilligen Manipulationen durch Computersysteme hinweg nicht ausgegangen wurde, vgl. [24]). Es konnte ferner davon ausgegangen werden, dass Probleme durch

[7] soweit von den Endknoten bzw. verwendeten Programmen unterstützt

Manipulationen an Verbindungen unter Militäraufsicht ohnehin erst zu einem Zeitpunkt auftreten, da es bereits wesentlich dringlichere Probleme als die Netzwerkstörungen gäbe.

Seit Bestehen des Internets hat es viele Entwicklungen auf dem Gebiet der routingbasierten Lastverteilung gegeben; praktisch genutzt Algorithmen zur Lastverteilung sind dabei in der Lage, Daten paketweise auf redundante Verbindungswege zu verteilen ([25], [70]). Hierbei können sowohl technische (z. B. gleichmäßigere Auslastung) als auch wirtschaftliche (Bevorzugung einer kostengünstigeren Verbindung) Kriterien eine Rolle spielen. Nachteil dieser Entwicklung ist allerdings, dass traditionelle Werkzeuge wie z. B. `traceroute` nicht mehr gewohnt zuverlässig arbeiten und auf Optimierungen zurückzuführende Artefakte wie Zyklen zeigen. Solchen Problemen kann mit neuen Werkzeugen begegnet werden [26].

Betreiber von Netzwerken haben ein natürliches Interesse daran, dass ihre Netze nicht überlastet sind und daher ausfallen – sei es, weil Kunden eine bestimmte Ausfallsicherheit zugesichert ist oder schlicht aus Gründen des Renommees. Denn im Fall nicht ausreichender Lastverteilung oder einer anders ausgelösten Überlastung kann es zu Ausfällen kommen (indem z. B. einzelnen Komponenten mehr Daten zugesandt werden als diese verarbeiten können (Congestion)). Folgen hieraus können langsamere Weiterleitung, geringerer Durchsatz oder auch komplette Nichterreichbarkeit von Teilen des Netzwerkes sein. Ein Ausfall eines Netzwerks aufgrund einer Congestion ist für einen Beobachter am Endpunkt einer Verbindung auf Grund der Routenabhängigkeit nicht von anderen Ursachen (z. B. einer Neutralitätsverletzung oder einem Serverausfall) unterscheidbar. Gleichzeitig kann in der Ausfallsicherheit auch eine Motivation für Neutralitätsverletzungen im Sinne der Beschränkung des Durchsatzes, der bestimmten Arten von Datenübertragungen auferlegt wird[8] vermutet werden: Durch das Reglementieren bestimmter Lasten bleiben Kapazitäten frei und Congestions werden vermieden.

Im Internet Protocol ist zwar die Möglichkeit des Source Routings, also der Vorgabe eines bestimmten Routingpfades unabhängig von einzelnen Systemen vorgesehen; aus Sicherheitsgründen wird dies heute jedoch nicht mehr angewandt. Im Ergebnis hat der Absender eines Paketes keinen Einfluss auf den Routingpfad.

## 2.4. Quality of Service (QoS)

Mit dem Begriff „QoS“ wird umgangssprachlich die Zusicherung bestimmter Netzwerkparameter (potentiell zulasten anderer Verbindungen) verstanden; zum anderen bezeichnen die QoS die Qualitätsparameter einer Netzwerkverbindung. Grundsätzlich gibt es (nach [21]) eine vorgegebene Menge von Qualitäten einer Netzwerkverbindung[9]:

- Durchsatz – Menge der übertragenen Daten pro Zeiteinheit.

[8] z. B. dem Ausschluss ressourcenintensiver Anwendungen wie VoIP im Einführungsbeispiel 1
[9] Liste aus [21], Erläuterungen durch den Autor.

- Latenz – Zeitliche Verzögerung bei der Datenübertragung; oft angegeben als Round-Trip-Time.
- Jitter – Je nach Detailliertheit der Betrachtung kann grob vereinfachend von der Änderung der Latenz über der Zeit gesprochen werden oder aber der Jitter kann als ein statistisches Maß betrachtet werden, das sich aus der leicht unterschiedlich lange andauernden Verarbeitung jedes einzelnen Datenpaketes bei der Weiterleitung ergibt.
- Zuverlässigkeit – Anzahl der während der Übertragung durch technische Fehler zerstörten Datenpakete.
- Kollisionen je Zeiteinheit – Durch z. B. gleichzeitiges Aussenden von Paketen[10] benötige Neuübertragungen.

Unterschiedliche netzbasierte Anwendungen haben unterschiedliche Anforderungen hinsichtlich dieser Qualitäten, wie in Tabelle 2.1 illustriert: Es gibt Anwendungen wie VoIP und Echtzeitspiele, für die eine geringe Verzögerung bei der Datenübertragung und eine Konstanz dieser Verzögerung wesentlich ist. Hiervon sind Anwendungen wie ein reiner Datentransfer zu unterscheiden, bei denen diese Parameter im Vergleich zum verfügbaren Durchsatz vernachlässigbar sind.

| Anwendung | Durchsatz | Latenz | Jitter |
|---|---|---|---|
| Webradio | moderat | kein Einfluss | möglichst klein |
| Echtzeitspiel | relativ gering | möglichst klein | möglichst klein |
| Dateiübertragung | möglichst hoch | kein Einfluss | kein Einfluss |
| Chat | relativ gering | kein Einfluss | kein Einfluss |
| Hochauflösendes Fernsehen | möglichst hoch | kein Einfluss | möglichst klein |
| VoIP-Telefonie | moderat | möglichst klein | möglichst klein |

Tabelle 2.1.: QoS-Anforderungen einiger populärer Netzanwendungen

Im allgemeinen Sprachgebrauch wird bei „QoS“ meist die Optimierung oder Zusicherung eines oder mehrerer dieser Netzwerkparameter gemeint, beispielsweise die Sicherstellung eines bestimmten Durchsatzes für einen Datentransfer oder einer möglichst geringen Latenz für Anwendungen wie etwa netzwerkbasierte Mehrbenutzerspiele. Ein anschauliches Beispiel ergibt sich aus der Verwendung eines kombinierten Internet- und Fernsehanschlusses. Damit das Fernsehprogramm auch während eines umfangreichen Downloads nicht abbricht, muss dem IPTV-Datenstrom eine Mindestqualität gewährt werden.

Häufig wird auch von „QoS-Klassen“ gesprochen. Eine solche Klasse ist stets durch identische QoS-Vorgaben (z. B. möglichst geringe Latenz) gekennzeichnet. In einer Installation können beispielsweise drei QoS-Klassen definiert werden: Eine mit hoher Priorität,

[10] Versuch von zwei Sendern, gleichzeitig auf ein gemeinsames Medium schreibend zuzugreifen.

z. B. für VoIP-Anwendungen; eine mit niedriger Priorität, z. B. für Dateiübertragungen; eine für die restlichen Datenpakete. Die Datenpakete können entsprechende Markierungen erhalten, die von Routern erkannt werden und zur Priorisierung einzelner Datenpakete verwendet werden. Wie diese Markierung von Datenpaketen erfolgt, ist Thema des unmittelbar folgenden Abschnitts.

QoS wird häufig mit Problemstellungen aus dem Bereich der Netzneutralität korreliert, da sich Methoden zur QoS-Zusicherung als Werkzeuge zur Verletzung der Netzneutralität anbieten. In diesem Kontext wird bisweilen sehr plakativ das Bild des Zwei-Klassen-Internets gemalt: Die (QoS-)Klasse derer, die ein High-Priority-Internet genießen und die Klasse derer, die das verbleibende Low-Priority-Internet behalten.
Alleinig die Bereitstellung von QoS, also einer unterschiedlichen Behandlung von Paketen durch Router ist noch nicht problematisch (denn in einem solchen Szenario könnte jede Anwendung ihre Pakete potentiell auch selbst als „wichtig“ oder „weniger wichtig“ klassifizieren). Probleme ergeben sich erst aus dem (vom Nutzer unbemerkten) Einsatz von Methoden zur Identifikation von Datenströmen (s.u.), mit deren Hilfe Datenströme unterschiedlichen QoS-Klassen zugeordnet werden und aus der sich somit – je nach Anwendung – unterschiedliche Verbindungscharakteristiken ergeben, ohne dass der Endanwender dies beeinflussen kann.

Beeinflussungen der Netzqualität lassen sich dann auf Einschränkungen durch QoS-Regelungen (die andere Datenströme privilegieren und eine Neutralitätsverletzung darstellen können) oder auf Congestion einzelner Netzwerksegmente oder -knoten bzw. aus Kombinationen aus beidem zurückführen. Die Differenzierung beider möglicher Ursachen ist wesentliches Problem des Nachweises von Neutralitätsverletzungen.

## 2.5. Identifikation von Datenströmen

Als ein „Datenstrom“ wird die Menge aller Datenpakete verstanden, die aus Anwendungssicht semantisch eine „Verbindung“ (also einen Kommunikationskanal zwischen zwei Anwendungen) bilden. Dabei ist (für die Identifikation eines Datenstroms) unerheblich, ob sich die Anwendungen eines Protokolls bedienen, das selbst einen Verbindungsbegriff besitzt (wie etwa TCP) oder nicht (wie etwa UDP).

Fragestellungen der Netzneutralität befassen sich zumeist mit der Diskriminierung oder Privilegierung von Datenströmen, nicht einzelner Datenpakete (wenngleich das eine auf dem anderen aufbaut). Hierzu ist die Identifikation solcher zusammengehöriger Ströme notwendig. Eine auf TCP oder UDP basierenden Kommunikation zwischen zwei Computern wird durch das Viertupel $(A_s, A_d, P_s, P_d)$ bestehend aus den Adressen von Sender und Empfänger $(A_s, A_d)$ sowie Ziel- und Quellport $(P_s, P_d)$ zu einem bestimmten Zeitpunkt eindeutig identifiziert. Einige Protokolle, insbesondere solche aus dem Filesharing-Bereich verwenden deutlich mehr als nur eine einzelne Verbindung zum Datenaustausch.

Werden standardisierte Ports genutzt[11], lässt sich aus dem verbindungsidentifizierenden Viertupel auf das verwendete Protokoll – allerdings nicht auf konkrete Kommunikationsinhalte – schließen. Erfasst werden jedenfalls die Kommunikationsumstände.

Um Datenströme zu bestimmten Anwendungen oder Protokollen zuzuordnen, gibt es zwei grundlegende Herangehensweisen: Die Deep Packet Inspection (DPI) und die statistische Protokollidentifikation (SPID). Beide sollen im Folgenden kurz vorgestellt werden. Eine abgeschwächte Form der DPI, bei der nur der erste Header in den Nutzdaten des IP-Pakets untersucht wird, wird als Shallow Packet Inspection (SPI) bezeichnet.

### 2.5.1. Deep Packet Inspection

Bei der Deep Packet Inspection wird die Kenntnis über den Aufbau bekannter Protokolle genutzt, um diese sowie die zugehörigen Datenpakete zu identifizieren; eine Identifikation kann auch anhand übertragener Inhalte erfolgen. Die Betrachtung von Headerinformationen des ISO/OSI-Modell-Layers 4 ermöglicht oft eine schnelle Einordnung nach verwendetem Protokoll, z. B. anhand der verwendeten Portnummern (im Beispielpaket in Abb. 2.3 ab Bit $5 \times 32$ bzw. $5 \times 32 + 16$). Eine ausschließliche Verwendung solcher Headerinformationen wird als SPI bezeichnet. Eine Deep Packet Inspection betrifft das gesamte Datenpaket, es kann z. B. nach bestimmten Schlüsselworten durchsucht werden.

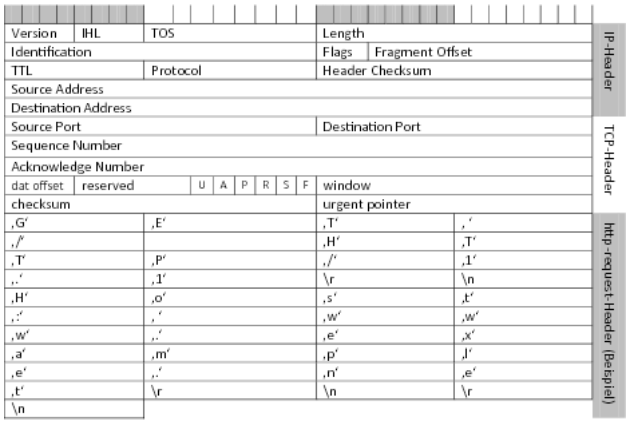


Abbildung 2.3.: Beispiel eines Datenpaketes einer HTTP-Anfrage

Aufgrund der hierzu notwendigen Kenntnis des verwendeten Protokolls gibt es natürliche Grenzen für den Einsatz von DPI: Unbekannte Protokolle können nicht analysiert werden. Hierbei ist bei „unbekannte Protokolle“ zu beachten, dass eine DPI nicht nur die

[11] also solche, für die eine Zuordnung zu einem Protokoll üblich ist wie z. B. Port 80 TCP zum Protokoll HTTP

für den Transport der Datenpakete verwendeten (und zu diesem Zweck vereinheitlichten) Protokolle verstehen muss, sondern auch die von der Anwendung selbst verwendeten Protokolle. Letztere sind im Gegensatz zu ersteren nicht notwendigerweise offen zugänglich spezifiziert[12]. Analog kann auch eine Suche nach Schlüsselworten nur dann erfolgreich sein, wenn die Schlüsselworte im Klartext übertragen werden – dies ist z. B. bei Ende-zu-Ende verschlüsselten Verbindungen nicht der Fall.

Eine genaue Beschreibung gegenwärtig verwendeter Techniken der Deep Packet Inspection findet sich bei Chaudhary [27], eine abstraktere Einführung in DPI bietet Lenka [28]. Forschung auf dem Gebiet der DPI befasst sich mit der Verwendung mächtiger Sprachen[13], um im Datenstrom zu suchende Ausdrücke zu beschreiben [28] und der performanten Änderung der entsprechenden Regelsätze [29].

*Betrachtet man die Einführungsbeispiele, so lassen sich beide zuordenen: Im Fall des schlechten Durchsatzes zu Servern des Skype-Dienstes könnten die Datenströme bereits anhand einer SPI identifiziert worden sein; im Fall der blockierten Downloads waren es die identifizierten Bytefolgen, die von einer Deep Packet Inspection als Schlüsselwort gesucht (und gefunden) wurden. Das Ergebnis dieser Suche wurde im Anschluss ausgewertet – und führte in diesem Fall zur Unterbrechung der Datenübertragungen.*

### 2.5.2. Statistical Protocol Identification

Bei der statistischen Protokollidentifikation werden sowohl Metadaten von Datenströmen wie etwa Paketgrößen und zeitliche Abfolgen wie auch Paketinhalte ohne Kenntnis des Aufbaus jenseits der Transportschicht ausgewertet. Abb. 2.4 illustriert eine statistische Protokollanalyse verschiedener Datenübertragungen anhand der zeitlichen Muster.

Hier sind symbolisch fünf Verbindungsmuster skizziert, bei denen lediglich zwischen Daten die zum Nutzer geschickt werden und Daten, die vom Nutzer geschickt sowie dem Volumen entsprechender Daten unterschieden wird:

1. Muster eines Anwenders, der sich Webseiten anzeigen lässt: Eine kurze Anfrage des Nutzers führt zu einer Antwort, die dann wiederum weitere Anfragen auslöst, die unterschiedlich große Antworten zur Folge haben (Nachladen von Graphiken o. ä.). Nach einer relativ langen Zeitspanne wird vom Nutzer die nächste Seite angefordert.

2. Muster eines Datentransfers, bei dem große Datenmengen vom Server zum Client transportiert werden. Der Client bestätigt dem Server durch kurze Acknowledgement-Nachrichten den Erhalt der Datenpakete.

[12] Beispielsweise ist das „Skype“-Protokoll nicht offengelegt und es ist keine durch Reverse-Engineering ermittelte Spezifikation verfügbar. In der Konsequenz bedeutet dies, dass eine DPI nicht in er Lage ist, „Skype“-Anrufe anhand der angerufenen Gegenstelle zu blockieren, sondern lediglich, weil aufgrund der SPI erkannt werden kann, dass die Gegenstelle eine „Skype“-Station ist – ausgenommen der Hersteller kooperiert und gewährt Zugriff auf Inhalte

[13] z. B. regulärer Ausdrücke

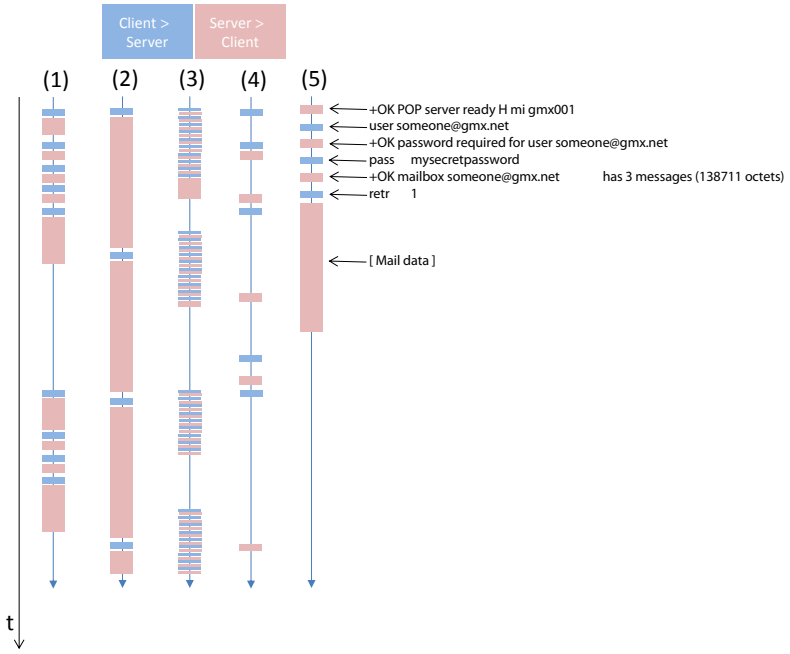


Abbildung 2.4.: Stilisierte Darstellung der Kommunikation zwischen Endnutzer und ISP als Beispiel für Muster verschiedener Datenübertragungen, die bei einer statistischen Protokollanalyse verwendet werden können – detaillierte Erläuterung der Beispieldatenflüsse (1)—(6) im Text.

3. Muster einer interaktiven Shell (z. B. Unix-Login): Viele kleine Datenpakete einzelner Tastenanschläge gefolgt von ebenso kleinen Datenpaketen, mit denen die Tastenanschlägen visualisiert werden (entferntes Echo); einzelne Tastendrücke bewirken größere Ausgaben (z. B. Absenden eines Befehls, der zu einer umfangreichen Ausgabe führt).

4. Muster ohne weitere Informationen nicht zuordnbar – denkbar wäre allerdings ein Chat per IRC, ICQ, Jabber oder dergleichen.

5. Muster einer POP3-Sitzung (in der Abb. zur Veranschaulichung um die jeweiligen Kommandos ergänzt).

Andere Methoden der statistischen Analyse betrachten die Verteilung relativer Häufigkeiten bestimmter Zeichen im Datenpaket – diese Methode lässt nicht nur Rückschlüsse zu, ob menschenlesbare Daten übertragen werden (in diesem Fall dominieren druckbare Zeichen), sondern auch (anhand der Häufigkeitsverteilung) eine grobe Einschätzung der verwendeten Sprache.

Eine grundlegende Einführung in die Technik der statistischen Protokollanalyse unter

Verwendung eines Bayes'schen Klassifikators bietet Ali [30], die Mächtigkeit der statistischen Protokollanalyse, auch mit verschlüsselten Inhalten umzugehen demonstriert Liu [31]; hier wird die Identifikation verschiedener genutzter Webanwendungen trotz ungebrochener Verschlüsselung des Übertragungskanals ermittelt. Die Präzision bei unverschlüsselten Datenübertragungen ist entsprechend höher, wie von Archibald gezeigt [32].

Gerade die Arbeit von Archibald verdeutlicht, dass es bei Kommunikation durchs Internet viele offene Fragestellungen aus dem Bereich des Datenschutzes gibt. Eine wichtige Frage ist hierbei die im nächsten Abschnitt mitbehandelte Frage der Personenbeziehbarkeit von Adressen behandelt. Denn während übertragene Daten potentiell durch den Nutzer vor Einblick geschützt werden können (z. B. durch Verschlüsselung), bleiben Adressdaten zwangsläufig unverschlüsselt.

## 2.6. Rechtliche Aspekte

Dieser Abschnitt gibt einige rechtliche Aspekte im Kontext von Netzwerken wieder. Begonnen wird mit einigen Begriffsdefinitionen nach Telekommunikationsgesetz (TKG); dem folgt ein kurzer Abriss der Gesetzesregelung zur Netzneutralität; schließlich wird die Personenbeziehbarkeit von Adressen erörtert.

### 2.6.1. Begriffsdefinitionen

Da einige Begriffe aus dem TKG regelmäßig in dieser Arbeit vorkommen werden, sollen sie hier kurz eingeführt werden.

**Telekommunikationsnetz** („Netzwerk“) als alle zur Signalübermittlung verwendeten aktiven und passiven Komponenten; im Sinne der Legaldefinition nach § 3(27) TKG ist ein

> „Telekommunikationsnetz“ die Gesamtheit von Übertragungssystemen und gegebenenfalls Vermittlungs- und Leitwegeinrichtungen sowie anderweitigen Ressourcen, einschließlich der nicht aktiven Netzbestandteile, die die Übertragung von Signalen über Kabel, Funk, optische und andere elektromagnetische Einrichtungen ermöglichen, einschließlich Satellitennetzen, festen, leitungs- und paketvermittelten Netzen, einschließlich des Internets, und mobilen terrestrischen Netzen, Stromleitungssystemen, soweit sie zur Signalübertragung genutzt werden, Netzen für Hör- und Fernsehfunk sowie Kabelfernsehnetzen, unabhängig von der Art der übertragenen Information.

Diese Arbeit beschäftigt sich ausschließlich mit öffentlichen Telekommunikationsnetzen, da eine Regulierung nichtöffentlicher Netze wenig sinnvoll erscheint, da diese nicht Teil des Internet sind.

**Öffentliches Telekommunikationsnetz** ist ein Telekommunikationsnetz, dass von der Öffentlichkeit genutzt werden kann; im Sinne der Legaldefinition nach § 3(16a) ist ein

> „Öffentliches Telekommunikationsnetz“ ein Telekommunikationsnetz, das ganz oder überwiegend der Bereitstellung öffentlich zugänglicher Telekommunikationsdienste dient, die die Übertragung von Informationen zwischen Netzabschlusspunkten ermöglichen.

Die Debatte um die Netzneutralität betrifft insbesondere die Betreiber von Netzwerken, sie erbringen Telekommunikationsdienste.

**Telekommunikationsdienste** sind die Übertragung von Daten gegen Geld; im Sinne der Legaldefinition nach § 3(24) sind

> „Telekommunikationsdienste“ in der Regel gegen Entgelt erbrachte Dienste, die ganz oder überwiegend in der Übertragung von Signalen über Telekommunikationsnetze bestehen, einschließlich Übertragungsdienste im Rundfunk.

Hiervon zu differenzieren sind die Inhalteanbieter; auch sie mögen Dienste gegen Geld (z. B. Pay-TV) oder ohne direkte Bezahlung (z. B. „YouTube“ als von „Google“ betriebene Videoplattform) erbringen.

### 2.6.2. Regelungen zur Netzneutralität

In der Bundesrepublik Deutschland existiert mit § 41a ein gesetzlicher Rahmen für eine Regelung der Netzneutralität durch eine Verordnung; allerdings ist bislang keine entsprechende Verordnung in Kraft. Ein im Sommer 2013 vorgelegter Entwurf des damaligen Bundeswirtschaftsministeriums wurde nicht weiter verfolgt. Entsprechend ist die Netzneutralität in der BRD weiterhin jenseits anderer Vorgaben (z. B. zum Wettbewerbsschutz oder zum Schutz freier Meinungsäußerung oder der Informationsfreiheit) unreguliert.

Die Kommentarlage zu dieser Regelung wird im folgenden Kapitel betrachtet; technische Details dieser Regelung werden im Kapitel 4 diskutiert.

Auf Gemeinschaftsebene existiert ein Entwurf für eine Verordnung des Europäischen Parlaments und des Rates zur Regelung der Netzneutralität[14] aus dem September 2013. Der Entwurf ist hinsichtlich der Ziele bzgl. der Netzneutralität mit der letzten TKG-Novelle vergleichbar. Auch hier wird nicht von einer absoluten Neutralität im Sinne einer Verpflichtung zum Einsatz von Best-Effort-Routing ausgegangen. Stattdessen wird von freiem Zugang zu Diensten und Anwendungen nach Wahl der Nutzer ausgegangen, der

[14] Dokument COM(2013) 627 final

durch einheitliche Vorschriften gewährleistet werden soll. Gleichzeitig sollen aber auch Spezialdienste („Managed Services“) möglich sein, die Anwendungen mit besserer Dienstequalität erlauben[15], wovon positive Auswirkungen auf Innovationen und die Erschließung neuer Märkte erhofft wird. Der Zugang zum im Entwurf als „offenes Internet“ bezeichneten Best-Effort-Bereich soll durch die Verordnung gewährleistet sein[16]. Dies wird in Art. 23 (1) konkretisiert, in der Endnutzern die Freiheit der Nutzung beliebiger Dienste oder Anwendungen über ihren Internetzugang zu verwenden zugesichert wird. In Art. 23 (2) findet sich die Verpflichtung der Netzbetreiber auf die Rückwirkungslosigkeit von (priorisierten) Spezialdiensten auf das „offene Internet“ – wenn auch nur eingeschränkt, denn hier sind nur dauerhafte oder sich wiederholende Einflüsse benannt. Kurzzeitige Auswirkungen von Spezialdiensten auf andere Anwendungen scheinen demnach tolerabel. In Art. 23(3) wird zunächst eine generelle Diskriminierungsfreiheit als Verbot von Blockaden und Verlangsamungen postuliert, ausgenommen sie sind angemessen. Als angemessen werden Eingriffe klassifiziert, die

a) eine rechtliche Grundlage besitzen oder ein Verbrechen abwehren oder verhindern,

b) zur Sicherstellung von Integrität und Sicherheit des Netzes (und der durch dieses erbrachten Dienste u. Anwendungen) notwendig sind,

c) nach Wunsch des Nutzers bestimmte Mitteilungen [...] unterbinden.

Diese Eingriffe haben transparent für den Nutzer zu erfolgen.

Art 24. schließlich trägt die „genaue Überwachung“ der Vorgaben des Art. 23 (1) und (5) den nationalen Regulierungsbehörden, im Fall der Bundesrepublik Deutschland der BNetzA, auf.

Nachdem in diesem Kapitel die technischen Grundlagen erläutert wurden, wird das nächste Kapitel diese Grundlagen zur Aufarbeitung des aktuellen Forschungsstandes zum Themenkomplex „Netzneutralität“ verwenden, auf dessen Basis dann die Transparenz von Netzwerken eingeführt wird.

[15] ebd., 2.3 Folgenabschätzung
[16] ebd., 3.4 Grundrechte

# 3. Forschungsstand zur Netzneutralität

> Network neutrality is a technical principle
> about the configuration of Internet routers.
> —*Timothy Lee [71]*

In diesem Kapitel werden die grundlegenden Aspekte des bestehenden Forschungsstandes zur Netzneutralität zusammengeführt. Die systematische Einordnung und Bewertung dieses Forschungsstandes findet sich als Eigenbeitrag dieser Arbeit im folgenden Kapitel 4.

In diesem Kapitel werden zunächst verschiedene Definitionen der Netzneutralität referiert; im Anschluss werden ausgewählte Nachweisverfahren beschrieben.

## 3.1. Definitionen von Netzneutralität

Je nach Betrachtungswinkel können sich unterschiedliche Definitionen ergeben; ein wesentlicher Unterschied ist auch der Grad der Abstraktion. In diesem Abschnitt werden Beiträge von Bullinger [12], Crowcroft [1], Dischinger [13], Felten [72], Gersorf [33, 20], Mengering [2], Schlauri [3], Tariq [14] und Wu [11] betrachtet.

Eine der ältesten Definitionen der Netzneutralität (network neutrality) stammt von Tim Wu aus dem Jahr 2003 [11]. Er definierte ein Netzwerk als neutral, wenn bei der Übertragung keine Anwendungen bevorzugt werden (Wu nennt als Beispiel das Browsen im World Wide Web etwa in Konkurrenz zu E-Mail).

Gersdorf beschrieb dies 2010 als Verpflichtung, „im Internet sämtliche Daten gleichberechtigt und unverändert zu übertragen“ [20]. Bullinger führt die Definition der Netzneutralität etwas technischer aus: „Alle Datenpakete werden gleichberechtigt übertragen, unabhängig davon, woher sie stammen, welchen Inhalt sie haben oder welche Anwendungen die Pakete generiert haben“ [12], so auch Mengering 2013 [2], Rn. 7.

Einige Definitionen heben auf die ausschließliche Verwendung des „Best Effort“-Ansatzes in einem neutralen Netzwerk ab (etwa [33]). Das bedeutet, dass jeder Bestandteil der Netzwerkinfrastruktur stets nur in Abhängigkeit von der Zieladresse des einzelnen Datenpakets die jeweils optimale Weiterleitung wählt und dass in der Folge die zur Verfügung stehende Netzkapazität unter den Nachfragern gleichmäßig verteilt wird [1, 3, 33]. Wird hiervon abgewichen und werden Datenpakete bestimmter Datenströme von einem

Netzbetreiber bevorzugt oder benachteiligt, so ist von einer Verletzung der Netzneutralität auszugehen [20].

Als besonders differenzierter und technikorientierter Definitionsvorschlag soll hier die von Crowcroft [1] vorgeschlagene Abkehr vom Begriff einer universellen Netzneutralität vorgestellt werden. Nach Crowcroft sollte Neutralität nur unter Bezugnahme auf eine konkrete Schicht der Internetarchitektur erfolgen. So schlägt folgende Begriffe vor:

- Verbindungsneutralität als Neutralität bezüglich der jeweiligen Endpunkte einer bi- oder multidirektionalen Kommunikation,
- Performanceneutralität als Neutralität der klar zu definierenden Leistungsfähigkeit einer Netzwerkverbindung (insb. im Fall eines Endnutzers und seines Internetzugangsanbieters) im Rahmen einer vertraglichen Vereinbarung,
- Diensteneutralität als Neutralität der Netzwerkanbindung hinsichtlich des genutzten Standes der Technik: Soweit dies technisch möglich ist, sollen auch neue Entwicklungen mit einer bestehenden Verbindung genutzt werden können. Als Beispiele benennt Crowcroft hierfür Multicasting oder mobiles Internet[1],
- Schichtenneutralität als Neutralität bezüglich der Nutzung beliebiger Dienste mit einem bestehenden Netzwerkanschluss.

Crowcroft selbst sieht seine Neutralitäten als platonische Illusionen an, die zwar erstrebt, jedoch nicht erreicht werden können.

Tariq et al. [14] definieren Netzneutralität als Eigenschaft des Internetservicebetreibers: Ein ISP sei dann neutral, wenn er Datenpakete unabhängig (neutral) von Inhalt, Anwendung oder Absender behandelt; Tariq beruft sich hierbei auf Felten [72].

Felten wiederum sieht drei Aspekte der Neutralität im Internet [72][2]:

**Netzneutralität als eine Ende-zu-Ende Designfrage** Netzwerke werden dafür bezahlt, Daten von einem Endpunkt zu einem anderen Endpunkt zu transportieren – Entscheidungen über Prioritäten oder Protokolle sind von den Endpunkten zu treffen.

**Netzneutralität als ein nichtexklusives Geschäftsprinzip** Neutralität ist ein ökonomisches Prinzip; Provider sollten Angebote [über Bevorzugung] nicht einzelnen Firmen, sondern wenn dann allen Firmen zugänglich machen.

**Netzneutralität als Nichtdiskriminierung von Inhalten** Diese Perspektive von Neutralität bezieht sich auf die von Nutzern Ausgetauschten Inhalte. Netzwerkbetreiber sollen eine „freie Rede“ nicht einschränken; sie sollen keine Entscheidung über die Transporteigenschaften einer Nachricht basierend auf deren Inhalt treffen.

[1] Ein Gegenbeispiel wäre eine Netzwerkverbindung, die auf den Transport von IPv4-Paketen beschränkt ist [Anm. d. Autors].

[2] sinngemäße Übersetzung und Zusammenfassung durch den Autor

Dischinger et al. schließlich betrachten Netzneutralität als die Frage, ob Netzwerkbetreiber Klassen von Datenübertragungen unterscheiden dürfen (um einzelne im Anschluss aus technischen oder ökonomischen Motiven zu drosseln) [13].

Eine systematische Einordnung und Bewertung der Aspekte dieser Definitionen findet als Beitrag dieser Arbeit in Kapitel 4 statt.

## 3.2. Verfahren zum Nachweis von Neutralitätsverstößen

Zusammen mit der Debatte über die Einhaltung oder Nichteinhaltung mehr oder weniger scharf umrissener Neutralitätskonzepte wurden diverse Ansätze zum Nachweis von Verstößen gegen ebendiese Neutralitätsvorstellungen entwickelt. In diesem Abschnitt werden die Ansätze Glasnost [34, 13], NANO [35, 14], Fathom [15], Herdict [16] und Nooter [17] vorgestellt. Eine systematische Einordnung findet sich ebenfalls im folgenden Kapitel.

### 3.2.1. Fathom

Bei Fathom [15] handelt es sich im Kern nicht um die Implementation eines Verfahrens zum Nachweis von Verletzungen der Netzneutralität, sondern ein Framework zur Untersuchung des Verhaltens von Webseiten im Browser „Mozilla Firefox“. Dennoch soll Fathom hier und in diesem Kapitel besprochen werden, da das Projekt einige interessante Ideen und Eigenschaften mitbringt. Unter anderem lässt sich mit dem Fathom-Framework eine Neutralitätsanalyse entwickeln. Hierauf wird in Kap. 5 wieder Bezug genommen werden.

Fathom ist der Grundidee nach ein im Browser integriertes Framework zur Analyse von Webapplikationen und deren Verhalten. Hierbei wurde das Framework allerdings so weit offen gehalten, dass auch eine detaillierte Analyse des Netzwerkverhaltens bei der Verwendung einer konkreten Anwendung möglich ist. Das Projekt wendet sich sowohl an Forscher, die der Gesamtkontext einer Anwendung vom Server bis hin zum Browser des Endnutzers interessiert, wie auch an Entwickler von Webanwendungen, die ein detailliertes Profiling und eine Identifikation von Engpässen wünschen.

Fathom wurde als ausschließlich auf JavaScript basierendes Browser-Plugin für den Webbrowser „Mozilla Firefox“ entwickelt und ist daher für alle Plattformen verfügbar, auf denen Mozilla Firefox verfügbar ist. Als Plugin klinkt sich Fathom in diverse Routinen und Aufrufe ein. Abb. 3.1 stellt ein Beispiel für die Beobachtung des Empfangens von Daten dar. Gezeigt ist die Einbindung von Fathom in ein auf der Webseite ablaufendes JavaScript. Aus der Webseite werden anstelle der durch NSPR bereitgestellten Aufrufe die von Fathom angeboten, so dass eine genaue Beobachtung des Seitenverhaltens möglich ist. Fathom wiederum leitet die Aufrufe in die NSPR-API weiter. Obwohl konzeptionell rein passive Messungen vorgenommen werden können (Fathom kann auch aktiv via JavaScript angesprochen werden), bewirken diese und der mit ihnen verbundene Overhead, bezogen auf de Aufruf einer Webseite Verzögerung im Bereich von Millisekunden.

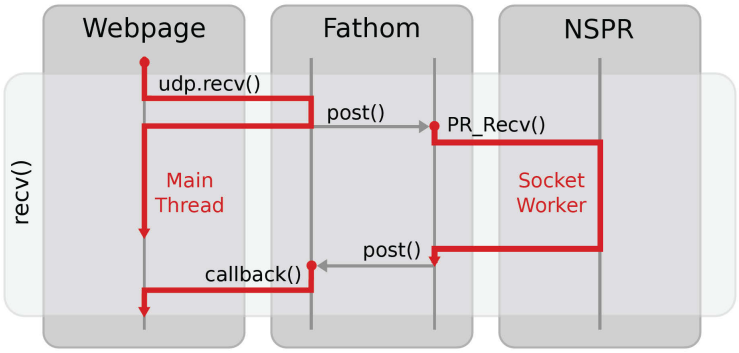


Abbildung 3.1.: Beispiel für einen Aufrufstack eines recv()-Calls, entnommen aus [15].

### 3.2.2. Glasnost

Das Programm Glasnost ([34], beschrieben in [13]) stellt sich dem Anwender in Form eines Java-Applets dar, welches im Browser ausgeführt wird und anschließend eine Bewertung der verwendeten Internetverbindung ausgibt. Zur Messung wird eine Reihe vordefinierter Server verwendet.

Die Grundidee von Glasnost ist der Austausch von Daten mit Testservern. Auf Grund der aus Java heraus eher schwierigen feingranularen clientseitigen Betrachtung[3] werden die Datenaustauschprozesse auf der Serverseite detailliert betrachtet. Auf der Suche nach Netzwerkpolicies, die zwischen verschiedenen Protokollen differenzieren, wird versucht, eine zu Grunde liegende Deep Packet Inspection nachzuweisen. Hierzu werden zwei unterschiedliche Formen von Testdaten zwischen Client und Server ausgetauscht, wie in Abb. 3.2 dargestellt.

Diese Daten sind jeweils in Größe und zeitlichem Ablauf des Handshakes zwischen Server und Client identisch; sie enthalten jedoch in einem Fall Zufallsdaten und im anderen Fall dem Protokoll entsprechende, valide Daten. Eine auf Deep Packet Inspection basierende Neutralitätsverletzung bezüglich des getesteten Protokolls liegt vor – so die Annahme – wenn sich diese beiden Datenströme hinsichtlich einzelner Qualitäten wie etwa Durchsatz, Latenz, Jitter oder Zuverlässigkeit statistisch signifikant unterscheiden.

Weiteres Designziel bei der Entwicklung von Glasnost war eine niedrige Barriere für auf dem Gebiet der Netzwerktechnik unerfahrene Nutzer. Hieraus motiviert sich auch die Entwicklung des Client-Programms als Java-Applet: Über einen Browser verfügt praktisch jeder Nutzer. Daneben wurde die Oberfläche möglichst einfach gestaltet und die Testdauer auf ein Maß herabgesetzt, dass sich in Studien als für Nutzer akzeptabel be-

[3] Insbesondere ein Java-Applet im Browser kann auf Betriebssystemfunktionen wie etwa Netzwerkverbindungen nur von einem sehr hohen – auf Unabhängigkeit vom unterliegenden Betriebssystem hin optimierten – Abstraktionsgrad zugreifen und daher z. B. keine genauen Beobachtungen über das Eintreffen einzelner Datenpakete anstellen

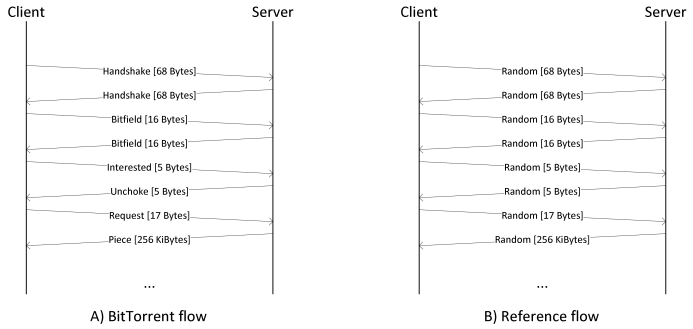


Abbildung 3.2.: Von Glasnost verwendete Testdatensätze unterscheiden sich zwar im übertragenen Inhalt, sind aber in Paketgröße und Timing identisch (Graphik nach [13]).

währt hat. Dies geschieht zwar zu Lasten der Messgenauigkeit – dafür wird die Messung von weniger Personen aus Zeitgründen wegen abfallenden Interesses abgebrochen.

In der Praxis wurde Glasnost von anderen Forschern in der Vergangenheit erfolgreich für Tests eingesetzt, die auch die Aufmerksamkeit der Populärmedien erreichte [73]; darüber hinaus wurde Glasnost 2013 in einem Projekt der Bundesnetzagentur zur Evaluation der Qualität deutscher Festnetzinternetzugänge verwendet [74].

### 3.2.3. Herdict

Herdict [16] ist eine vom Berkman Center for Internet and Society der Universität Harvard betriebene Webseite, die als Sammelstelle für Meldungen über von Nutzern beobachtete Einschränkungen konzipiert ist. Dabei wird komplett auf Crowdsourcing als Datenquelle gesetzt – es gibt keine technische Absicherung einzelner Problemmeldungen; eine Absicherung erfolgt allenfalls indirekt durch Häufung von Störungsmeldungen für einen bestimmten Kontext durch mehrere Nutzer. Jeder Benutzer kann Beobachtungen beitragen, es gibt keine Mindeststandards z. B. hinsichtlich technischer Detailliertheit. Meldungen werden mit den für den Server verfügbaren Daten über den Einsender einer Meldung korreliert. Ein Missbrauchsschutz gegen falsche Einträge ist nicht implementiert (Stand 2012).

### 3.2.4. Nano

NANO [14, 35] besteht aus zwei wesentlichen Komponenten: Einer dezentralen Monitoring- und einer zentralen Auswertungskomponente. Die Monitoringkomponente kann von Endanwendern auf Linux-Systemen installiert werden; die Auswertung findet auf einem

zentralen System des Georgia Institute of Technology statt. Nachdem die Installation abgeschlossen ist, muss der Nutzer NANO einige Rahmeninformationen über die betrachtete Verbindung bereitstellen: Etwa die Art der Internetverbindung (DSL oder Kabel?), die Art der physikalischen Anbindung des Rechners auf dem der NANO-Agent ausgeführt wird zum Router (Ethernet oder WLAN?) sowie Informationen zum abgeschlossenen Vertrag (1MBit oder 100MBit?). Schließlich wird der NANO-Agent als Monitoringdienst ausgeführt und überwacht sämtliche ein- und ausgehenden Datenpakete. Eine Teilmenge der hierbei gewonnenen Informationen wird zur zentralen Auswertung weitergeleitet. Der Nutzer kann den Messagenten für eine gewisse Zeit deaktivieren, um „unbeobachtet“ das Internet zu benutzen.

NANO ist gegenwärtig nur für Linux verfügbar, daher bleibt seine Anwendung auf einen relativ kleinen Kreis interessierter Nutzer beschränkt. Eine zentrale Auswertung ist erforderlich; dies bedeutet, dass Informationen über die von einem Nutzer kontaktierten Gegenstellen weitergegeben werden müssen. Hierauf wurde von den Entwicklern mit der Möglichkeit des temporären Deaktivierens des NANO-Agenten reagiert.

### 3.2.5. Shaperprobe

Das Programm Shaperprobe [36] besteht aus einem Client, der vom Anwender ausgeführt werden kann und zugehörigen Servern.

Die grundlegende Idee des Projekts Shaperprobe ist die Erkennung eines bestimmtes Verhaltens von Shapingalgorithmen, konkret eines Einschwingverhaltens, wie in Abb. 3.3 skizziert. Dieses führt zu einem beobachtbaren „Burst“ zu Beginn der Datenübertragung. Anschließend tritt die Durchsatzbegrenzung in Aktion und (um den Soll-Wert einzuhalten) wird der Datenstrom deutlich nachlassen – hieraus ergibt sich die abfallende Flanke nach dem anfänglichen Burst.

Die konkrete Umsetzung von Shaperprobe verwendet aus UDP-Paketen bestehende Datenströme, die mit dem Testserver ausgetauscht werden.

### 3.2.6. Nooter

Ein für diese Arbeit grundlegender Ansatz wurde von Dan Kaminsky unter dem Namen „Nooter“ vorgestellt [17]. Da die in Abschnitt 5 vorgestellte Methode auf Nooter basiert, soll dieser Ansatz hier ausführlich besprochen werden.

Ein wesentlicher Grundgedanke der Konzeption von Nooter ist, dass Provider grundsätzlich nicht zur Neutralität gezwungen werden sollen; vielmehr wird durch Nooter dafür gesorgt, dass jeder nichtneutrale Eingriff vom Nutzer bemerkt werden kann[4]. Darüber

[4] „Get transparent or bust!” [17]

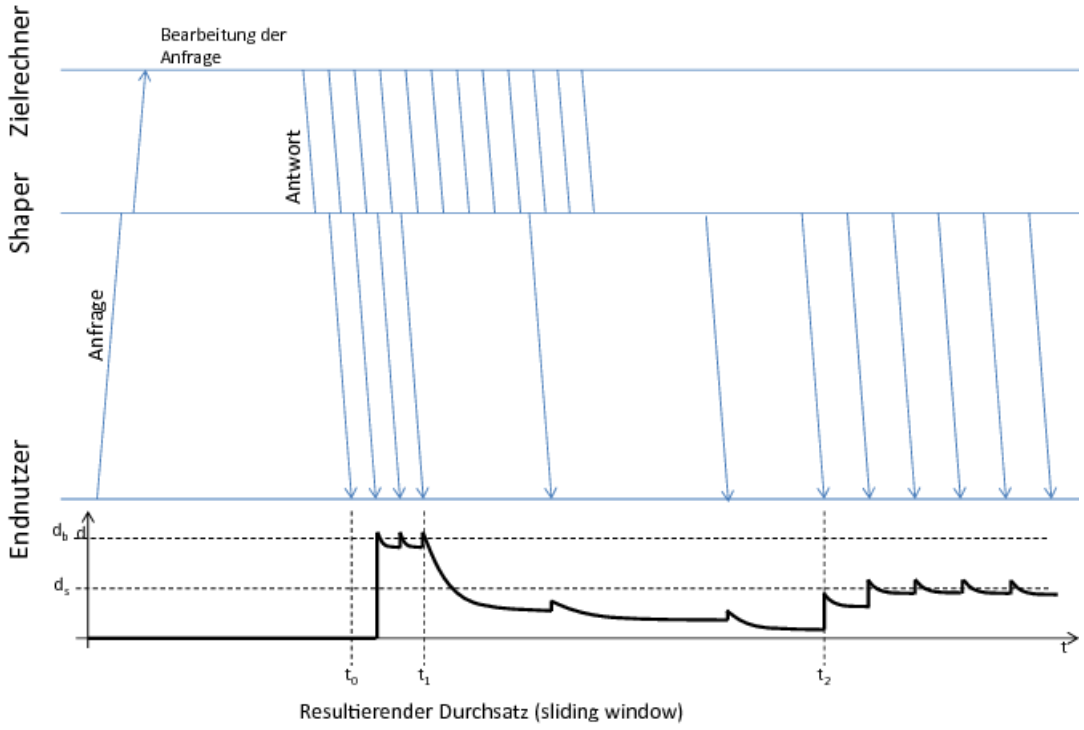


Abbildung 3.3.: Nachweis der Anwesenheit eines Trafficshapers anhand von Latenzänderungen: Oben der Datenaustausch zwischen Endnutzer und Zielsystem z. B. mit UDP-Paketen; unten angedeutet der sich aus der Position des Endnutzers ergebende Durchsatz (bei Berechnung mit einem Sliding-Window-Algorithmus). Der Datenstrom vom Server zum Nutzer lässt sich in drei Phasen einteilen: 1. Burst (von $t_0$ bis $t_1$), hier wird der ungedrosselte Burst-Durchsatz von $d_b$ erreicht; 2. Delayphase (von $t_1$ bis $t_2$), der Durchsatz wird – gerechnet auf die gesamte bisherige Übertragung – dem Sollwert $d_s$ durch Verzögerung einzelner Datenpakete angenähert; 3. ab $t_2$ werden die Datenpakete so versandt, dass sich der vom Shaper sicherzustellende Durchsatz $d_s$ ergibt.

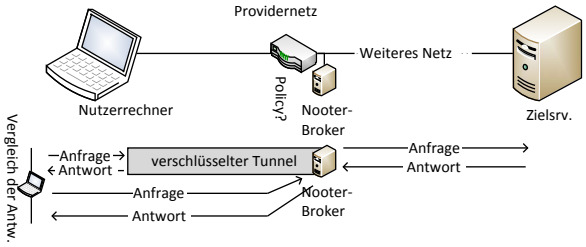


Abbildung 3.4.: Aufbau für eine Messung mit dem hybriden Nooter-Ansatz: Es wird ein verschlüsselter Tunnel zu einem Broker aufgebaut. Dieser Broker fungiert als Proxy und sendet die Antwort des tatsächlichen Ziels gleichzeitig durch den Tunnel und durch das Betreibernetz zum Nutzer. Dieser kann nun zwischen beiden eingehenden Datenströmen vergleichend messen.

hinaus ist Nooter so konzipiert, dass der Provider die Verwendung von Nooter zwar bemerken kann, dies ihn jedoch nur vor die Wahl stellt, dass nichtneutrale Policies detektiert werden oder aber er für Verbindungen dieses Nutzers auf ein neutrales Policy-Set wechseln muss. In jedem Fall gewinnt der Endanwender: Entweder in der Form von weiteren Informationen über das (nichtneutrale) Verhalten seines Providers – oder aber durch einen Policy-Wechsel des Providers hin zu einer öffentlich vertretbaren (und meist anwenderfreundlicheren) Policy.

Die grundsätzliche Vorgehensweise von Nooter wird anhand von Abb. 3.4 erläutert. Die Abbildung zeigt im oberen Teil vereinfacht die Topologie zwischen einem Endnutzer und einem (beliebigen) Zielrechner. Dieser bietet einen beliebigen Dienst an. Hierbei wird zunächst das Providernetz passiert, dessen Netzwerkpolicies untersucht werden sollen. Im weiteren Verlauf des Routingpfades gibt es einen sog. Nooter-Broker. Topologisch ist zunächst wichtig, dass sich dieser Router außerhalb des Einflusses des Betreibers der zu untersuchenden Netzwerke befindet. Daneben benötigt der Broker die Fähigkeit, beliebige Datenpakete an den Endnutzer senden zu können, insbesondere Datenpakete mit beliebigen Absenderadressen.

Im unteren Teil der Abbildung sind die während einer Messung aufgebauten tatsächlichen und scheinbaren Verbindungen skizziert.

Um das Netzwerk des Providers zu untersuchen, wird eine verschlüsselte Verbindung zwischen Nutzer und Nooter-Broker aufgebaut. Je nach „Ausbaustufe“ von Nooter kann nun eines der folgenden Vorgehen gewählt werden:

1. Im einfachsten Szenario wird die abzusetzende Anfrage (an einen beliebigen Server im Internet) an den Nooter-Broker übermittelt. Diese stellt daraufhin eine Verbindung zum eigentlichen Zielserver hier und übermittelt die Antwortpakete an den Anfrager. Dies allerdings doppelt: Einmal durch die verschlüsselte Verbindung (also für den Provider unsichtbar) und einmal durch das zu untersuchende Netzwerk (also potentiell für den Provider sichtbar). Hierbei wird vom Nooter-Broker als Ab-

senderadresse der Datenpakete die Adresse des eigentlichen Zielservers eingetragen. Für eine eventuelle DPI des Zugangsproviders sind diese Pakete nicht von Paketen zu unterscheiden, die tatsächlich vom Zielserver stammen. Von der Position des Endnutzers aus kann nun das Verhalten der unverschlüsselt übertragenen mit dem Verhalten der durch den Tunnel übertragenen Datenpakete verglichen werden und nach unterschiedlichem Verhalten in Abhängigkeit von Ziel, Protokoll oder Inhalt gesucht werden. Vorteil einer solchen Implementation ist eine relative Einfachheit.

2. Optional können auch die Anfragepakete vom Endnutzerrechner aus zusätzlich abgeschickt werden; hierbei muss sichergestellt werden, dass diese Pakete den eigentlichen Zielrechner nicht erreichen. Kaminsky schlägt das bewusste Einsetzen zu niedriger TTLs oder ungültiger Prüfsummen[5] vor. Im Falle einer zu geringen TTL würde das Paket von einem Router auf dem Pfad zum eigentlichen Ziel verworfen (und der Endnutzer hierüber durch ein ICMP-Paket informiert werden); im Fall einer falschen Prüfsumme würde (in einem absolut neutralen Netz) der Zielrechner das Paket verwerfen; in einem realen Netz ist vorstellbar, dass ein Paket mit einer ungültigen Prüfsumme bereits vorher durch einen Router verworfen würde. Ziel wäre in beiden Fällen, dass im zu untersuchenden Provider-Netzwerk der Eindruck eines tatsächlichen bidirektionalen Datenaustauschs erweckt wird. Hierzu müssen ferner für Protokolle wie TCP die zu verwendenden Sequenz- und Bestätigungsnummern koordiniert werden.

Das Verfahren kann in Abhängigkeit von einer konkreten Implementation unsichtbar gegenüber Protokollen ab Ebene 4 des ISO-/OSI-Modells sein, da die Paketweiterleitung ebenso wie die Messung der Netzwerkperformance auf IP-Basis geschieht. In der Praxis scheint dies einige Probleme zu provozieren, die nun betrachtet werden sollen.

Bei verbindungslosem Datenaustausch fällt zunächst auf, dass aus Sicht des ISP der Datenaustausch nur vom entfernten Server zum Endnutzer stattfindet – ohne ein einziges Anfrage- oder Antwortpaket. Dies könnte ein IDS oder andere Sicherungseinrichtungen zu Aktivität veranlassen, weil es ein stark von der Normalität abweichendes Verhalten wäre. Vorteil für den Endnutzer: Wenn es zu einem Eingriff kommt, gibt es einen Wissensgewinn über das Verhalten des Netzwerks.

Wird in dieser Situation Variante 2 verwendet, so kann es dazu kommen, dass vom Endnutzer ausgehende Pakete trotz Gegenvorkehrungen beim eigentlichen Zielserver eintreffen[6]. Diese können natürlich ebenfalls zu einer Reaktion führen – schlimmstenfalls jedoch zu einer Doppelung des Datenverkehrs, wovon die eigentlichen Messpakete nicht betroffen sind. Gegebenenfalls muss auf der Endanwenderseite im Rahmen der Messung

[5] Die IP-Header-Prüfsumme wird von jedem Router geprüft; Pakete mit falscher IP-Header-Prüfsumme werden vom nächsten Router verworfen. Datenpakete mit falscher TCP- (oder UDP-)Prüfsumme werden hingegen bis zum Zielrechner transportiert und erst dort verworfen.

[6] Z. B., weil falsche Checksums von Routern korrigiert werden (dies führte zu weiterer Kenntnis über die vom Netzbetreiber eingesetzte Technik) oder eine eigentlich als ausreichend niedrig angenommene TTL nicht niedrig genug war.

ein Filter verwendet werden, um doppelte Pakete zu entfernen.

Wird bei Testvariante 1 als Layer-4-Protokoll TCP eingesetzt, verschärfen sich die geschilderten Probleme: TCP ist etabliert und das Vorhandensein statusbehafteter Firewalls im Provider-Netz erscheint nicht abwegig. Eine solche Firewall würde jedoch keine TCP-Pakete einer – aus Sicht der Firewall – nicht ordnungsgemäß eröffneten Verbindung weiterleiten. Auch wenn so zunächst keine Messungen möglich sind, dennoch steht am Ende ein Erkenntnisgewinn über eingesetzte Netzwerkinfrastruktur.

Leicht anders stellt sich die Situation dar, wenn Variante 2 mit TCP verwendet wird. Hier wäre für das ISP-Netzwerk eine vollwertige TCP-Verbindung sichtbar; zu Problemen käme es jedoch, wenn aus o.g. Gründen eigentlich nur für das untersuchte Netzwerk generierte Anfragepakete den entfernten Server erreichen. Entweder erreichen diese den Server von Beginn, dann kommt es zum Beginn eines Dreiwegehadshakes. Der Endnutzerrechner wird jedoch nicht auf das Verbindungsangebot des entfernten Servers sondern das des Nooter-Brokers eingehen[7]. Im weiteren Verlauf kann es dann dazu kommen, dass der entfernte Server Datenpakete erhält, die er keiner aktiven Verbindung zuordnen kann – als Reaktion wird er ggf. Pakete mit gesetztem `RST`-Flag senden, um die scheinbar noch halboffene Verbindung zu terminieren. Auch diese Pakete könnten herausgefiltert werden, würden jedoch die Provider-Infrastruktur passieren und könnten dabei Reaktionen hervorrufen. Alternativ erreichen die Pakete den Server nicht, womit dieses Szenario entfällt.

Reale Verbindungen existieren während einer Messung

- Zwischen Endnutzerrechner und Nooter-Broker als verschlüsselter Tunnel,
- Zwischen dem Nooter-Broker und dem tatsächlichen entfernten Testserver.

Ferner wird einem Beobachter im untersuchten Netz der Eindruck erweckt, dass eine weitere Verbindung zwischen dem entfernten Testserver und dem Endnutzerrechner existieren würde. Die hierzu ausgetauschten Datenpakete dienen jedoch ausschließlich zur Messung der Verbindungsqualität.

Eine Kenntnisnahme des Providers von einer verschlüsselten Verbindung zu einem Nooter-Broker ist dem Verfahren zwar als Messverfahren abträglich (denn diese Kenntnis könnte das untersuchte Verhalten verändern) – es ist jedoch im Sinne der Ausrichtung von Nooter kein Problem. Denn ein beobachteter Provider wird kein Interesse haben, einem testenden Nutzer Indizien für eine Neutralitätsverletzung an die Hand zu geben. Entsprechend stünde ein Netzwerkbetreiber vor der Wahl, entweder ein neutrales Policy-Set auf diesen einen Kunden anzuwenden – oder aber den Nachweis nichtneutralen Handelns zu riskieren. In beiden Fällen gewinnt der Nutzer.

[7] Welches eine andere Sequenznummer verwenden wird; Ausnahme: Es kommt hier zu einer Kollision.

Aus jedem Verhalten des Netzwerks sollen sich– im Sinne der Nooter-Idee – Schlüsse über das Verhalten des Providernetzwerks und eventuelle Traffic-Engineering-Maßnahmen ziehen lassen.

Kaminsky berichtet [17] von einem erfolgreichen Einsatz mit einem ISP; eine öffentlich verfügbare Version von Nooter gibt es allerdings bislang nicht.

## 3.3. Rechtliche Aspekte

In diesem Abschnitt wird der Literaturstand zu zwei für diese Arbeit wesentlichen Aspekten von Netzwerken wiedergegeben: Zunächst zur Personenbeziehbarkeit von Adressen und im Anschluss zum § 41a Telekommunikationsgesetz (TKG).

### 3.3.1. Personenbeziehbarkeit von Adressen

Eine wesentliche Fragestellung im Kontext von Internet und Datenschutz ist die Frage der Personenbeziehbarkeit von Adressen. Ein Datum gilt nach § 3(1) Bundesdatenschutzgesetz (BDSG) als personenbezogen, wenn es sich um „Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person“ handelt. Aus Routing ergibt sich, dass potentiell jeder Router alle ihn durchquerenden Datenpakete mitlesen kann; mithin der Router-Betreiber mittels Beobachtung der Datenpakete ein akkurates Bild vom Nutzerverhalten[8] erhält, beispielsweise welche Webseiten von einer bestimmten Adresse aus genutzt wurden. Datenschutzrechtlich relevant wird es immer dann, wenn sich einer solchen Adresse (einem Profil besuchter Seiten) ein Name zuordnen lässt. Mit der Entwicklung der „sozialen Online-Medien“ (Facebook, Twitter, ...) und den von ihnen genutzten Like- und Share-Buttons fällt ein entsprechendes Profil noch an anderer Stelle an: Da diese Buttons stets von den Seiten der Betreiber der sozialen Medien geladen werden (vgl. Abb. 3.5) und bei diesen Abrufen auch stets in der Anfrage enthalten ist, über welche Seite diese Anfrage erfolgt[9], entstehen auch hier entsprechende Nutzungsprofile, da der Betreiber des eingebundenen Inhalts einen Datensatz erhält, welcher Inhalt von einem Client mit welcher IP gerade über welche Seite eingebunden wurde. Dies ist allerdings nur der letzte Schritt einer längeren Entwicklung; zuvor wurden solche Tracking-Services über unsichtbare 1x1-Pixel-Graphiken realisiert, die von Webseitenbetreibern eingebettet werden konnten. Im Gegenzug erfuhren diese von den Betreibern der Tracking-Dienste mehr über ihre Besucher (Überblick und rechtliche Bewertung vergleichbarer Techniken durch Alich/Voigt in [37]).

Dass die geräteeindeutige Hardwareadresse eines Endgeräts (insbesondere die Hardwareadresse eines WLAN-fähigen Smartphones) hierzu gehört, wird angesichts des Forschungsstands wohl kaum zu verneinen sein: Die Arbeiten von Tan et al. [38, 39] zeigen eindrücklich die Möglichkeiten, einzelne Personen anhand der Protokolle einzelner Access-Points zu vereinzeln (bzw. zu identifizieren). In diesem Kontext genügt bereits

[8] mit der IP-Adresse als Pseudonym einzelner Nutzer oder Nutzergruppen

[9] sog. „Referrer“ des HTTP-Requests

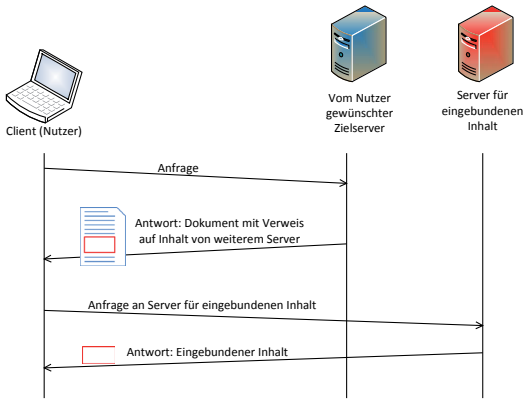


Abbildung 3.5.: Illustration des Anfallens von Nutzungsprofilen bei den Betreibern sozialer Medien: Attraktive Inhalte (z. B. „Like-Button“) werden als einfach einzubindende externe Seiteninhalte bereitgestellt. Ein Nutzer ruft eine Internetseite ab (blau). Diese beinhaltet den Verweis auf ein entsprechendes Fragment (rot), welches von einem Server des Betreibers abgerufen wird. Durch den Abruf erhält der Betreiber des roten Servers Kenntnis darüber, welche Seiten vom Nutzer besucht wurden. Das Vorgehen ist nicht auf soziale Netze beschränkt (auch ist es deutlich älter – klassischerweise wurden zum Nutzer-Tracking unsichtbare Ein-Pixel-Grafiken verwendet).

die Anwesenheit mit einem aktiven Smartphone, um einem Beobachter das Verfolgen einer Person zu erlauben. Tan et al zeigen auf, dass auch nach Entfernung der eigentlichen Hardwareadressen in einem Log von WLAN-Zugangssystemen ausreichend Datenspuren verbleiben, um Nutzer eindeutig zu identifizieren. Ghosh et al [40] zeigen ferner im Rahmen eines großangelegten Experiments an der ETH Zürich die Möglichkeiten der Bildung von Nutzerprofilen. Derartige Experimente zeigen die Möglichkeit, Bewegungsprofile anzulegen, aus denen Schlüsse über das soziale Verhalten der Besitzer bestimmter Adressen (mithin der Nutzer) abgeleitet werden können[10].

Die Personenbeziehbarkeit der IP-Adresse wird in der Rechtswissenschaft teilweise noch kontrovers diskutiert; Krüger und Maucher [41] geben hier einen Überblick über die vorhandenen Literaturpositionen. Kern der Frage ist stets die Personenbeziehbarkeit; die Sammlung pseudonymer Profile (die stets nur der abstrakten IP-Adresse zugeordnet sind) durch Betreiber von Webseiten wird als unproblematisch oder sogar als technisch notwendig betrachtet. Eine Zuordnung von IP-Adressen zu Personen ist den Zugangsanbietern möglich. Gegner einer Personenbeziehbarkeit der IP-Adresse argumentieren hierauf basierend, dass diese Zuordnung Dritten in Ermangelung der Daten des Providers nicht möglich ist[11]. Eine Personenbeziehbarkeit sei nur gegeben, wenn dem Dritten auch die Mittel zur Herstellung eines Personenbezugs zur Verfügung stünden[12]. Von Befürwortern einer Personenbeziehbarkeit wird diese als eine sich aus der Eigenschaft der IP-Adresse als Einzelangabe über eine bestimmbare natürliche Person[13] ergebend angenommen. Ebenfalls wird als Argumentation angenommen, dass bereits die Möglichkeit der Zuordnung (durch einen Dritten, hier den Zugangsprovider) ausreicht, um den Personenbezug zu begründen.

Im Ergebnis verneinen Krüger und Maucher eine Personenbeziehbarkeit der IP-Adresse. Zur Begründung verweisen Sie darauf, dass die notwendigen Zuordnungsdaten nur den Providern vorliegen und dass es in deren wirtschaftlichen Interessen und rechtlichen Verpflichtungen läge, diese Zuordnung Dritten nicht zugänglich zu machen.

### 3.3.2. Rezeption des § 41a TKG

Mit der letzten Novellierung des TKG wurde 2012 mit § 41a erstmals eine gesetzliche Regelung der Netzneutralität in der Bundesrepublik angestrebt. Allerdings wird an dieser Stelle keine Regelung im eigentlichen Sinne getroffen; Adressat ist in Absatz (1) die Regierung, die ermächtigt wird, eine Verordnung zur konkreten Regelung der Netzneutralität (mit Zustimmung des Bundesrates) zu erlassen, deren Adressat dann Betreiber

[10] Auf einem Universitätscampus lassen sich anhand derartiger Profile beispielsweise leicht Mitarbeiter von Studenten (Anwesenheitszeiten) und Studenten untereinander nach Studiengang (Anwesenheit in bestimmten Gebäudeteilen zu bestimmten Zeiten außerhalb der vorlesungsfreien Zeiten) ableiten.
[11] AG München vom 30.09.2008; MMR 2008, 860
[12] LG Wuppertal vom 19.10.2010; MMR 2011, 66 – ebenso OLG Hamburg vom 03.11.2010; MMR 2011, 281
[13] AG Berlin-Mitte vom 27.03.2007 und LG Berlin vom 06.09.2007; K&R 2007, 600

von Telekommunikationsnetzen wären.

Bereits dieses Vorgehen wird formal hinsichtlich der europarechtlichen Vorgaben angezweifelt; konsequent (und konsistent mit §§ 9ff. TKG) erschiene eine ausschließliche Ermächtigung der BNetzA als Regulierungsbehörde[14]. Um diesen Vorgaben weiterhin zu entsprechen, seien die faktischen Schwerpunkte einer Sicherung der Netzneutralität (auch durch eine entsprechende Verordnung) bei der BNetzA zu erwarten. In Absatz (2) wird die BNetzA zur Herausgabe einer technischen Richtlinie ermächtigt, in der von ihr technische Mindeststandards für Internetverbindungen deklariert werden können – dies allerdings nur in enger Kooperation mit den europäischen Gremien.

Hervorgehoben wird von Kommentatoren[15] die Formulierung in § 41a(1) als reine Ermächtigung der Bundesregierung. Hieraus ergäbe sich ein weitgehender Ermessensspielraum, da die Entscheidung über ein Eingreifen (durch Erlass einer Verordnung) faktisch einzig dem Ermessen der Regierung überlassen bliebe. Seine Grenzen habe dieser Rahmen durch § 41a(1) 2. Satz – hier werden die in § 2 normierten Regulierungsziele des TKG referenziert, insbesondere die „Möglichkeit der Endnutzer, Informationen abzurufen und zu verbreiten oder Anwendungen und Dienste ihrer Wahl zu nutzen“.

## 3.4. Zusammenfassung

In diesem Kapitel wurden zunächst Definitionen für Netzneutralität zusammengetragen, die teilweise unterschiedliche Kriterien berücksichtigen. Im Anschluss wurden unterschiedliche Verfahren zum Nachweis von Verstößen gegen die Netzneutralität dargestellt. Diese sind in Tabelle 3.1 noch einmal zusammengefasst. Ausgenommen ist hierbei Fathom, da es sich bei Fathom um kein konkretes Messverfahren, sondern lediglich eine Plattform handelt, auf deren Basis ein Nachweisverfahren implementiert werden kann.

Im folgenden Kapitel werden die Ergebnisse dieses Kapitels kritisch reflektiert und systematisch eingeordnet.

[14] [33], Rn 12 bezugnehmend auf Art 3. Abs 3a RRL
[15] [33], Rn. 13; [2], Rn. 25

| | Glasnost | Herdict | NANO | Nooter | Shaperp. |
|---|---|---|---|---|---|
| Unterstützung durch Netzbetreiber | Nein | Nein | Nein | Ja | Nein |
| Verwendet zugehörige Testserver | Ja | Nein | Nein | Nein | Ja |
| Zentrale Auswertung | Nein | Nein | Ja | Nein | Nein |
| Statistische Absicherung d. Aussage | Ja | Nein | Ja | Ja | Ja |
| Betriebssystem/Plattform | unabh./Java | unabh./Web | Linux | k.A. | unabh./f. mehrere BS verfügbar |
| Messverfahren | Differenzmessung Daten/Rauschen | Beobachtung d. Nutzers | Monitoring aller Verb. | Differenzmessung Tunnel/Plain | Test auf Shap.-Alg. |

Tabelle 3.1.: Zusammenfassung der betrachteten Verfahren zum Test auf Neutralitätsverstöße.

# 4. Analyse des Forschungsstands und weiterführender Fragestellungen

Fanatiker wollen mit dem Kopf durch das Brett, das sie vor ihm haben.
—*Wolfgang Weidner*

In diesem Kapitel wird der im letzten Kapitel zusammengetragene Forschungsstand systematisch bewertet. Zunächst wird auf die Schwierigkeiten eingegangen, die sich in Verbindung mit der Suche nach einer Definition der Netzneutralität ergeben; im darauffolgenden Abschnitt werden die Testverfahren kategorisiert und hinsichtlich ihrer Stärken und Schwächen beurteilt.

Der nach diesem Kapitel bewertete Forschungsstand zur Netzneutralität wird die Grundlage der hierauf aufbauende Konstruktion des Begriffs der „Transparenz von Netzwerken“ im Folgekapitel bilden. Die Übersicht und Bewertung der technischen Nachweisverfahren ist eine erweiterte Fassung des im Rahmen dieser Dissertation als [42] publizierten Surveys.

*Handelte es sich bei den Einführungsbeispielen um Neutralitätsverletzungen? Im Beispiel 1 wurden Verbindungen hinsichtlich ihres Durchsatzes beschränkt, weil sie sie scheinbar einen bestimmten Server als Gegenstelle hatten. Nicht zuletzt die Auflösung durch eine Zahlung scheint eine Neutralitätsverletzung nahezulegen.*

*Das zweite Beispiel ist für diese Diskussion ergiebiger, weil weniger eindeutig zuordenbar. Als Martin das Phänomen einer nicht herunterladbaren Datei das erste Mal beobachtete, stand auch ein technisches Problem des ausliefernden Servers im Raum – denkbar wäre, dass der Server aufgrund eines Festplattenschadens in einem bestimmten Sektor nicht in der Lage war, die entsprechende Datei selbst komplett auszuliefern. Der erfolgreiche Download von anderem Anschluss als der Universität Rostock ließ uns diese Hypothese verwerfen. Wenn der Server die Datei reproduzierbar durch Routingpfad A ausliefern kann, nicht aber durch Pfad B, dann kann die Ursache im Routingpfad B liegen. Ausgenommen, es gibt Serverregeln, die Pfad B diskriminieren. Diese konnten spätestens mit der Platzierung der Testdatei auf einem kontrollierten Testserver ausgeschlossen werden – und es blieb nur noch der beiden Servern gemeinsame Routingpfad B' als Ursache. Damit scheint das Urteil eindeutig: Neutralitätsverletzung, denn einige Router haben bestimmte Datenpakete nach enthaltenem Inhalt (oder Inhalt* ***vorangegangener*** *Pakete) anders behandelt als andere.*

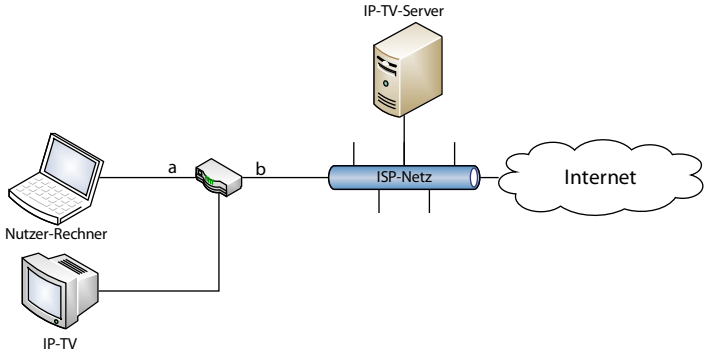


Abbildung 4.1.: Skizze eines typischen Anwender-Internetanschlusses mit zwei möglichen Bezugspunkten (a, b) für eine Untersuchung der Neutralität.

## 4.1. Der Begriff der „Netzneutralität“

Wie im Abschnitt 3.1 beschrieben, gibt es eine Vielfalt an Definitionen. Eine zentrale Rolle spielt – ausgesprochen oder unausgesprochen – bei den meisten Definitionen die Verwendung des Best-Effort-Ansatzes, also eine ausschließliche Abhängigkeit der Paketbehandlung vom jeweiligen Zielrechner, wobei diese Behandlung sich nur auf die Weiterleitung des Paketes und nicht auf eine mögliche Priorisierung hierbei bezieht.

Die Vielfalt an Definitionen ist offensichtlich zwei Tatsachen geschuldet: Netzneutralität ist sowohl ein multidisziplinäres und umstrittenes wie auch ein recht junges Phänomen. Ersteres ist leicht verständlich, betrachtet man die unterschiedlichen Interessen, die nach der Entwicklung des Internets in den letzten zwanzig Jahren hierauf einwirken – das anfängliche Interesse an einem reibungslosen Betrieb ist mittlerweile nur noch eines neben massiven ökonomischen und (geographisch unterschiedlich stark ausgeprägten) rechtlichen und politischen Interessen.

Insbesondere die Teildefinitionen von Crowcroft verdeutlichen, dass das Verständnis von Netzneutralität wesentlich vom Bezugspunkt abhängig ist. Zur Illustration sei auf die in Abb. 4.1 skizzierte Topologie eines Internetzugangs verwiesen, mit dem gleichzeitig ein IP-TV-Paket gebucht sei. In diesem Fall werden die Daten des IP-TV durch QoS-Mechanismen höher priorisiert als die Daten der Internetnutzung. Bei Untersuchung der Fragestellung, ob es sich um einen neutralen Internetzugangspunkt nach Definition von Bullinger handelt, hängt die Antwort ganz wesentlich davon ab, an welchem Punkt der Verbindung die Untersuchung stattfindet.

Würde die Neutralität des Netzwerk an der Position `(a)` untersucht, so könnte es sich um ein neutrales Netz handeln, in dem alle Datenpakete (vom PC des Nutzers ins Internet) gleich behandelt werden (wenn der Provider neutral agiert). Wird das Netz

hingegen an der Position (b)[1] untersucht, so würde eine Priorisierung von Datenpaketen des IP-TV-Datenstroms (was den Kriterien Inhalt, Protokoll oder Ziel bzw. Absender zugeordnet werden könnte) nachweisbar sein. Dieses Beispiel ist besonders deshalb wichtig, weil ein Endkunde bei der Durchführung von Messungen typischerweise nur eine Messung an der Position (a) durchführen können wird.

Eine Argumentation, dass Netzneutralität nur für Diensteanbieter außerhalb des Netzes des Zugangsproviders gelten solle, ist in zweierlei Hinsicht widersprüchlich. Einerseits ist ein solches Kriterium z. B. in der Definition nach Bullinger weder wörtlich noch nach Sinn und Zweck enthalten. Daneben führt eine solche Definition schnell ad absurdum, weil sie eine Privilegierung der vom ISP mit weiteren Unternehmenszweigen betriebene Dienste (z. B. ein Videoportal) nicht mehr als Neutralitätsbruch erfassen würde, wenn er sie nur in seinem eigenen Netz betriebe. Ein solches Konstrukt ist jedoch klar von allen Autoren als Bruch der Neutralität beschrieben.

Ferner widerspräche eine solche Definition auch dem Zweck von Routing als grundlegendem Mechanismus des Internet: Routing durchbricht die Grenzen lokaler Netzwerke, eine solche Definition würde sie wieder bedeutsam machen. Das Beispiel verdeutlicht auch die Notwendigkeit einer präzisen Definition: Wird auf eine präzise Definition verzichtet und der Fokus stattdessen nur auf die Entwicklung immer feinerer Messverfahren gelegt, kommt es zu systematischen Fehlern.

Im Ergebnis ist festzuhalten, dass die Wahl des Bezugspunktes der Netzneutralitätsdefintion eine besondere Bedeutung erhält. Dies ergibt sich insbesondere aus der Verbreitung des Angebots mehrerer Dienste (Telefon, Fernsehen) als Netzwerkanwendungen in „next generation networks“ (NGN; hiermit wird die Ablösung der bislang eigenständigen Transportkanäle für Fernsehen, Telefonie und Internet durch einen Netzwerkanschluss, der als Transportkanal für Fernsehen, Internet und Telefonie verwendet wird bezeichnet). Eine Forderung nach Neutralität bezogen auf den Übergabepunkt (Anschlussdose) hat andere Konsequenzen als die Forderung nach Neutralität an der Netzwerkbuchse des Einwahlendgerätes (je nach Technologie z. B. DSL- oder Kabelmodem). Die erstgenannte Forderung erzwingt die Gleichbehandlung von Datenpaketen von VoIP- oder IPTV-Diensten welche in letzter Konsequenz zu einer geringeren Zuverlässigkeit dieser Dienste bei gleichzeitiger starker anderweitiger Netznutzung führen können. Letzterer Bezugspunkt ist streng dogmatisch nicht mit der Definition der Netzneutralität verträglich (denn es findet eine inhaltsabhängige Differenzierung statt, wobei die Inhalte pauschal durch das jeweilige Endgerät bzw. dessen Anschluß identifiziert werden).

An dieser Stelle soll ein weiterer Definitionsvorschlag in Form einer Ergänzung der Definition nach Bullinger unterbreitet werden, der den sich verändernden Realitäten Rechnung trägt:

> **Ein Netz ist neutral, wenn alle Datenpakete gleichberechtigt übertragen werden, unabhängig davon, woher sie stammen, welchen In-**

[1] die sich ebenfalls räumlich in der Wohnung des Endkunden befindet

**halt sie haben oder welche Anwendungen die Pakete generiert haben, soweit eine Abweichung nicht objektiv zur Sicherstellung der Funktion von Diensten notwendig ist.**

Diese Definition berücksichtigt die neue Rolle des Internets als Verbreitungsweg des linearen Rundfunks in hoher Auflösung (IPTV) oder als Medium für Telefoniedienstleistungen (VoIP). Diese Anwendungen können entweder vom Internetdienstleister erbracht und zusammen mit dem Internetzugang geleistet oder aber unter Benutzung eines Internetzugangs durch einen Dritten genutzt werden[2]. Hiermit unterscheidet sich auch die Zuständigkeit für die Sicherstellung der notwendigen Eigenschaften des Zugangsnetzes. In ersterem Fall ist der Zugangsanbieter auch für die Sicherstellung der vom Kunden zusätzlich beauftragten Leistungen zuständig. Dies wird er mit geeigneten Mitteln erreichen: Datenpakete für die Zusatzleistungen wie IPTV oder VoIP werden eine andere, höhere Priorisierung gegenüber Datenpaketen der weiteren Internetnutzung erhalten. Im Falle der Erbringung durch einen Dritten wird diese Priorisierung nicht auftreten, da der Internetzugangsprovider nicht involviert ist. Entsprechend liegt die Aufgabe der Sicherung entsprechender Netzqualitäten beim Anwender – und müsste ihm auf technischer Ebene ermöglicht werden.

Diskussionskern wäre bei einer solchen Definition voraussichtlich immer die Frage, welche Mittel „objektiv zur Sicherstellung der Funktion von Diensten“ notwendig sind – und welche nicht.

Im Folgenden wird von oben vorgestellter Definition ausgegangen, denn sie erlaubt es, die Betrachtung der „Netzneutralität“ auf das zu beschränken, was der Endkunde typischerweise als „Internetanschluss“ begreift – die Verbindung seines Rechners zum Internet ohne Berücksichtigung weiterer Dienste, die aus Sicht des Endnutzers hiervon unabhängig funktionieren. Ferner erlaubt sie, bei Tests der Netzneutralität von „üblichen“ Plattformen und Protokollen auszugehen.

### 4.1.1. Historischer Kontext

Auch wenn der Begriff der „Neutralität“ von Netzwerken erst um 2005 populär wurde, ist der Grundgedanke der Differenzierung von Datenströmen in solche von hoher und solche von niedriger Priorität nicht neu. Entsprechend ist auch bereits in der RFC 791 [43], mit der IPv4 standardisiert wurde, ein `TOS`-Feld u.a. zur Identifikation der Priorität des Services enthalten. Dies wurde allerdings als zu starr für eine praktische Nutzung kritisiert und fand nur geringe Verbreitung [44].

[2] Z B. VoIP als Angebot des ISP im Kontrast zu VoIP durch einen durch das Internet erreichbaren Dienstleister wie „Skype“.

Bereits vor Zeiten der heutigen Knappheit wurde ebenfalls detailliert über verschiedene Abrechnungsmodelle für knappe Netzwerkressourcen nachgedacht [45]. Von MacKie-Mason und Varian wurde dabei unter anderem der Vorschlag einer von der aktuellen Nachfrage abhängigen Preisgestaltung unterbreitet: Wenn es keine Ressourcenknappheit gibt, sollten Datenübertragungen günstig sein; kommt es allerdings zu Ressourcenknappheit, so sollten Auktionsmechanismen über die Priorität einzelner Datenpakete entscheiden. Hierzu schlugen die Autoren von „Pricing the Internet“ die Ergänzung eines `Bid`-Feldes in den Kopfdaten eines Datenpaketes vor. Hierin wäre der jeweils für eine Behandlung des Datenpakets gebotene Betrag enthalten, anhand dessen dann über die Abarbeitungsreihenfolge entschieden werden sollte [44].

Die damalig vorgeschlagenen Verfahren unterscheiden sich jedoch in zwei wesentlichen Details von den heute mit Netzneutralitätsverstößen assoziierten Vorgehensweisen:

**Information des Anwenders.** Der Netzwerkbenutzer sollte jederzeit im Klaren darüber sein, nach welchen Kriterien die Abarbeitungsreihenfolgen sortiert würden – im Fall des oben skizzierten Vorschlages sind keine weiter impliziten Abhängigkeiten neben dem gegenwärtigen Auslastungszustand des Netzes (welcher im Minutentakt bestimmt werden sollte) und dem `Bid`-Feld des Datenpaketes vorhanden. Beides ist für den Nutzer zugänglich gestaltet.

**Unabhängigkeit von den übermittelten Inhalten.** Die Frage, welche Daten hoch und welche Daten niedrig priorisiert übertragen werden sollten, wurde komplett dem Endnutzer und seiner Entscheidung welche Einträge im `Bid`-Feld vorgenommen werden überlassen werden. Der Nutzer könnte sowohl einen Dateitransfer mit einem beliebigen Protokoll oder auch eine interaktive Kommandozeilenanwendung mit besonderer Priorität versehen. Dies steht in deutlichem Kontrast zur heutigen Neutralitätsdebatte, in der die Identifikation hoch- bzw. niedrigpriorisierter Datenströme mittels Deep-Packet-Inspection oder statistischer Protokollidentifikation seitens des Internetzugangsproviders erfolgt. Darüber hinaus hat eine solche Klassifikation die Tendenz zur deutlichen Abhängigkeit von den jeweils verwendeten Protokollen oder gar den übermittelten Inhalten.

Aus historischer Perspektive ist also ein wesentlicher Punkt der Netzneutralitätsdebatte die Verlagerung der Entscheidung über hochpreisige (wobei auch verschlechterte Netzwerkqualität als Preis betrachtet werden kann) Dienste aus der Hand des Nutzers in die des Internetzugangsanbieters, wobei dem Nutzer sowohl der Einfluss auf und Kenntnis entsprechender Vorgänge entzogen werden.

### 4.1.2. Öffentliche Debatte

Die öffentliche Debatte um die Netzneutralität ist stark von Radikalforderungen, Zuspitzungen, Polemik und gegenseitigen Schuldzuweisungen gekennzeichnet. Wie schlüpfrig die Debattenlage ist, lässt sich leicht am Eingangsbeispiel darstellen.

*Die Aussage „Es bleibt der Eindruck, dass von Netzbetreibern nur das zugegeben wird, was ihnen unbestreitbar nachgewiesen wurde“ ist eine unwiderstehliche Einladung für einen Disput. Im Fall des ersten Einführungsbeispiels der Benutzung von „Skype“ mit Fokus um die rechtlichen Details eines Nutzungsvertrags; im zweiten Einführungsbeispiel um einen Download durch das vom Universitätsrechenzentrum betriebene Netzwerk um die Legitimität von Bemühungen, möglichst ohne Einfluss auf die Nutzer ein sicheres Netz zu betreiben. Wieder einmal illustrieren die Beispiele das Spektrum der möglichen Ausformungen von Netzneutralität sowie der mit ihr verbundenen Auswirkungen; sind es hier mit dem Absatz eines weiteren Tarifs und der Absicherung eines lokalen Netzes zwei sehr unterschiedliche Motive, die beide zu einem Verhalten führten, dass als Neutralitätsverletzung betrachtet werden kann.*

In diesem Abschnitt soll daher nur ein kurzer Abriss über die Forderungen in der deutschen Öffentlichkeit gegeben werden.

Von Seiten der **Inhalteanbieter** wird die Neutralität des Internets gerne als Innovationsantrieb beschrieben, durch den es einem kleinen Start-Up-Unternehmen möglich ist, im Internet die gleiche Präsenz zu besitzen wie ein multinationaler Konzern. Einige der Erfolgsgeschichten des Internets[3] werden regelmäßig als entsprechende Beispiele präsentiert; Forschungsergebnisse [46] stützen diese These.

Von Seiten der **Netzbetreiber** wird hingegen kritisiert, dass die Inhalteanbieter ihre Gewinne auf dem Rücken der Infrastrukturbetreiber erwirtschaften würden und dass es nur fair wäre, diese an den Gewinnen entsprechend der verursachten Netzlast zu beteiligen.

Von **Personen die sich selbst als Interessensvertreter einer postulierten schweigenden Mehrheit (der „Netzgemeinde“) exponieren** wird regelmäßig im Kontext der Debatten um die Netzneutralität vor dem Schritt zum „Zwei-Klassen-Internet“ gewarnt. Die Vorstellung hinter dieser Warnung ist die Befürchtung, dass Internetzugangsanbieter die Nutzer günstiger Zugänge auf weniger Dienste zugreifen lassen könnten als die Nutzer hochpreisige Zugänge bzw. dass sie Aufschläge für einzelne Dienste[4] verlangen könnten.
Ferner wird das Thema Netzneutralität von dieser Seite regelmäßig in einen Kontext mit Befürchtungen einer „Zensur“ des Internets durch staatliche oder private Stellen gebracht; dabei steht die Befürchtung im Vordergrund, dass eine Infrastruktur eingerichtet werden könnte, die später für Eingriffe in eine „Freiheit des Internets“ [75], gemeint ist vornehmlich die Informationsfreiheit und der Schutz unliebsamer Meinungen vor Unterdrückung, verwendet werden könnte.

Diese Diskussion ist im Verlauf der letzten fünf Jahre aus einer Nische in das öffent-

[3] „Facebook“, „StudiVZ“, „Amazon“
[4] z. B. Dateiaustausch

liche Bewusstsein gerückt. Insbesondere die im Frühjahr 2013 publik gewordenen Pläne der Deutschen Telekom, bisherige Flatrates durch Datenvolumenkontingente mit nachfolgend gedrosseltem Durchsatz zu ersetzen und einzelne Dienste von der Berechnung des Kontingentsverbrauchs auszunehmen, hat die Debatte um die Netzneutralität auch nicht technikaffinen Bevölkerungskreisen bewusst gemacht und führte möglicherweise zu einem Entwurf einer Verordnung des Bundeswirtschaftsministerium wie nach § 41a Telekommunikationsgesetz vorgesehen (vgl. Abschnitt 4.3.1).

Die Arbeit beteiligt sich nicht an dieser Debatte.

### 4.1.3. Unumgehbarkeit von Neutralitätsverletzungen in Zugangsnetzen

Naheliegend ist zunächst der Gedanke, Neutralitätsverletzungen durch den Internetzugangsanbieter aus Sicht des Endnutzers technisch umgehen zu wollen.

*Im zweiten Einführungsbeispiel ist ein Umgehen nichtneutraler Teile des Routinggraphen möglich: Per SSH-Zugriff auf weitere Rechner, die über einen anderen ISP als dem Universitätsrechenzentrum an das Internet angebunden waren, und nachfolgende verschlüsselte Übertragung konnten die durch die Infrastruktur eigentlich auszufilternden Inhalte dennoch geladen werden.*

Es scheint also denkbar, auch in größerem Stil durch Verschlüsselung die Analyse des Inhalts zu erschweren und Identifikation von Datenströmen durch DPI ins Leere laufen zu lassen. Auch die Benutzung von Verschleierungsdiensten wie TOR erscheint denkbar, um dem Provider die Analyse von Kommunikationspartnern zu erschweren.

In der Praxis sind solche Herangehensweisen jedoch wenig erfolgversprechend: Eine Ende-zu-Ende Verschlüsselung (die also zwischen dem Contentanbieter und dem Rechner des Endnutzers besteht) mag den Inhalt in gewissem Grad einem Mitlesen durch den Internetzugangsprovider entziehen, allerdings bleiben die Informationen über Absender und Empfänger der Datenpakete enthalten. Daneben kann einer Identifikation durch statistische Verfahren nicht wirksam begegnet werden; Wright et al demonstrierten 2008 die Erkennung einzelner Phrasen innerhalb verschlüsselter VoIP-Verbindung durch Verwendung von HMMs [47]; Liu demonstriert die Erkennung bestimmter Webanwendungen bei Verwendung von HTTPS [31] – jeweils ohne die Verschlüsselung zu brechen.

Ein weiterer Ansatz könnte in der generellen Nutzung von Overlaynetzwerken liegen. Die Grundidee von Overlaynetzwerken ist die Nutzung bestehender Netzwerkverbindungen, um in einer weiteren Abstraktionsschicht ein weiteres Netzwerk aufzubauen, das wiederum eine andere Topologie besitzen kann. Diese neue Topologie ist dabei begrenzt von der unterliegenden (realen) Topologie abhängig: Die Netzqualitäten können sich nicht verändern, nur die logische Struktur. Innerhalb eines Overlaynetzwerkes können Datenströme nach optimierten Kriterien verteilt und sich ergebende Redundanzen genutzt werden. Diese können unter Kenntnis der unterliegenden Architektur zur schnellen Umgehung von Engpässen genutzt werden [48], führen bei Nutzung verschiedener Techniken

jedoch zu Interoperabilitätsproblemen, welche gegenwärtig nur durch Gatewaykonstrukte gelöst werden können [49]. Einen detaillierten Überblick über Overlaynetzwerke und verwendete Werkzeuge bietet Ding [50]. Im Sinne der Umgehung von Neutralitätsverletzungen sind Overlaynetzwerke in soweit interessant, als dass sie Zugriff auf Netzwerkressourcen bieten können, ohne dass dem Betreiber des unterliegenden Netzwerks die Zieladressen vorliegen – dieser erfährt lediglich die Adressen anderer Knoten im Overlaynetzwerk, nicht jedoch das eigentliche Anfrageziel. Dieser Undurchschaubarkeit stünde von Seiten eines aggressiven Internetzugangsanbieters eine Klassifikation als nicht identifizierter und damit nicht privilegierter Datentransfer – eine Einteilung in die unterste Klasse von Netzwerkdatenübertragungen – gegenüber. Somit bliebe gegenüber dem Provider zwar die Verschleierung des eigentlichen Anfrageziels, dies muss aber keinen Vorteil in der Klassifikation bedeuten. Auch TOR kann als Overlaynetzwerk angesehen werden. Der Fokus des Tor-Projekts liegt primär auf der Verschleierung von Sendern und Empfängern, nicht aber auf Verschlüsselung der Daten. Die Nutzung von TOR [76] würde sich zusätzlich durch ein hohes Datenaufkommen zu einem oder mehreren der bekannten TOR-Einstiegsknoten verraten.

## 4.2. Nachweis von Netzneutralitätsverletzungen

In diesem Abschnitt soll der gegenwärtige Forschungsstand zum Nachweis von Neutralitätsverletzungen systematisiert und bewertet werden; er wird Ausgangspunkt für Kap. 5 sein.

Hierzu werden die vorgestellten Verfahren in drei Gruppen klassifiziert: In die Gruppe der **aktiven Ansätze**, die der **passiven** und die der **hybriden** Ansätze. Diese Trennung ergibt sich aus dem Vorgehen einzelner Verfahren.

*Um die Einführungsbeispiele zu bemühen: Am Anfang des ersten Beispiels stand die reine Beobachtung, dass ein bestimmter Dienst („Skype“) eine besonders schlechte Leistung aufwies bzw. dass im zweiten Beispiel ein Download fehlschlug. Solche, rein beobachtende Untersuchungen ohne nur zum Zweck der Messung von Leistungsdaten – wie Durchsatz, Latenz oder Jitter – durchgeführte Datentransfers werden im folgenden als passive Messung klassifiziert. Der nächste Schritt war in beiden Fällen ein Wechsel des Vorgehens: Es wurden Datentransfers zum Zweck des Messens bestimmter Parameter durchgeführt (Download von* `skype.com` *bzw. Nutzung eines Testservers, von dem gezielt bestimmte Bytesequenzen heruntergeladen wurden, um das vom Filter gesuchte Muster zu identifizieren). Die Verwendung eines Datentransfers, der nur zur Messung seiner Qualitäten erzeugt wird, ist kennzeichnend für Verfahren des aktiven Ansatzes.*

*Das gesamte Vorgehen im Beispiel 2 – angefangen mit der neugierige Beobachtung der alltäglicher Netznutzung bis hin zum expliziten Verfolgens interessanter Artefakte mit einem Testsetup – führt zu einem möglichen hybriden Ansatz als Kombination aktiven und passiven Vorgehens.*

Eine solche Klassifikation in aktive, passive und hybride Verfahren erweist sich auch deshalb als sinnvoll, da sich die Aussagen von Tests aus unterschiedlichen Klassen maximal innerhalb einer Klasse vergleichen lassen[5]. Im Rahmen der Darstellung der einzelnen Gruppen von Testverfahren werden auch die durch einen Test erreichbaren Aussagen untersucht.

Im Folgenden wird zunächst das generelle Nachweisproblem der Netzneutralität sowie dessen mögliche Lösungen erörtert, bevor die drei grundlegende Herangehensweisen beschrieben und eingeordnet werden. Schließlich werden die Rückwirkungen der jeweiligen Messverfahren auf die Messgröße und die Auswirkungen anderer (gleichzeitig zur Messung stattfindender) Datenübertragungen betrachtet. Zum Abschluss dieses Abschnitts werden die im vorangehenden Kapitel vorgestellten Vertreter den drei Herangehensweisen zugeordnet.

### 4.2.1. Grundlegendes Nachweisproblem von Neutralitätsverletzungen

Um die Frage nach der Neutralität eines Netzwerkes abschließend zu beantworten, müsste nur die tatsächlich in der gesamten betrachteten Infrastruktur eingesetzte Firmware auf die ausschließliche Verwendung von Best-Effort-Algorithmen, deren einziges Kriterium die Zieladresse eines Datenpaketes ist, geprüft werden. Dies wäre das einzig sichere Verfahren, um die Neutralität eines Netzes positiv nachzuweisen. Da ein solches Vorgehen nicht realisierbar ist, wird stattdessen die Frage nach der Abwesenheit von Neutralitätsverletzungen untersucht. Da eine Neutralitätsverletzung aber auch bereits dann besteht, wenn nur eine einzige Diskriminierung oder Privilegierung vorliegt, entsteht ein Dilemma: Für einen tatsächlichen Nachweis der Neutralität eines Anschlusses müssten (gemäß Definition nach Bullinger) Datentransfers mit sämtlichen denkbaren Verbindungspartnern mit sämtlichen Protokollen bzw. Diensten mit sämtlichen Inhalten untersucht werden; nach der im Rahmen dieser Arbeit aufgestellten Definition wäre für jeden Dienst, der als privilegiert identifiziert werden kann zusätzlich zu prüfen, ob diese Privilegierung objektiv technisch gerechtfertigt wäre.
Solche Tests sind zwar denkbar, würden jedoch unweigerlich an der Durchführung scheitern, schon weil keine Vergleichsdaten zur Verfügung stehen, welches Verhalten von welcher Gegenstelle bei einer Verbindung mit welchem Protokoll zu erwarten wäre bzw. welche Verbindungsqualität in einem unbeeinflussten Netzwerk zu erwarten wäre. Weitere Probleme ergeben sich auch aus der Vielzahl zu prüfender Verbindungen (Hosts im Internet × Layer-4-Protokolle × Layer-4-Ports × Upper-Layer-Protokolle × Inhalte[6]).

Aus der Architektur des Netzwerkes, insbesondere aus dem Routing, ergibt sich ein weiteres Problem: Aus der Position eines Endnutzerrechners (eines Endknotens im Routingpfad) sind die folgenden Fälle nicht unterscheidbar, weil sie das gleiche Symptom

[5] und auch innerhalb der Gruppe ist eine Vergleichbarkeit schwierig
[6] Bereits diese allein sind allenfalls auf- jedoch nicht abzählbar.

(Timeout beim Versuch eines TCP-Verbindungsaufbaus bzw. keine Reaktion) erzeugen würden:

- Die entfernte Maschine reagiert nicht (z. B. weil das Betriebssystem keine Ressourcen für die Annahme einer weiteren Verbindung aufbringen kann),
- Datenpakete zur Zielmaschine werden von einem Knoten des Routingpfades nicht weitergeleitet,
- Datenpakete von der Zielmaschine werden von einem Knoten des Routingpfades nicht weitergeleitet,
- die entfernte Maschine existiert überhaupt nicht bzw. ist nicht mit dem Internet verbunden oder die Internetadresse ist keiner Maschine zugewiesen.

Diese Ununterscheidbarkeiten basieren maßgeblich auf der Weiterleitung von Datenpaketen über mehrere Router: Da es nicht möglich ist, vom Endknoten aus einen alternativen Routingpfad vorzugeben und so einzelne Router als Ursache auszuschließen, werden Datenpakete zu einem bestimmten Ziel immer die gleichen Routingknoten passieren. Da sich die Einflüsse der einzelnen Routingknoten auf das Datenpaket kumulieren (und Identifikation der Einzeleinflüsse ohne weitere Informationen ein Problem vergleichbar der Zerlegung einer Summe in ursprüngliche Summanden ein unlösbares Problem darstellt), kann am Endknoten ohne weitere Messungen nur vermutet werden, ob die Beobachtung vom beabsichtigten Host stammt oder auf eine der Routing-Stationen zurückzuführen ist.

Weitere Effekte kommen durch Optimierungstechniken hinzu, die beispielsweise eine gleichmäßige Auslastung mehrerer Verbindungen anstreben. Diese führen dazu, dass die vom Endknoten aus bestimmbaren Informationen über Netzwerkrouten inkonsistent erscheinen können [26].

Weitere Anforderungen an ein Messverfahren ergeben sich aus der Annahme eines böswilligen Netzbetreibers, der Messungen zu erkennen und zu manipulieren versucht, um keine Nachweis eines Verstoßes gegen Neutralitätsgrundsätze entstehen zu lassen.

Gemeinsam ist den im Folgenden beschriebenen Ansätzen die Notwendigkeit des Aufstellens mehrerer Datenreihen, zwischen denen im Anschluss unterschieden werden kann. Das Ziel der Analyse ist die Feststellung der kausalen Abhängigkeit eines oder mehrerer Qualitätsparameter von einer oder mehreren der Eigenschaften Quell- und Zieladresse, Quell- und Zielport, Protokoll oder Inhalt.

Schließlich soll noch die Notwendigkeit von Zeitstempeln im Zusammenhang mit der Dokumentation von Neutralitätsverstößen erwähnt werden: Sollte es zum Streit zwischen Kunde und ISP kommen, muss die Dokumentation eines Verhaltens mit einem Zeitstempel versehen sein, der eine Nachvollziehbarkeit erlaubt – und anhand dessen alternative Ursachen für ein bestimmtes Netzwerkverhalten (z. B. eine Congestion auf Grund eines

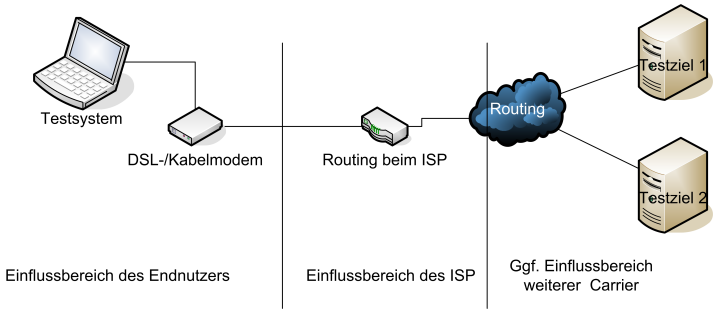


Abbildung 4.2.: Aufbau für eine Messung nach dem passiven Ansatz. Optional ist eine zentrale Auswertung der Daten möglich.

Angriffs) zeitlich ausgeschlossen werden können. Da sich Begründungen für ein bestimmtes Netzverhalten in Größenordnungen von Minuten oder Sekunden bewegen, scheint eine Zeitsynchronisation per NTP ausreichend.

Nichttechnische Ansätze wie Crowdsourcing (z. B. im Rahmen des Projekts Herdict [16]) werden im Rahmen dieser Arbeit nicht weiter beschrieben; der Fokus liegt bei der folgenden Klassifikation auf rein technischen Nachweisverfahren. Zunächst wird die passive, die aktive und schließlich die hybride Herangehensweise beschrieben.

### 4.2.2. Passiver Ansatz

Beim passiven Ansatz wird nur der vom Nutzer ohnehin verursachte Datenaustausch beobachtet. Aus der Auswertung solcher Beobachtungen (mehrerer Nutzer) werden dann Rückschlüsse auf die Neutralität der verwendeten Netze gezogen.

Der passive Ansatz besteht entsprechend aus einer Monitoring-Komponente und einer Auswertungskomponente. Die Monitoring-Komponente beobachtet die ein- und ausgehenden Netzwerkdaten und zeichnet deren Eigenschaften auf; die Auswertungskomponente versucht aus diesen Daten Rückschlüsse auf das angeschlossene Netzwerk zu ziehen, wobei diese Komponente auch zentral angesiedelt sein kann, um die Daten mehrerer Versuchsaufbauten (Monitoring-Komponenten) auszuwerten. Die benötigte Infrastruktur ist in Abb. 4.2 skizziert.

Passive Verfahren sind in der Regel für Protokolle ab Ebene 4 des ISO-/OSI-Modells unsichtbar, da die Messung der Netzwerkperformance auf IP-Basis geschieht.

Die Durchführung eines Tests besteht aus der „normalen“ Nutzung eines Netzwerkanschlusses und gleichzeitiger Beobachtung durch die Monitoring-Komponente. Dabei werden die ein- und ausgehenden Pakete beobachtet und Daten wie Quell- und Ziel-Adresse, Quell- und Zielport sowie Protokoll festgehalten. Durch einen ausreichend langen Be-

obachtungszeitraum[7] (oder die Nutzung von Daten mehrerer Monitoring-Komponenten) wird eine ausreichende Menge von Testdaten aggregiert, um eine statistische Auswertung zu erlauben. Es gibt entsprechend keine zeitlichen Fenster, innerhalb derer eine Messung durchgeführt wird; die Monitoring-Komponente ist immer aktiv.

**Einfluss der Messung auf die Messgröße** Passive Verfahren weisen Prinzip bedingt weniger Ungenauigkeiten auf, die durch die Messung selbst verursacht werden, denn die Messung erzeugt keinen zusätzlichen Datenverkehr (abgesehen von gelegentlicher Übermittlung gebündelter Messdaten zu einer zentralen Auswertstelle, soweit erforderlich); entsprechend kann es hier auch keinerlei Einfluss geben. Denkbar sind dennoch andere Einflüsse, die allerdings dem Rechner entstammen, auf dem die passive Beobachtung des Datenverkehrs stattfindet: Kommt es zu einer erhöhten Auslastung kritischer Systemressourcen wie z. B. der CPU durch einen sehr rechenaufwändigen anderen Prozess, so ist nicht mehr sichergestellt, dass eine Beobachtung aller Datenpakete mit hinreichender Genauigkeit erfolgen kann.
Der Einfluss passiver Verfahren auf die Messwerte steigt allerdings nennenswert, wenn sich die Monitoring-Komponente auf einem eingebetteten System wie etwa einem Router befände, da diese gegenwärtig noch über deutlich geringere Ressourcen verfügen als klassische Workstations. Hieraus resultierende Effekte setzten allerdings eine hinreichend viel Rechenlast erzeugende Implementation voraus. Eine allgemeine Aussage jenseits eines Trends ist ohne weitere Spezifikation eines Systems (sowohl der Hardware wie auch der verwendeten Software) nicht möglich.

**Vorteile des passiven Ansatzes** Die Vorteile des passiven Ansatzes liegen vor allem in der Ausweitung der beobachteten Verbindungen: Anstelle eines vorgegebenen Testsets werden (im Idealfall) sämtliche vom Nutzer versandten und empfangenen Pakete in die Auswertung einbezogen. Ausgenommen sind dabei Daten, deren Anzahl für eine statistisch hinreichend sichere Antwort nicht ausreichen würde.
Da sich die mit Hilfe des passiven Ansatzes getroffenen Aussagen auf alle beobachteten Verbindungen beziehen[8], ist die Aussage aus Nutzersicht mit einer pauschalen Aussage über den Internetanschluss als solchen identisch.

**Nachteile des passiven Ansatzes** Die Nachteile des passiven Ansatzes ergeben sich vor allem aus der (in der Praxis notwendigen) zentralen Datensammlung und Datenlage. Die erhobenen Daten, bestehend aus Quell- und Zieladressen bzw. -ports sowie verwendeten Protokollen erlauben Rückschlüsse auf das Nutzungsverhalten[9]. Diesen Bedenken kann begegnet werden, indem die Monitoring-Komponente über eine

[7] Abhängig von der Intensität der Netznutzung – für eine Auswertung müssen hinreichend viele Datensätze zur Verfügung stehen.

[8] bei Vernachlässigung von für eine Auswertung zu selten kontaktierten Gegenstellen

[9] Angedeutet seien hier nur Verbindungen zu One-Click-Hostern oder Webseiten mit pornographischen Inhalten sowie zeitliche Profile; daneben sei auf die Betrachtungen zur Personenbeziehbarkeit von IP-Adressen verwiesen.

Deaktivierungsfunktion verfügt. Diese erlaubt dem Nutzer, sich für einen gewissen Zeitraum „unbeobachtet“ im Internet zu bewegen. Eine solche Option verringert natürlich die Zahl der beobachteten Verbindungen, folglich entspricht die beobachtete Verbindungsmenge nicht mehr den tatsächlich genutzten Verbindungen. Eine Aussage über die Gesamtqualität der Internetverbindung hinsichtlich der Neutralität ist nicht mehr möglich.
Insbesondere würde ein solches Vorgehen wahrscheinlich zu einer systematischen Lücke bei der Beobachtung bestimmter Dienste führen – es wäre nur plausibel, wenn viele Anwender die Monitoring-Komponente während ihrer Besuche auf Seiten von zweifelhaftem Ruf (Pornographische Inhalte, illegaler Dateiaustausch) deaktivierten.

**Aussage des passiven Ansatzes** Die Aussage des passiven Ansatzes kann als

> Die getestete Internetverbindung zu allen Zielen $t_1, ..., t_n$, mit denen mithilfe der Protokolle $p_1, ..., p_m$ zu den Zeitpunkten $(s_1, ..., s_o)$, $m, n, o, i, j, k \in \mathbb{N}$; $s_i$ geeigneter Zeitstempel, die Datensätze $d(s_i, p_j, t_k)$ ausgetauscht wurden, ist für diese Daten [nicht] neutral. Die Ziele $t_1, ..., t_n$ ergeben sich aus der Netznutzung des Anwenders.“

beschrieben werden.

### 4.2.3. Aktiver Ansatz

Im Unterschied zum gerade beschriebenen passiven Vorgehen verfolgen aktive Ansätze die Idee des direkten Testens bestimmter Verbindungen unter detaillierter Beobachtung der Verbindungscharakteristika mit anschließendem Vergleich.

Die Durchführung des Tests besteht aus (meist) mehreren Datentransfers von oder zu entfernten (bekannten und ggf. auch kontrollierten) Zielen; der klassische Infrastrukturaufbau ist in Abb. 4.3 skizziert. Eine Beobachtung der Verbindung ist sowohl auf Seiten des testenden Systems wie auch des entfernten Testservers möglich, da beide zum Versuchsaufbau gehören und entsprechend detaillierte Aufzeichnungen vornehmen können. Um eine Differenzierung durch Netzbetreiber zu identifizieren, werden unterschiedliche Daten ausgetauscht, die jeweils charakteristisch für eine bestimmte Nutzungsform sind, wie etwa Daten eines BitTorrent-Handshakes oder eines HTTP-Dateitransfers.

Im Anschluss an die Datentransfers werden die dabei ermittelten Daten (wie etwa Durchsatz, Latenz, Jitter) statistisch analysiert, um – für den Fall einer Neutralitätsverletzung – eine Abhängigkeit von Ziel- oder Quelladressen, Ports, Protokoll oder übertragenen Daten (Inhalt) nachzuweisen.

**Einfluss der Messung auf die Messgröße** Auf Grund ihrer Fundierung im eigenständigen Versenden von Testdaten, die als einzige Grundlage der Auswertung dienen,

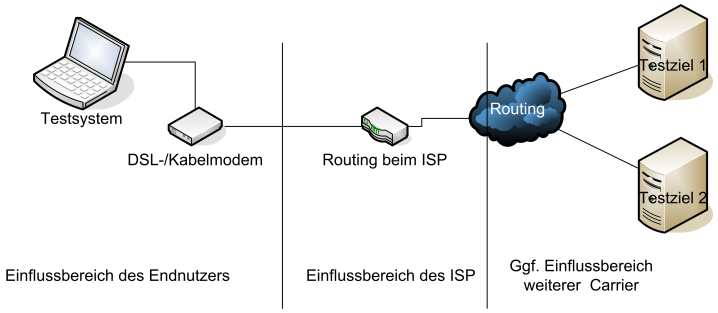


Abbildung 4.3.: Beispiel für eine Nutzung eines Tests nach dem aktiven Ansatz von einem Endkundenanschluss aus.

sind aktive Verfahren besonders anfällig für Einflüsse der Messungen auf die Messgrößen. Für eine Beeinflussung ergeben sich mehrere Möglichkeiten:

- Einflüsse auf dem Testsystem: Zunächst besteht potentiell eine Konkurrenz um den Netzwerkzugriff mit anderen, gleichzeitig ausgeführten Programmen. Im Rahmen aktiver Tests werden ferner erhöhte Datenmengen mit anderen Peers ausgetauscht. Hierzu müssen diese Daten auf dem Testrechner generiert und versandt werden. In diesem Sinne sind bereits die ausgesandten Testdaten den Scheduling- und Priorisierungsalgorithmen des Betriebssystems ausgeliefert; die Erfassung von Messwerten sowohl ausgehender wie auch eingehender Pakete ist von der Auslastung des PCs abhängig.
- Einfluss im lokalen und in weiteren Netzwerken: Ausgesandte Daten können in Netzwerken – wie alle anderen Datenpakete auch – zu Engpässen führen oder diese weiter zuspitzen. In diesem Sinne kann ein aktiver Test das Testergebnis bei einem ohnehin bestehenden Engpass weiter verschlechtern.
- Auslösung von Abwehrmaßnahmen seitens des Netzwerkbetreibers: Netzwerkbetreiber könnten Tests detektieren und entsprechend andere, neutral erscheinende, Policies aktivieren.

**Vorteile des aktiven Ansatzes** Die Vorteile des aktiven Ansatzes liegen in einer relativ leichten Umsetzbarkeit, die schnell (bezogen auf die Testdauer) zu eindeutigen Ergebnissen führt. Ebenso führen diese Eigenschaften dazu, dass ein solcher Test auch von Personen ohne allzu tiefe Sachkenntnis verwendet werden kann, da sich seine Verwendung bei den Umsetzungen dieses Ansatzes typischerweise auf die Ausführung eines Programms durch den Endnutzer beschränkt.

**Nachteile des aktiven Ansatzes** Die Nachteile des aktiven Ansatzes liegen in der Abhängigkeit der Ergebnisse von der Auswahl des genauen Testverfahrens und der Test-Peers. Letztere sind als bekannt anzunehmen, daher könnte ein bösartiger ISP Netzwerkpolicies einrichten, die diese Peers von jeder Diskriminierung ausnehmen.

In einer derart präparierten Netzwerkumgebung würde ein Test immer zu einem – vom ISP gewünschten – negativen Ergebnis führen (vgl. hierzu auch Anhang D, in dem die Nutzbarkeit dieses Nachteils demonstriert wird).
Einen vergleichbaren Einfluss hat auch die Auswahl der verwendeten Testdaten und -protokolle, denn sie können entweder Mechanismen seitens des Providers auslösen bzw. zu einer Einordnung der Datenströme in eine diskriminierte oder privilegierte Kategorie führen – oder nicht. Schließlich ist es auch denkbar, dass der Netzbetreiber die Testdaten erkennt.

**Aussage des aktiven Ansatzes** In der Summe führen die Nachteile dazu, dass das Ergebnis eines Tests nach dem Prinzip des aktiven Ansatzes keinen Anspruch auf Allgemeingültigkeit haben kann. Die Aussage eines solchen Tests kann nur lauten

> „Die getestete Internetverbindung zu Testziel $t_1, ..., t_n$ mit den Testdaten (die von Ziel und Protokoll abhängig sein dürfen) $d(p_i, t_i)$ die mittels der Protokolle $p_1, ..., p_m$ zum Zeitpunkt $s_1, ..., s_o$ ($i, n, m \in \mathbb{N}$, $s_i$ sei geeigneter Zeitstempel) ausgetauscht wurden verhält sich **für diese Daten** $d(p_i, t_i)$ [nicht] neutral.“.

Eine generalisierte Aussage der Form „der getestete Internetanschluss verhält sich neutral“ ist durch den durchgeführten Test nicht seriös belegbar.

### 4.2.4. Hybride Ansätze

Hybride Ansätze vereinen aktive mit passiven Komponenten. Als konkretes Beispiel wurde im vorangehenden Kapitel Nooter vorgestellt.

**Einfluss der Messung auf die Messgröße** Hybride Verfahren kombinieren wie schon bei den Vorteilen allerdings auch die Probleme des Einflusses auf die Messgröße: Aktive Teile laufen regelmäßig Gefahr, vom Netzwerkbetreiber als solche erkannt zu werden (konkret auf das skizzierte hybride Verfahren bezogen hieße dies Identifikation der Kommunikation mit der Gegenstelle im Providernetzwerk). Das hier vorgestellte hybride Verfahren Nooter löst dieses Probleme nicht auf der technischen Ebene. Stattdessen ist es so ausgerichtet, dass die möglichen Konsequenzen aus Einflussnahmen stets zu Gunsten des Nutzers ausfallen – und auffallen.

**Vorteile hybrider Verfahren** und

**Nachteile hybrider Verfahren** sind entsprechend eine Kombination der zu Grunde liegenden aktiven bzw. passiven Komponenten. Ziel bei der Entwicklung eines hybriden Verfahrens wird stets die Kombination der Vorteile sein.

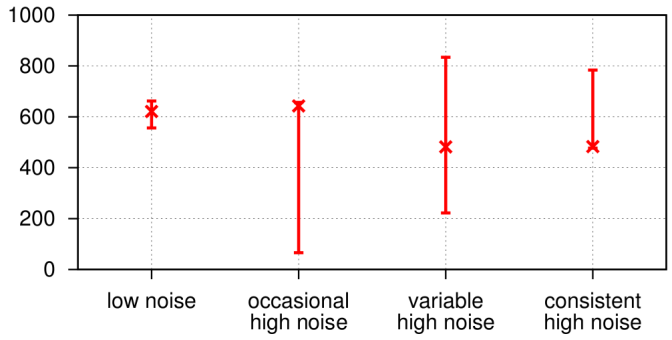


Abbildung 4.4.: Übersicht der im Projekt Glasnost entdeckten Ausformungen von Rauschen, Graphik entnommen aus [13]. Angegeben sind Minimum, Maximum und Median; Angaben in kbps.

### 4.2.5. Einfluss weiterer Datentransfers auf Messungen

Um möglichst präzise Aussagen zur Kausalität zwischen angenommenen Einflüssen und dem Messergebnis treffen zu können, sind unerwünschte andere Einflüsse weitest möglich auszuschließen. Im Falle eines aktiven Tests wäre das Netzwerk idealerweise bis auf die Tests ungenutzt. Unter realen Bedingungen gibt es jedoch praktisch immer andere Datenströme, die eine Messung beeinflussen und die sich in der Auswertung als Rauschen bemerkbar machen. Diese Datenströme können sowohl vom Rechner auf dem der Test ausgeführt wird ausgehen, wie auch dem selben lokalen Netzwerk (andere Nutzer) oder Nutzern in Netzwerken, die nicht unter Kontrolle des Endanwenders sind. Somit lassen sich diese Effekte auch nicht immer durch Hinweise an den durchführenden Benutzer eliminieren. Bei Nichtbeachtung von Phänomenen, die sich aus anderen Datentransfers ergeben, steigt die Gefahr falscher Schlüsse: Was in Wirklichkeit nur eine Verringerung des Durchsatzes auf Grund einer konkurrierenden Datenübertragung war (also ein klassischer Fall einer Congestion im Best-Effort-Ansatz), wird als Neutralitätsverletzung klassifiziert. Der Umgang mit dem Einfluss anderer Datenübertragungen wird ausführlich in [13] betrachtet und soll hier kurz wiedergegeben werden.

Das durch andere Datenströme hervorgerufenes Rauschen lässt sich in vier Klassen einteilen (vgl. Abb. 4.4, [13]):

- Grundsätzlich invariantes (und geringes) Rauschen („low noise“); alle Datenströme des aktiven Tests werden ähnlich beeinflusst.
- Meist geringes Rauschen, einzelne Störungen haben jedoch erheblichen Einfluss („occasional high noise“). Die einzelnen Ausreißer sind bei entsprechender Auswertung tolerierbar.
- Hohe Streuung, der Median ist entfernt von Maximum wie Minimum („variable

high noise“). Im Projekt Glasnost werden derartige Messungen verworfen, da eine Beeinflussung auf Grund einer Verletzung der Netzneutralität nicht von einer Beeinflussung durch andere Datentransfers unterschieden werden kann.

- Einzelne Phasen geringen Rauschens („consistent high noise“), ansonsten hohe Last, die für geringen Durchsatz sorgt. In diesem Fall ist eine Analyse schwierig bzw. nicht möglich, da die Verbindung weite Teile des Tests über eine geringe Leistung zeigte – dieser aber in einzelnen Peaks anstieg; aus nicht feststellbaren Gründen ist ein höherer Durchsatz als im Median möglich.

Im Falle passiver Testverfahren addieren sich die beschriebenen Phänomene zu ohnehin schon Prinzip bedingt höherem Rauschen, da der Test keinen Einfluß auf die Datenbasis hat.

### 4.2.6. Einordnung und Bewertung der vorgestellten Nachweisverfahren

In diesem Abschnitt werden im vorangehenden Kapitel vorgestellten Nachweisverfahren in die Kategorien der aktiven, passiven und hybriden Verfahren eingeordnet.

Eine auf einem Test basierende Aussage über die Neutralität oder Nichtneutralität eines Netzwerkes ist unabhängig vom Testverfahren erst dann richtig einzuschätzen, wenn auch Kenntnis über die Sicherheit besteht, mit der diese Aussage getroffen werden kann. Konkretes Ziel aller Implementierungen ist eine möglichst geringe Zahl von False-Positives (also fälschlich als nichtneutral erkannter Konstellationen, die tatsächlich aber neutral sind) sowie von False-Negatives (der als neutral erkannten Konstellationen, die tatsächlich aber nichtneutral sind).

Vergleiche dieser Kennziffern unterschiedlicher Verfahren sind schwierig, da die unterschiedlichen Verfahren (auch innerhalb einer Gruppe von Testverfahren) stets eine unterschiedliche Aussage treffen. Aus diesem Grund wird bei den im folgenden Abschnitt dargestellten existierenden Testinstrumenten stets die genaue Bedeutung getroffener Aussagen fokussiert.

#### Fathom

Mit Hilfe des Browserframeworks Fathom sind sowohl aktive wie auch passive Messverfahren umsetzbar.

Fathom bietet mit NANO vergleichbare Vorteile der passiven Beobachtung von im Netzwerk ausgetauschten Daten. Da das Framework jedoch auf „Mozilla Firefox“ aufbaut, wird es auch nur diejenigen Datenflüsse beobachten und analysieren können, die auch von diesem Browser verwendet werden. Dies limitiert die beobachteten Protokolle

deutlich. Von Vorteil ist jedoch die plattformunabhängige und als Plugin leicht einzupflegende Implementation, die es – gerade im Vergleich zu NANO – auch weniger versierten Nutzern erlauben würde, Messungen durchzuführen. Die Zukunft von Fathom hängt offensichtlich maßgeblich von der Zukunft des Browsers „Mozilla Firefox“ ab.

Durch eine passive Untersuchung mit Fathom erlangte Aussagen werden sich systematisch auf die von „Mozilla Firefox“ unterstützten Protokolle und die vom Nutzer mit ebendiesem Browser besuchten Webseiten beschränken. Ferner sind Messungen unmittelbar von der Leistung des Browsers und der Performance der JavaScript-Ausführungsumgebung abhängig.

### Glasnost

Glasnost ist den aktiven Verfahren zuzuordnen.

Hervorzuheben und sicherlich maßgeblich für die Verbreitung von Glasnost ist die Konzentration auf eine Oberfläche, die auch Benutzern ohne Sachkenntnis die Bedienung erleichtert. Wünschenswert (wohl aber mit der zeitlichen Obergrenze eines Tests kollidierend) wäre die Ausweitung auf weitere Testdaten, mit denen eine granularere Differenzierung zwischen Statistischer Protokollidentifikation und Deep Packet Inspection möglich wäre. Derzeit kann es im Fall einer allein auf Pakettimings und -größen abzielenden statistischen Protokollanalyse zu einer fehlerhaften Gleichbehandlung kommen.

Im Zuge der Entwicklung des Betriebssystem-, Software- und Browsermarktes hat sich eine Stärke mittlerweile in eine Schwachstelle gewandelt: Die Implementation als Java-Applet. Wurde Java noch 1999 als künftige Lösung der durch unterschiedliche Plattformen hervorgerufenen Probleme betrachtet (vgl. [51]), so verwenden heute immer weniger Nutzer einen Browser, der über eine Java-Ausführungsumgebung verfügt. Als Gründe sind wenigstens Sicherheitslücken und die Dominanz von Adobe Flash und HTML5/JavaScript zu erwägen; dies soll hier jedoch nicht weiter untersucht werden.

### NANO

NANO ist ein klassischer Vertreter der passiven Verfahren.

NANO ist gegenwärtig nur für Linux verfügbar, daher bleibt seine Anwendung auf einen relativ kleinen Nutzerkreis beschränkt; ferner bleibt ein systemimmanentes Datenschutzproblem bestehen: Eine zentrale Auswertung beim Georgia Institute of Technology ist erforderlich; dies bedeutet aber dass Informationen über die von einem Nutzer kontaktierten Gegenstellen weitergegeben werden müssen. Auch die Möglichkeit des temporären Deaktivierens des NANO-Agenten ist nur eine scheinbare Lösung – schafft sie doch einen Bereich von systematisch unüberwachten Verbindungen. Schließlich ist es plausi-

bel, anzunehmen, dass es einen gemeinsamen Schwerpunkt geben wird, zu welchen Zielen Verbindungen keine Betrachtung finden sollen: Zwielichtige Angebote. Angedacht seien die Webseiten, die Dateiaustausch oder Pornographie anbieten.

### Nooter

Bei Nooter als hybridem Verfahren sollen hier die Vor- und Nachteile besonders differenziert betrachtet werden.

Die Vorteile des Verfahrens liegen vor allem im Verzicht auf eine zentrale Auswertungsstelle, da es für jeden einzelnen Anwender – so die Annahme – zwei unabhängige Datenkanäle gibt, deren Charakteristik im neutralen Fall idealerweise[10] kaum voneinander abweicht. Ferner bezieht sich die bei der Auswertung getroffene Aussage nicht nur auf einzelne Testseiten, sondern auf vom Anwender tatsächlich genutzte Angebote. Auswuchs eines solchen Nutzerverhaltens wäre das Aufbauen eines verschlüsselten Tunnels zum Nooter-Broker in jedem Fall nebst Erzeugung eines gewissen Datenverkehrs in dieser Verbindung mit dem Ziel den Provider zum Verzicht auf nichtneutrale Policies zu zwingen. Das Dilemma des Providers wäre seine Unkenntnis, welche unverschlüsselten Datenpakete des Nutzers zu einem Nooter-Test gehören – und welche nicht[11].

Die Nachteile dieses Verfahrens sind vornehmlich durch den Einsatz einer Gegenstelle im Providernetzwerk bedingt: Damit das Messverfahren funktionieren kann, muss diese Gegenstelle Datenpakete mit beliebigen Absenderadressen in das Providernetzwerk versenden können; eine solche Praxis kann grundsätzlich als risikoreich eingeschätzt werden, da eine Maschine mit dieser Fähigkeit in der Lage ist, auf Ebene der IP-Adresse beliebige andere Identitäten anzunehmen, ohne dass dies vom Gegenüber mit einfachen Mitteln nachgewiesen werden kann. Bei der Gegenstelle im Providernetz bestehen die selben Datenschutzprobleme, die auch schon beim passiven Ansatz auftreten, wenn dieser umfassend verwendet wird (um die gesamte verwendete Netzwerkumgebung auf Neutralitätsverstöße zu untersuchen). Ferner wird die Annahme getroffen, dass es dem Internetzugangsprovider nicht möglich sein wird, Rückschlüsse auf die im verschlüsselten Kanal ausgetauschten Daten zu ziehen – fraglich ist, in wie weit diese Annahme angesichts des Fortschritts auf dem Gebiet der statistischen Protokollanalyse haltbar ist. Gegenwärtig werden jedoch [31] Verfahren entwickelt, welche auch den Inhalt von mittels TLS bzw. SSL geschützten Datenkanälen mittels statistische Protokollidentifikation erkennen.

Weitere Nachteile ergeben sich aus der Positionierung des Nooter-Brokers in komplexeren Netzwerktopologien, die die Aussagekraft einschränken können.

[10] abgesehen von Performance-Einbußen durch die Verschlüsselung

[11] Ein Problem, welches sich wahrscheinlich aus Providersicht mit SPID (statistischer Protokollidentifikation) adressieren ließe; es brächte dem Netzwerkprovider allerdings nur geringen Nutzen im Vergleich zum notwendigen Aufwand um Nooter-Nutzer hinreichend sicher zu erkennen.

#### Shaperprobe

Konstruktionsbedingt ist Shaperprobe nur zum Auffinden von Trafficshapern in der Lage, welche dem beschriebenen Modell entsprechen. Folglich ist ein Umgehen einer Prüfung eines Netzwerks mittels Shaperprobe durch Wechsel des Shapingalgorithmus leicht zu bewerkstelligen – in einem spontan durchgeführten Testlauf war es Shaperprobe nicht möglich, einen mit dem Toolkit Dummynet [52] aufgebauten Trafficshaper zu erkennen (vgl. Anhang D); stattdessen wurde die Abwesenheit eines Shapers gemeldet.

Das Ergebnis eines Testlaufs mit Shaperprobe ist nur im Fall eines Shapers, der sämtliche Daten (oder die von Shaperprobe verwendeten Testdaten) beeinflusst aussagekräftig; in allen anderen Fällen ist es der Test eines (ggf. nicht betroffenen) Spezialfalls.

## 4.3. Einordnung rechtlicher Regelungen

In diesem Abschnitt werden die im vorangehenden Kapitel beschriebenen gesetzlichen Regelungen und Regelungsvorschläge eingeordnet und bewertet.

### 4.3.1. Regelung der Netzneutralität in § 41a TKG

Wie in der Abgrenzung und auch in der Rekapitulation der öffentlichen Debatte geschildert, sind Forderungen nach einer gesetzlichen Regelung der Netzneutralität nicht neu (oder überraschend). Weltweit wurde solchen Forderungen in einigen Staaten bereits gefolgt. Die gesetzlichen Regelungen unterscheiden sich jedoch erheblich. So findet sich in Chile beispielsweise eine gesetzliche Festschreibung der Verpflichtung von Netzbetreibern zur Neutralität, während in Großbritannien mit dem Beschluss, dass es keinen Bedarf einer Regelung gäbe, de facto das Gegenteil umgesetzt wurde. Für einen ausführlichen Überblick sei auf [2], Rn. 15ff. verwiesen; eine tabellarische Zusammenfassung findet sich im Anhang E.
Dieser Abschnitt wird sich im Folgenden mit zwei konkreten Ansätzen befassen: Den mit Wirkung zum 10.05.2012 ins TKG aufgenommenen §§ 41a, 43a. § 41a gibt einen Rahmen für eine konkrete Regelung vor. Auch der Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Regelung der Netzneutralität[12] aus dem September 2013 wird diskutiert.

Ein im Sommer 2013 vom Bundeswirtschaftsministerium vorgelegter Entwurf einer Verordnung nach § 41a TKG wird in dieser Arbeit nicht reflektiert – er wird als Teil des Bundestagswahlkampfes 2013 und als in diesem Licht zu betrachtende Reaktion auf Ankündigungen der deutschen Telekom zur Kontingentierung ihrer Flatrate-Tarife betrachtet. Seitdem ist kein neuer Entwurf vorgelegt worden.

Aus informatischer Sicht sind drei Aspekte des § 41 bemerkenswert: In (1) wird eine Verordnung umrissen, die

[12] Dokument COM(2013) 627 final

- willkürlicher Verschlechterung von Diensten und
- ungerechtfertigter Behinderung oder Verlangsamung des Datenverkehrs

adressieren soll; mit (2) wird der Bundesnetzagentur die Möglichkeit gegeben, in Zusammenarbeit mit den europäischen Gremien

- Mindestanforderungen der Dienstqualität

festzulegen.

Aus technischer Sicht klafft zwischen dieser Regelung und einer Neutralitätsdefinition wie der von Bullinger eine weite Lücke. Auch das Paradigma des Best-Effort als grundlegendem Prinzip findet sich hier nicht wieder. Stattdessen wird Traffic-Engineering als (wohl nicht mehr zurückzunehmender) Teil heutiger Internetinfrastruktur akzeptiert und auf die Motivation einzelner Beeinflussungen von Datenübertragungen abgestellt: Sie sollen nicht **willkürlich** oder **ungerechtfertigt** sein; hierzu wurde im Vorfeld die Differenzierung zwischen „Ungleichbehandlung“ und „Diskriminierung“ untersucht [53], wobei eine Ungleichbehandlung als aus technischem Grund hinzunehmen angenommen wird; eine Diskriminierung hingegen als ungerechtfertigt angesehen.

Als (objektiv) gerechtfertigt und als nicht willkürlich wird ein Einsatz von Policies nur dann gelten, wenn er nicht der Privilegierung einzelner Dienste (nach entsprechender Gegenleistung) oder der Unterdrückung bestimmter Inhalte (z. B. Konkurrenz) dient. Eine solche Negativabgrenzung verbietet das Angebot von Diensten mit zugesicherten höheren Dienstegüten nicht[13] – die Regelung versucht sicherzustellen, dass sich hieraus keine negativen Rückwirkungen auf die übrigen Internetverbindungen ergeben. Die Regelung trägt damit auch den im Abschnitt 4.1 geschilderten Problemen des faktisch existenten Angebots von Managed Services[14] Rechnung.

Es kann von einer Formalisierung der (eben nicht hierdurch eingeführten sondern schon mit der Einführung bestimmter Dienste begonnenen) Trennung der nach Best-Effort organisierten Paketweiterleitung von Managed Services ausgegangen werden. Bemerkenswert ist der Vorstoß in so weit, als dass er sich des Bereichs des „Best-Effort-Internets“ annimmt und versucht, diesen zu schützen – damit dieser nicht dem wirtschaftlichen Druck geopfert wird und zur „Dirt-Road“ des Internets verkommt [33]. Gemeint ist hiermit eine Aufteilung der Netze in einen großen privilegierten Bereich und den „Rest“, der nach dem Best-Effort-Prinzip organisiert wird. Die Folge wäre, dass ein Nutzer, der keine privilegierten Dienste in Anspruch nimmt, mit einem unzuverlässigen Internet von schwankend niedriger Qualität vorlieb nehmen müsste.

Flankierend trat am 10.05.2012 mit § 43a eine Regelung in Kraft, die eine Offenlegung bestimmter Traffic-Engineering-Praktiken verlangt. Konkret wird vom Gesetzgeber

[13] es kann als hinzunehmende Ungleichbehandlung auftreten
[14] wie VoIP oder IP-TV für den Endkunden

gefordert, dass dem Nutzer Informationen über Einschränkungen des Zugangs zu einzelnen Diensten oder der Nutzung von Anwendungen, die Deklaration eines Mindestniveaus der Dienstqualität und Informationen über Traffic-Engineering-/Monitoring-Regeln zugänglich zu machen sind. Dabei werden die Informationen über Traffic-Engineering-/Monitoring-Regeln auf solche eingeschränkt, die zur Vermeidung von Aus- oder Überlastung von Netzverbindungen genutzt werden. Dies lässt im Rahmen des § 43a einen Spielraum für Maßnahmen, die nicht diese Ziele haben. Dieser Spielraum kann allerdings durch eine Verordnung nach § 41a geschlossen werden, da diese Traffic-Engineering ohne eine objektive technische Notwendigkeit zulasten des Nutzers adressiert.

§ 43a (3) delegiert die Konkretisierung der von den Netzbetreibern zu leistenden Angaben an die BNetzA, die entsprechende Vorgaben nach Beteiligung der betroffenen Verbände und Unternehmen in einem Amtsblatt erlassen kann. Daneben ist die BNetzA ermächtigt, Netzbetreiber zur Entwicklung von Hilfsmitteln zu verpflichten, mit deren Hilfe eigenständige Messungen unternommen werden können.

### 4.3.2. EU-Verordnungsentwurf

Insbesondere die Transparenzregelug des Verordnungsentwurfs leidet an bekannten Schwächen:
Obgleich der Nutzer Kenntnis über die Gründe nehmen kann, finden sich hier zwei Regelungen, die zum einem dem Netzbetreiber und zum anderen dem Staat Werkzeuge zur Blockade von Diensten und Anwendungen bieten. Die Entscheidung über Integrität und Sicherheit des Netzes wird sich praktisch immer erst nach einer Blockade prüfen lassen – und auch eine solche Prüfung wäre wesentlich von kaum nachprüfbaren Angaben des Netzbetreibers abhängig. Maßnahmen zur Abwehr oder Verhinderung von Verbrechen sind regelmäßig erst nach ihrer Installation (z. B. durch Anrufung eines Gerichts oder einer Aufsichtsbehörde) prüfbar.

### 4.3.3. Personenbeziehbarkeit von IP-Adressen

Die ablehnende Position von Krüger und Maucher ist vor aktuellen Entwicklungen kritisch zu hinterfragen, da die Argumentation ausschließlich auf der Annahme beruht, dass es neben den Zugangsanbietern keine weiteren Instanzen gäbe, die zu einer entsprechenden Zuordnung in der Lage wären. Diese Annahme übersieht allerdings die Existenz weiterer Betreiber, die zu einer solchen Zuordnung in der Lage sind[15]. Als Gegenbeispiel sei hier explizit auf das soziale Netzwerk „Facebook“ und auf diverse Anbieter anderer Dienste verwiesen.

„Facebook“ beispielsweise forciert seit langem durch einen „Klarnamenszwang“[16] eine Identifikation seiner Nutzer. Ähnliches gilt für Anbieter anderer Dienste, bei denen Nutzer

[15] Wiewohl nur beim ISP von vorliegender Zuordnung für **alle** von ihm vergebenen IP-Adressen ausgegangen werden kann.
[16] AGB-Klausel und automatisierte Befragung von „Freunden“ vermuteter Pseudonymträger

Daten hinterlegen, die auf ihre Person schließen lassen (Bestellsysteme aber auch Mail-Accounts). All diese Anbieter haben das notwendige Wissen, um die Zuordnung einer IP-Adresse zu einer natürlichen Person durchzuführen. Eine entsprechende Dienstleistung kann in einem Drittland stattfinden und muss nicht dem deutschen Datenschutzrecht unterliegen.

Die Möglichkeiten der Anbieter entsprechender Dienste können als hinreichend vorhanden angesehen werden, entsprechend liegt der Schluss einer Personenbeziehbarkeit der IP-Adresse nahe.

Dass Dienste zur Zuordnung von IP-Adressen zu Namen natürlicher Personen dennoch nicht als Dienstleistung verfügbar sind[17], erscheint nach Beobachtungen aus dem Marketing plausibel, die zeigen, dass eine subtile Werbung ohne direkte Adressierung des Nutzers effektiver ist als eine Werbung, die dem Nutzer zu verstehen gibt, wie viel tatsächlich über ihn bekannt ist[18] [77]. Entsprechend plausibel scheint es, davon auszugehen, dass entsprechende Schattenprofile von Nutzern auch mit Klarnamen gekennzeichnet sind – auch wenn diese nicht angezeigt werden.

## 4.4. Zusammenfassung

Es existieren drei grundlegende Herangehensweisen bezüglich des Nachweises von Neutralitätsverletzungen: Der aktive Ansatz, bei dem ein eingegrenzter Test bestimmte Eigenschaften des verwendeten Netzwerks prüft und auch nur hierüber Aussagen treffen kann; der passive Ansatz, bei dem die Beobachtung des Netzwerks während der alltäglichen Nutzung im Vordergrund steht und der hybride Ansatz, bei dem aktive und passive Komponenten kombiniert werden. Gemeinsam ist allen Ansätzen das Ziel, die Abhängigkeit einer Qualität des beobachteten Netzes von den ausgetauschten Daten nachzuweisen.

Tabelle 4.1 fasst einige wesentliche Punkte der Bewertung nochmals zusammen.

| | aktive Verfahren | passive Verfahren | hybride Verfahren |
|---|---|---|---|
| Vorgestellte Vertreter | Glasnost, Shaperprobe | Nano | Nooter, Herdict[19] |
| Vorteile | einfache Implementation, Testserver kann Daten liefern | umfassende Aussage, keine Testserver | implementationsabhängig |
| Nachteile | Beschränkte Aussage | Datenschutz | implementationsabhängig |

Tabelle 4.1.: Zusammenfassung der betrachteten Nachweisverfahren

[17] oder dass die Existenz derartiger Dienste nicht bekannt ist

[18] Extrembeispiel wäre eine Bannerwerbung für eine Seite mit erotischen Inhalten, die den Nutzer unter Verweis auf seine sexuellen Vorlieben beim Namen anspricht.

[19] je nach Nutzerverhalten und -kenntnissen werden bei Herdict teilweise rein beobachtende Einträge,

teilweise auf expliziten Tests basierende Einträge vorgenommen. Eine automatisierte Testung von Einträgen ist nicht vorgesehen.

# 5. Von Neutralität zu Transparenz von Netzwerken

> Derjenige, der sich mit Einsicht für beschränkt erklärt, ist der Vollkommenheit am nächsten.
> —*Johann Wolfgang v. Goethe*

Der bis hier vorgestellte und systematisierte Forschungsstand betrachtet Netzneutralität als eine binäre Eigenschaft: Ein Netz kann neutral sein – oder nicht. An dieser Stelle wird die herkömmliche Betrachtung der Netzneutralität verlassen: Es wird stattdessen davon ausgegangen, dass Netzbetreiber zu einer kompletten und nachprüfbaren Deklaration sämtlicher Netzpolicies verpflichtet wären, vergleichbar einem Beipackzettel. Dieses Kapitel beschreibt die Spezifikation, eine mögliche Überprüfung angegebener Spezifikationen und die von diesem neuen Blickwinkel ausgehenden Implikationen.

*Im Fall des ersten Einführungsbeispiels (S. 17) hätte eine Spezifikation der Netzwerkeigenschaften eine Iteration der Ursachensuche erspart: An die Stelle des Kontaktierens des Providers und eigenständiger Tests wäre der Blick in die Spezifikation (und ggf. ein Test auf von der Spezifikation abweichendes Verhalten) getreten.*

*Im zweiten Einführungsbeispiel hätte es ebenfalls die Suche nach der Fehlerursache deutlich vereinfacht, wenn Netzbetreiber zu einer einheitlichen und maschinenlesbaren Deklaration der von ihnen angewandten Policies verpflichtet gewesen wären. In diesem Fall hätte lediglich nach einem Regelsatz gesucht werden müssen, der die Unterbrechung eines Datenaustauschs als Reaktion auf die zu diesem Zeitpunkt bereits identifizierte Bytefolge vorsieht. Idealerweise wäre ein solches Durchsuchen von Regelsätzen automatisiert möglich gewesen.*

*Auch die praktische Analyse wäre um ein Vielfaches schneller gewesen, hätte jedes genutzte Teilnetz eine Gegenstelle enthalten, von der aus die Bytesequenz testweise heruntergeladen werden hätte können.*

**„Netztransparenz“ sei definiert als Bekanntheit der Traffic-Engineering-Regeln eines Netzwerks**[1]; dieser Begriff wird im Rahmen dieser Arbeit für die Zusammenstellung von Eigenschaft von Netzwerken verwendet, die eine Auswirkung auf den Nutzer besitzen. Der Begriff wurde definitionsweise bereits durch Publikation in [10] international eingeführt.

[1] im Sinne einer „Durchschaubarkeit“ des Vorgehens des Netzwerkbetreibers

Transparenz und Neutralität von Netzwerken stehen in einem engen Verhältnis, da Transparenz stets auch Information über Neutralität impliziert, nicht jedoch das Vorhandensein von Neutralität.

Die Verwendung des Begriffs der Netztransparenz verdeutlicht einen zu Grunde liegenden Perspektivwechsel: Weniger die (dogmatische) Frage nach der absoluten Neutralität eines Netzwerkes als das Wissen um die Neutralität soll im Fokus des Interesses stehen. Im Sinne der unternehmerischen und vertragsgestalterischen Freiheit sei es Netzzugangsanbietern und -betreibern dabei komplett überlassen, beliebige Traffic Engineering Rules festzulegen. Wichtig erscheint im Interesse des Nutzers primär eine präzise und nachprüfbare wie auch verständliche Angabe und Möglichkeit zur Überprüfung dieser veröffentlichten Netzwerkcharakteristika.

Dazu wird zunächst die Reduktion der Komplexität der Routingpfade beschrieben; im Anschluss hieran werden die Anforderungen an eine Sprache zur einheitlichen Spezifikation von Netzwerkpolicies skizziert. Danach wird der praktische Einsatz dieses Protokolls mit den notwendigen Schnittstellen seitens der Netzbetreiber wie auch der Nutzer betrachtet. Der weitere Teil dieses Kapitels widmet sich der Überprüfung der von Betreibern angegebenen Policies mithilfe in Betreibernetzen integrierter Broker. Schließlich werden die Verträglichkeit bzw. Unterstützung der hier vorgeschlagenen Methoden durch Vorgaben des TKG bzw. des EU-Verordnungsentwurfs untersucht.

## 5.1. Reduktion des Routingpfades

Weder Nutzern noch Betreibern von Netzwerken wäre gedient, wenn in einer Spezifikation das Verhalten jedes einzelnen Routers offenzulegen wäre. Im Gegensatz hierzu ist die Zusammenfassung von Teilnetzwerken zu „Black Boxes“, deren Verhalten spezifiziert wird, eine sinnvolle Abstraktion, wenn die Zusammenfassung entsprechend der Zuordnung zu Betreibern erfolgt. Eine solche Zuordnung in organisatorische Einheiten dürfte auch dem Nutzerinteresse entsprechen, denn selbst wenn es gelänge, einzelne Beschränkungen immer zuverlässig einem Router[2] zuzuordnen, wäre das Interesse des Nutzers nicht, welches technische Gerät eine Beschränkung verursacht – sondern von welchem Betreiber die Beschränkung veranlasst wurde.

Abb. 5.1 illustriert die damit verbundene Zusammenfassung der Policies: Der Endnutzer-PC kommuniziert mit einem Ziel. Die Datenpakete passieren die Router 1 und 2 des Netzbetreibers 1 sowie Router 3, 4 und 5 als Infrastruktur von Netzbetreiber 2. Router 1 begrenzt den Durchsatz auf maximal 1 MBit/s; Router 2 lässt keine Datenpakete des Protokolls P passieren. Router 3 begrenzt Datenpakete des Protokolls Q auf einen Maximaldurchsatz von 5 MBit/s; Router 4 und 5 arbeiten nach dem Best-Effort-Prinzip.

[2] oder der Verbindung zweier Router

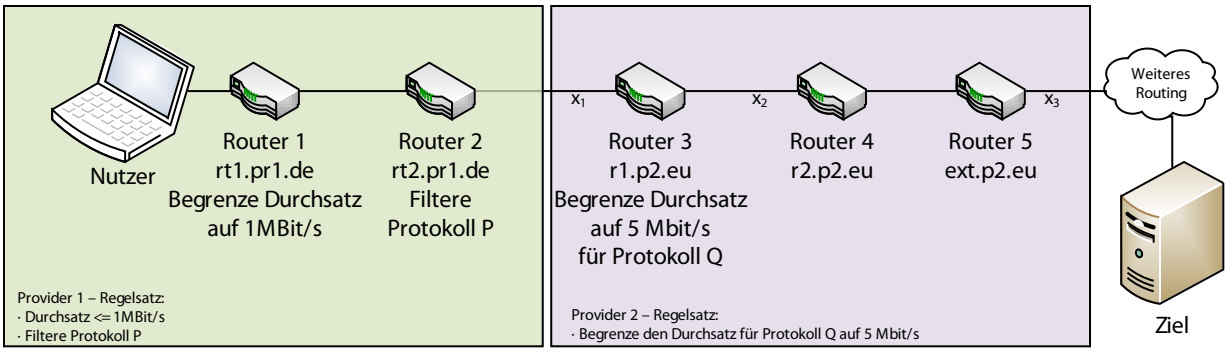


Abbildung 5.1.: Zusammenfassung von Netzwerken, der Nutzer wird auf Grund seiner öffentlichen IP-Adresse dem Providernetzwerk zugeordnet; ein mögliches privates Netzwerk des Nutzers ist hierdurch nicht ausgeschlossen.

Diese Regeln einzelner Router[3] lassen sich nach Providern zusammenfassen. Nach einer solchen Betrachtung ist es allerdings nicht mehr möglich, Effekte diskreten Geräten oder Verbindungen zuzuordnen. Betrachtet wird nur noch die Summe der von einem Provider verursachten Effekte; die einzelnen Summanden (einzelne Shaper) lassen sich nicht mehr identifizieren. Nach dieser Abstraktion bleibt von der Anordnung in Abb. 5.1 folgendes: Provider 1 begrenzt den Durchsatz auf 1 MBit/s und filtert das Protokoll P heraus; Provider 2 begrenzt Datenflüsse mit dem Protokoll Q auf maximal 5 MBit/s.

Eine solche Gruppierung von Routingknoten ist in der Praxis durch Nutzung von Routen-Traces als Ausgangsdaten durchführbar, wobei aus einem Routing-Pfad von vielen Elementen ein Pfad mit deutlich weniger Elementen entsteht. Eine Vereinfachung ist ferner pragmatisch für die nachfolgende Betrachtung, da die vom Punkt des Anwenders aus zu gewinnenden Routing-Informationen auf Grund moderner Netzwerkoptimierungen[4] nicht mehr verlässlich sein müssen (vgl. [26]).

In der Praxis stellt sich nach einem Routen-Trace die Frage der Zuordnung einzelner Knoten zu organisatorischen Einheiten. Es sind mehrere Lösungswege denkbar. Grundsätzlich muss eine Zuordnung jedoch nur die folgenden Kriterien erfüllen:

**Kriterium 1** Eine Zuordnung einzelner IP-Adressen zu organisatorischen Entitäten (z. B. Betreibern; es sind aber auch Konzepte wie eine Zuordnung nach Staatsgrenzen vorstellbar) muss gegeben sein,

**Kriterium 2** vollständige Abdeckung der im Routingpfad vorkommenden IP-Adressen,

[3] Netzwerkpolicies können natürlich auch von Geräten, die unterhalb des Netzwerk-Layers im ISO-/OSI-Modell angesiedelt sind realisiert werden; in dieser Arbeit werden die Effekte jedoch stets Routern als kleinste vom Endnutzer aus nachvollziehbare Infrastrukturentität zugeordnet.

[4] insbesondere innerhalb einzelner Netzwerke

**Kriterium 3** aus der Zuordnung muss sich der Ort (URL) der Deklaration der Netzwerkeigenschaften durch den Betreiber des jeweiligen Teilnetzes ableiten lassen.

Kriterium 2 kann wiederum auf zwei Arten variiert werden:

**Variation 2.1** Es ist ausreichend, wenn eine Zuordnung für die beiden ein Netz begrenzenden Router verfügbar ist[5] – in Abb. 5.1 wäre eine Zuordnung der IP von Router 4 verzichtbar, da auf seine Zugehörigkeit zu Provider 2 geschlossen werden könnte, wenn Informationen zu Router 3 und 5 vorliegen.

**Variation 2.2** Wenn die Begrenzung der Teilrouten vernachlässigt werden soll (wogegen prinzipiell nichts spricht), reicht es aus, wenn pro genutztem Netzwerk wenigstens ein Router eine Zuordnung besitzt.

Eine Zusammenfassung von Routing-Knoten ist anhand mehrerer Kriterien durchführbar, von denen hier zwei betrachtet werden sollen: Eine ist die Zuordnung von IP-Adressen zu autonomen Systemen (AS); eine andere Möglichkeit ist die Nutzung der inversen DNS-Auflösung der IP-Adressen der Routingknoten. Vorteil einer Zuordnung anhand inverser DNS-Auflösungen ist, dass diese meist mit betriebssystemeigenen Mitteln möglich ist; weiterer Vorteil ist das sofortige Vorliegen eines Domainnamens des Netzwerkbetreibers, der zum Bezug weiterer Informationen verwendet werden kann. Nachteil der Zuordnung nach Reverse-DNS-Auflösung ist die Möglichkeit fehlender Einträge und die Frage des Umgangs mit entsprechenden Routingpunkten. Die Zusammenfassung anhand der Zuordnung der IP-Adressen von Routingknoten zu AS hat den Vorteil kompletter Auflösung (jede vergebene IP-Adresse ist eindeutig einem AS zugeordnet); nachteilig ist hierbei die Notwendigkeit des Zugriffs auf eine entsprechende Zuordnungsliste sowie deren Aktualisierungen.

Aufgrund einer leichteren Umsetzbarkeit wird im Folgenden eine Zuordnung anhand einer Domainnamensauflösung angenommen. Dies ist als reiner Vorschlag zu verstehen und für die Kernfunktionalität nicht erheblich; die Gruppierung muss lediglich oben genannte Anforderungen erfüllen.

Die als Kriterium 3 genannte Verknüpfung mit weiteren Informationen zur Verbindung des identifizierten Netzbetreibers im Sinne einer Black Box mit dem von ihm veröffentlichten Policy-Satz muss nicht notwendigerweise in einem Schritt mit der Zusammenfassung der Netze ermittelt werden und könnte auch durch weitere Quellen zugeordnet sein. Die Vereinigung diese Schritte vereinfacht eine Umsetzung allerdings, so dass sie als obligatorisch angesehen werden sollte.

Im kommenden Abschnitt wird auf die Anforderungen an die Spezifikation der Netzwerkeigenschaften eingegangen; die Verwendung sowohl der hier vorgestellten Zusammenfassung von Netzen wie auch der Policy-Spezifikation wird im übernächsten Kapitel

[5] Wenn das Netz komplett durchquert wird; ferner bezieht sich die Angabe von zwei Routern auf die konkrete Verbindung, für die das Netzwerk (weder des ISPs noch des angefragten Gegenpeers) einen Eintritts- und einen Ausgangspunkt hat; für unterschiedliche Verbindungen können diese variieren.

geschildert.

## 5.2. Beschreibung von Netzwerkeigenschaften

Dieser Abschnitt beschreibt Anforderungen an eine geeignete formale Sprache zur Spezifikation von Netzwerkeigenschaften auf einem vom konkreten Netzwerk abstrahierten Niveau. Ziel dieses Kapitels ist das Skizzieren von Anforderungen an eine über Betreibergrenzen hinweg nutzbare Spezifikationssprache. Ziel ist im Rahmen dieser Arbeit die Plausibilität einer solchen „Beipackzettelsprache“ aufzuzeigen; eine formale Untersuchung und Implementation wäre ein Feld über diese Arbeit hinausgehender Forschung.

Ziel einer solchen „Beipackzettelsprache“ selbst ist die Beschreibung des Verhaltens von Netzwerken, wie es vom Nutzer beobachtet werden kann.

Durch die Zielstellung leidet eine Sprache mit diesem Aufgabenfeld jedoch bereits an einem Dilemma: Die enthaltenen Angaben sind als Beschreibung eines Verhaltensmodells eines Netzwerks zu verstehen – es gibt also sowohl Aussagen, die ein bestehendes System beschreiben und deren Zutreffen maschinell getestet werden kann – wie auch weitere Informationen für den Nutzer bzgl. der Motivation bestimmter Regeln[6]. Bei Tests handelt es sich um Versuche mit Datenübertragungen, deren Ziel stets ist, einen statistischen Nachweis für die Nichteinhaltung der Spezifikation zu finden.

Gleichzeitig stellt sich natürlich die Frage, wie einem solchen Nachweis im Erfolgsfall umzugehen ist (sind doch einige Anforderungen von Betreibern leicht sicherzustellen und andere nur mit hohem Aufwand) – diese Fragestellung wird im Zusammenhang mit der Überprüfung deklarierter Policies weiter vertieft.

In diesem „virtuellen Beipackzettel“ zu einem Netzwerk sind die Einflüsse auf den Datenaustausch abzubilden, die den Nutzer beeinträchtigen; hinzu kommen Angaben, die für den menschlichen Nutzer interessant sind, weil sie bestimmte Einflüsse erklären. Es ergeben sich somit drei wesentliche Kategorien von Informationen:

- Beeinflussung (Diskriminierung oder Privilegierung) von Datenaustausch (unabhängig von eingesetzten Kriterien oder Identifikationstechniken),
- Kontingentbeschränkungen (z. B. aktueller Verbrauch des gebuchten monatlichen Datenvolumens im Bereich mobiler Internetanschlüsse),
- sowie weitere vom Gesetzgeber vorgesehene Informationen bezüglich einzelner Verbindungen (z. B. temporäre Beschränkungen zum Schutz der Arbeitsfähigkeit des Netzwerks durch Filterungen im Zuge der Abwehr eines DDoS-Angriffs).

[6] z. B. weil dies gesetzlich vorgeschrieben ist, etwa Kontingentinformationen oder Sperren

Die Spezifikation betrachtet einzelne Netzbetreiber als Black Boxes; sie enthält keine konkreten Konfigurationsanweisungen oder technische Details, wie eine bestimmte Restriktion erreicht wird. Stattdessen wird das Netzwerkverhalten hiervon abstrahiert dargestellt, wobei alle innerhalb der Black Box auftretenden Phänomene zusammenwirkend beschrieben sind (vgl. auch 5.1) und sich nicht mehr auf einzelne Ursachen (einzelne Geräte) zurückführen lassen müssen.

Aus den drei Kategorien ergeben sich drei Arten von Statements, die eine Spezifikation enthalten kann:

**Regeln** Der Beschreibung eines Kriteriums folgt die Beschreibung, wie mit vom Kriterium erfassten Daten verfahren wird. Dieser Teil beschreibt das bestehende System und wird ggf. automatisch getestet. Die Regeln sind als Implikation der Form `Kriterium` → `Policy` mit der Semantik „Wenn Datenpakete `Kriterium` erfüllen, dann gelten für diese die als `Policy` spezifizierten Policies“ notiert, wobei Kriterien sich auf einzelne Datenpakete oder den Datenstrom beziehen können (s.u.). Ein Test einer Regel entspricht somit der Evaluation, ob ⟦`Kriterium`→`Policy`⟧ wahr oder falsch ist, wobei sowohl das Kriterium wie auch die Policy vom Nutzer evaluiert werden können. Kriterien können entweder von einzelnen Datenpaketen oder vom Datenstrom als solchem erfüllt werden. Hierbei ergibt sich schon aus der Methodik des Testens, dass eine absolute Einteilung in „wahr“ oder „falsch“ kaum möglich sein wird; stattdessen ergibt sich aus jedem Test eine statistische Sicherheit, mit der eine solche Aussage belegt oder abgelehnt werden kann.

**Kontingente** Der Beschreibung eines Kriteriums folgt eine Angabe über Verbrauch und zeitliche Konditionen des verbleibenden Kontingents (enthalten ist insbesondere, welche Daten vom Kontingent erfasst sind).

**Annotationen** Es können weitere informationelle Statements definiert werden; auch diese sind jeweils durch Kriterien identifizierten Datenströmen zugeordnet. Diese Angabe erfolgt ebenso wie die Angabe zu Kontingenten ausschließlich zur Information des Nutzers über die Hintergründe einzelner Regeln.

Zentral ist offensichtlich die Rolle der Kriterien, durch die Regeln, Kontingente oder Annotationen mit ausgetauschten Daten verknüpft werden. Die Kriterien müssen die vom Provider durchgeführte Datenstromidentifikation abbilden können, bleiben hierbei aber abstrakter als z. B. die Konfigurationssprache einer Infrastrukturkomponente. Entsprechend sind ausschließlich IP-basierte Regeln unbrauchbar. Aus den möglichen Formen der Datenstromidentifikation ergeben sich die folgenden Kriterien:

**Bedingungslos** Dieses Kriterium wird von jedem Datenpaket erfüllt.

**Ziel-IP-Adressen** Dieses Kriterium bezieht sich auf einzelne Datenpakete und entspricht einer Identifikation mittels SPI[7].

[7] Shallow-Packet-Inspection, Analyse der Header inkl. Layer-4-Header.

**Zielport** Dieses Kriterium bezieht sich auf einzelne Datenpakete und entspricht einer Identifikation mittels SPI, kann allerdings auch mit Protokollbezug verwendet werden[8].

**Quell-IP-Adresse** Auch dieses Kriterium bezieht sich auf einzelne Datenpakete und entspricht ebenfalls SPI; für ISP-Netze ist die Quell-IP-Adresse mit dem `Bedingungslos`-Kriterium identisch.

**Quell-Port** Auch dieses Kriterium bezieht sich auf einzelne Datenpakete und entspricht ebenfalls einer SPI.

**Inhalt** Dieses Kriterium bezieht sich auf den gesamten Datenstrom, da Fragmentierung von Daten ansonsten zu Problemen führen würde; erkannt werden bestimmte Bytesequenzen (DPI[9]).

**Protokoll** Identifikation eines bestimmten Protokolls; für die Spezifikation ist unerheblich, wie genau das Protokoll durch den Provider identifiziert wird, ob per Zielport, per DPI oder per SPID[10]. Hier sind zunächst Layer 5+-Protokolle gemeint; Tunnelkonstrukte (ssh über HTTP) stellen ein Problem dar – werden hier aber bewusst als Anwendung des „äußeren“ Protokolls betrachtet[11].

**Quota verbraucht** Kriterium für eine Regel, die beschreibt, wie Datenpakete gehandhabt werden, für die eine Kontingentierung erreicht worden ist. Dieses Kriterium kann nicht für die Spezifikation von Kontingenten verwendet werden (Ausschluss von Zyklen in der Spezifikation, da ansonsten die Aussage „Ein Datenstrom, dessen Quota verbraucht ist wird wie ein Datenstrom behandelt, dessen Quota verbraucht ist“ möglich wäre).

Die Kriterien sind mittels and- und or-Operatoren kombinierbar; hinzu kommt ein not-Operator, mit dem Kriterien negiert werden können, da dies die Spezifikation bestimmter Datenströme massiv vereinfacht, wie in den untenstehenden Beispielen verdeutlicht.

Die spezifischere Beschreibung hat eine höhere Priorität als die unspezifischere[12].

Die technische Realisation der Kriterien ist für die Spezifikation unerheblich, dies gilt insbesondere für Identifikation von Protokollen. Das Kriterium „Quota verbraucht“ erlaubt die einfache Zusammenführung der Regeln für verbrauchte Kontingente. Die Frage, ob ein Datenstrom (oder -paket) ein Kriterium erfüllt, ist zu einem bestimmten Zeitpunkt bei Kenntnis des Datenstroms stets eindeutig zu beantworten.

[8] Die Zielports 80 oder 443 können i.A. mit HTTP(s) als Protokoll belegt angenommen werden.

[9] Deep-Packet-Inspection, Analyse des gesamten Datenpaketes.

[10] Statistische Protokollidentifikation

[11] Dieses Vorgehen ist deshalb pragmatisch, da es ein Abbruchkriterium wieviele Tunnel-Lagen durchblickt werden müssen vermeidet.

[12] Die Spezifizität könnte sich z. B. am ISO-/OSI-Netzwerkmodell und nachfolgend dem Offset innerhalb eines Datenpaketes orientieren. Entsprechend wäre die Präzedenz von Kriterien (in aufsteigender Reihenfolge etwa IP-Adressen, Ports, L4-Protokoll, L5-Protokoll, Inhalt).

Praktische Kombinationen zur Identifikation von Datenströmen könnten beispielsweise lauten:

- `(protocol=="HTTP") AND (NOT dstip==139.30.1.202)`
- `(protocol=="Skype")`
- `(quotafull) AND (protocol=="FTP")`
- `(quotafull)`
- `(dstport==23) AND (contains=="rm -rf /")`
- `(contains=="Der Pirol pfeift heute Nacht")`

Der Blick auf die Anforderungen liefert nun drei mögliche Konstrukte in denen Kriterien genutzt werden: Regel, Kontingentierung oder Annotation. Zunächst sollen Regeln betrachtet werden. Hierbei handelt es sich um nachprüfbare, von der technischen Umsetzung abstrahierte und als Implikationsfolge beschriebene Policybeschreibungen. Die Policybeschreibung ist von der technischen Umsetzung losgelöst – es ist unerheblich, wie eine bestimmte Policy technisch erreicht wird; relevant ist nur, dass bei einer Betrachtung des zusammengefassten Netzes der Nutzer wie spezifiziert beeinflusst wird.

Im Einzelnen ergeben sich die folgenden Policies:

**Best-Effort** Die Daten werden weder privilegiert noch diskriminiert (Nachweisproblem[13]).

**Herausfiltern eines Paketes** Es wird nur das vom Kriterium erfasste Datenpaket herausgefiltert.

**Herausfiltern des Stroms** Es werden das vom Kriterium erfasste Datenpaket und alle im Datenstrom hierauf folgenden Pakete herausgefiltert.

**Zusicherung eines Mindestdurchsatzes** Beschreibt eine Privilegierung durch den Netzbetreiber.

**Zusicherung einer bestimmten Höchstlatenz** Beschreibt eine Privilegierung durch den Netzbetreiber.

**Zusicherung geringen Jitters** Beschreibt eine Privilegierung durch den Netzbetreiber.

**Begrenzung des Durchsatzes** Beschreibt eine Diskriminierung durch den Netzbetreiber.

**Angabe einer bestimmten Mindestlatenz** Beschreibt eine Diskriminierung durch den Netzbetreiber.

[13] Hier stellt sich das elementare Nachweisproblem der Neutralität, s.o..

**Manipulation des Datenstroms** Beschreibt eine Veränderung des Paketinhalts aus Sicht von ISO/OSI-Modell-Layer 3 (IP). Ausgenommen sind vorgesehene Veränderungen von TTL und Header-Prüfsumme. Veränderungen werden durch Angabe von Position und einzusetzendem Byte bzw. einzusetzender Bytefolge beschrieben.

Eine Regel ergibt sich somit aus einem Kriterium und einer Policy. Bei einem Test wird versucht, die Anwendung der Policy anhand der Beobachtung der ausgetauschten Daten anzufechten; Ergebnis eines Tests ist eine statistische Sicherheit, mit der die Policy angezweifelt werden kann.

Die Einführungsbeispiele könnten etwa wie folgt repräsentiert werden:

- `(protocol=="Skype")`→`(throughput`≤`0.5MBit/s)`
- `(protocol=="HTTP") AND (contains==[0x00 0x00 0x95 0xDC 0x00 0x02])` → `(FilterStream)`
- `(contains==[0x00 0x00 0x95 0xDC 0x00 0x02])`→`(FilterStream)`

Ein neutrales Netz würde durch den minimalen Ausdruck

- `(unconditional)`→`(BestEffort)`

beschrieben werden.

Als nächstes sollen die genauen Angaben zu Kontingenten betrachtet werden. Da mehrere einem Anschluss zugeordnete Kontingente denkbar sind, müssen auch Kontingente stets eine Spezifikation der von ihnen betroffenen Daten enthalten. Diese Angaben sind in erster Linie zur Information des Nutzers vorhanden.

**Kontingent erschöpft** Das Kontingent ist aufgebraucht. In Regelsätzen gelten die Policies, die dem Kriterium „Quota verbraucht“ folgen.

**Gesamtkontingent** Angabe des für den Kontingentszeitraum zur Verfügung stehenden Gesamtkontingents.

**Verbrauchsbeginn** Angabe des Zeitpunkts zu dem der Kontingentszeitraum beginnt.

**Bisher verbraucht** Angabe des seit Kontingentszeitraum verbrauchten Anteils.

**Folge nach Verbrauch** Angabe nur, wenn das Kontingent nicht bereits erschöpft ist. Es sind zwei Werte denkbar: Einschränkung oder Zahlung.

**Kontingentfrei** Um einzelne Datenströme bei pauschaler Kontingentierung auszunehmen.

Bei Kontingenten sollte jedes Kontingent stets mit Angaben zum Gesamt- und verbrauchten Kontingent sowie den Angaben zu Verbrauchsbeginn und den Folgen des Aufbrauchens ausgestattet sein, so dass Folgen der Nutzung abschätzbar sind. Die Existenz einer Regel zur Kontingentfreiheit ist systematisch nicht notwendig; erleichtert die Darstellung von netzinternen Services allerdings sehr (andernfalls müssten die Kriterien jeweils entsprechend des ersten Beispiels für Kriterienbildung stets unter Verwendung der Negation deklariert werden).

Eine Angabe könnte in der Praxis für einen mobilen UMTS-Internetzugang etwa wie folgt aussehen:

- `(unconditional)→(Quota==5242880kByte)`
- `(unconditional)→(Quotastart=="2013-05-01")`
- `(unconditional)→(Quotausage==48460kByte)`
- `(quotafull)→(Measure==LimitThroughput)`

Wie die in der letzten Deklaration angedeutete Durchsatzbeschränkung aussieht, wird im Regelteil spezifiziert sein:

- `(quotafull)→(Throughput≤10kBit/s)`

Wenn es netzinterne Dienste gibt, die von Kontingenten ausgenommen sind, kann dies mithilfe eines weiteren Kontingentstatements angegeben werden:

- `(dstip==1.2.3.4) AND (protocol=="HTTP")→(NoQuota)`

Schließlich existieren noch Annotationen. Auch sie sind stets nur auf bestimmte Datenpakete bezogen, die über die bereits eingeführten Kriterien beschrieben werden und bestehen aus Zeichenfolgen (die in der Praxis aus pragmatischen Gründen[14] auf eine bestimmte Länge begrenzt sein können). Eine Annotation könnte wie folgt aussehen:

- `(protocol=="Skype")`
  `→(note=="Nur nach Abschluss des VoIP-Paketes verfuegbar.")`

Regelsätze beschreiben das konkrete Verhalten des Netzwerks für den anfragenden Nutzer. Wenn ein Netzbetreiber seine Konfiguration in Abhängigkeit vom jeweiligen Nutzer gestaltet, muss sich dies auch in der an den Nutzer zurückgegebenen Spezifikation wiederspiegeln. Zwei technische Umsetzungen sind denkbar: Zum einen das Erstellen einer globalen Spezifikation, in der sämtliche Nutzer mit ihren jeweils separat gültigen Einschränkungen durch ihre jeweilige Absender-IP als von der Regel betroffen oder nicht betroffen adressiert werden. Alternativ ist auch eine dynamische Erstellung der Spezifikation in Abhängigkeit vom Anfragesteller – also eine für jeden Zugriff neu und in

[14] Um die Menschenlesbarkeit der Spezifikation zu erhalten und nicht durch das Einbetten kompletter HTML-codierter Webseiten zu gefährden.

Abhängigkeit von der Anfrage-IP generierte Spezifikation – denkbar. Dieses Vorgehen ist allerdings inhärent nicht nachvollziehbar, da sich die Spezifikation der Netzpolicies somit je nach Anfrager unterscheiden könnte.

Kontingentinformationen hingegen müssen stets in Abhängigkeit von der Anfrager-IP erstellt werden.

Eine Trennung der Spezifikation in die universell gültigen und hinreichend abstrakten Netzwerkregeln mit den weiteren Annotations einerseits und die personenbezogenen Kontingentangaben andererseits bietet sich als Lösung an.

Als weiteres Kriterium sei hier noch die Menschenlesbarkeit der Spezifikation genannt; nicht zuletzt um eine Abschätzung des Netzwerkverhaltens auch ohne weitere Programme nur mit der Routinginformation, DNS-Anfragen und einem Browser abrufen zu können; mithin einer maximalen Plattformunabhängigkeit.

Für eine Nichtabstreitbarkeit der tatsächlichen Urheberschaft einer veröffentlichten Policy-Deklaration sind die entsprechenden Dokumente zu signieren. So ist auch der Weg für eine Archivierung und damit eine Langzeitbeobachtung gegeben – andernfalls wäre die Urheberschaft nur für die aktuelle vom Provider hinterlegte Spezifikation nachweisbar.

## 5.3. Realisation der Deklaration von Netzwerkeigenschaften

Dieser Abschnitt beschreibt die Voraussetzungen einer praktischen Umsetzung transparenter Policygestaltung. Eine Umsetzung muss zwei Voraussetzungen erfüllen: Die Abrufbarkeit von Routeninformationen und den einheitlichen Zugriff auf die Policydeklarationen aller Netzbetreiber.

Um zu wissen, welche Netzwerkbetreiber (und damit auch, welche Policies) an einem bestimmten Datenaustausch beteiligt sind, muss der Endnutzer den Routingpfad ermitteln können. Dies kann wahlweise durch den direkten Test (mittels `traceroute` oder an heutige Routing-Vorgehen angepasste Verfahren wie Paris Traceroute[26]) oder durch Auswertung globaler Routinginformationen ermittelt werden. Dabei ist die erstgenannte Alternative zu bevorzugen, da sie mit einem geringeren Aufwand Ergebnisse liefern kann[15]. Hierzu ist (z. B. durch gesetzliche Vorgaben, s.u.) sicherzustellen, dass es vom Nutzerstandpunkt aus möglich ist, die Route zumindest bezüglich der beteiligten Netzbetreiber zu ermitteln und diesen dann auch z. B. per Reverse-DNS einen Domainnamen zuzuordnen.

Die Spezifikation der Netzwerkeigenschaften ist per HTTP (und HTTPs) unter einer für alle Netzwerkbetreiber einheitlich aufgebauten URL abrufbar zu machen. Eine solche

[15] Wenn entsprechende Testpakete nicht von einem bösartigen Provider manipuliert werden.

URL könnte bei Auftrennung der unterschiedlichen Deklarationsteile beispielsweise die Formen

```
http://domainname.tld/netspec
http://domainname.tld/kontingentspec
http://domainname.tld/notespec
```

besitzen. Im Fall der Verwendung von HTTPs stellen sich Fragen der Zertifikatsvalidierung, insbesondere die Frage, wie mit nicht validierbaren Zertifikaten verfahren werden soll. Des Weiteren ist der Abruf einer solchen Spezifikationsdatei statusfrei, also jederzeit und ohne Vorbedingungen durchführbar zu implementieren. Nur der Abruf der Kontingentinformationen soll in Abhängigkeit von der Anfrager-IP erfolgen.

Um die Spezifikation des Gesamtpfades zu ermitteln, sind zunächst die beteiligten Teilnetze und für diese dann Spezifikationen zu ermitteln. Im Anschluss müssen die jeweils für den betrachteten Pfad relevanten Policies ermittelt werden. Im Anschluss kann eine Schätzung der Netzqualität für die Gesamtverbindung aufgestellt werden. Entsprechende Berechnungen sind nicht Teil dieser Arbeit[16].

Ein an dieser Stelle noch offenes Problem ist die Frage, wie mit Veränderungen umgegangen werden soll – speziell im Fall einer Congestion durch Nutzerverhalten oder aber durch einen (DDoS-)Angriff wird ein Netzwerkbetreiber kurzfristig andere Prioritäten als die Validität der veröffentlichten Spezifikationen haben. Vergleichbares gilt im Fall von durch Defekt oder Havarie notwendigen Änderungen. Nicht weniger problematisch kann in der Praxis auch ein schnelles Verändern von Routen auf Grund ökonomischer Motive sein.

Als Lösung bietet sich das Einräumen eines Zeitraumes an, innerhalb dessen die Policies vom Provider aktualisiert werden müssen. Um ein Einhalten dieses Zeitraums sicherzustellen und einen Missbrauch zu vermeiden, sind geeignete Verfahren zur Erzeugung eines authentifizierten Zeitstempels (wie beispielsweise in RFC 3161 u. 5816 [56, 57] spezifiziert) zu verwenden. Andernfalls würde sich ein bösartiger Netzbetreiber immer mit dem Verweis auf die Herausgabezeit einer Spezifikation und eventueller Ungültigkeit zum Testzeitpunkt vor Konsequenzen schützen können.

Bei einer ausschließlichen Deklaration von Netzwerkspezifikationen wird diese Deklaration nie den Status einer Behauptung verlassen können – schließlich ist es dem Nutzer nicht möglich, die Einhaltung der Spezifikation zu prüfen. Da sich aber Netzwerkeigenschaften verschiedener Netze überlagern, ist aus einer Beobachtung und den gegebenen Spezifikationen hingegen auch kein Rückschluss möglich, welcher konkrete Provider für

[16] Es gibt hierzu in der Literatur bereits hinreichend viele Arbeiten (exemplarisch sei [54] genannt) – verwiesen sei aber auch auf die Grenzen deterministischer statistischer Netzwerkmodelle wie beschrieben in [55]; dennoch wird im Anhang C die Nutzbarkeit einer überschlagsmäßigen Abschätzung für die Werte von Latenz und Durchsatz mit einfachen Mitteln für einen im folgenden Abschnitt beschriebenen Anwendungsfall demonstriert.

eine bestimmte Einschränkung verantwortlich ist.

*Am zweiten Einführungsbeispiel (S. 17): Hätten Martin und ich uns bei der Analyse der aus unklarem Grund nicht funktionierenden Verbindungen die Mühe gemacht, die zuständigen Netzbetreiber anzurufen, die die von uns auf dem Weg zu unserem Testserver passierten Netze betreiben und wir hätten von allen die Antwort „unser Netz macht so etwas nicht“ erhalten – es hätte uns nicht geholfen, denn das Phänomen war nachweisbar vorhanden. Genau so erging es dem Nutzer im ersten Einführungsbeispiel.*

Der folgende Abschnitt beschreibt eine Infrastruktur, die eine Überprüfung der deklarierten Netzwerkeigenschaften ermöglicht.

## 5.4. Überprüfung von Netzwerkeigenschaften

In diesem Abschnitt wird ein Verfahren vorgestellt, mit dem es möglich ist, eine den in den letzten Abschnitten beschriebenen Ansprüchen genügende Spezifiaktion von Netzeigenschaften zu überprüfen.

*Im zweiten Einführungsbeispiel war die Identifikation des „schuldigen[17]“ Teils der Infrastruktur nur deshalb möglich, weil die Bytesequenz in beliebigen Paketen die Filterung aller weiterer Pakete desselben Datenstroms[18] inklusive eventueller ICMP-Antworten provozierte. Dadurch konnte jeder Router als Test-Peer dienen. Ein vergleichbares Verfahren soll an dieser Stelle (in auf das Szenario angepasster Form) zur Überprüfung der von Netzwerkbetreibern deklarierten Policies genutzt werden.*

Zunächst werden im Folgenden die notwendigen Anpassungen der Infrastruktur beschrieben, dem folgt die Beschreibung des Testvorgehens; schließlich werden die Beschränkungen des Verfahrens beschrieben.

### 5.4.1. Modifikation und Einfluss der Infrastruktur

Um eine Spezifikation zu validieren, in der Netze als Black Boxes beschrieben sind, ist in jeder dieser Black Boxes eine Gegenstelle vonnöten. Um das von Kaminsky vorgeschlagene Verfahren (vgl. Kap. 3) einzusetzen, muss es sich hierbei um Broker mit einer Proxy-Funktionalität[19] handeln, die für den Kanal zum Nutzer eine Vergleichsmessung zwischen einem verschlüsselten und einem unverschlüsselten Kanal ermöglichen.

[17] im Sinne des Verursachens der beobachteten Phänomene; verantwortlich wird immer der Betreiber der Infrastruktur sein und es wird auch nur selten gelingen, eine Beobachtung einer konkreten Hardwarekomponente zuzuordnen

[18] identifiziert anhand von Absender- und Ziel-IP möglicherweise auch Absender- und Ziel-Port oder Layer-4-Protokoll

[19] Im Sinne einer Paketweiterleitung, hier allerdings ohne Kenntnis des Upper-Layer-Protokolls.

Dieser Broker stellt einen invasiven Eingriff in das umgebende Netzwerk dar, denn zur Verwendbarkeit gibt es zwei wesentliche Randbedingungen:

- Das Herstellen von Verbindungen zu beliebigen anderen Adressen muss möglich sein,
- das Absenden von Netzwerkpaketen mit beliebigen Absenderadressen muss möglich sein, wobei sich „beliebig“ einschränken lässt: Weiterzuleiten sind ausschließlich Datenpakete mit Absenderadressen, die auch unter „Normalbedingungen“ das Netzwerk kreuzen würden.

Während die erstgenannte Anforderung kaum einen Netzwerkbetreiber vor Probleme stellen dürfte, beinhaltet die zweite Anforderung ein (abschreckendes) Missbrauchspotential, denn eine Maschine, deren Netzwerkanbindung derart konfiguriert ist, kann als Gefahr betrachtet werden. Durch die Fähigkeit, Pakete mit beliebiger Absenderadresse zu versenden (die von der Infrastruktur auch weitergeleitet werden), können schlimmstenfalls[20] beliebige andere Dienste impersoniert werden. Hierbei handelt es sich um ein systemimmanentes Problem, denn zur vergleichenden Messung ist das Absenden von Datenpaketen mit der IP des ursprünglichen Target-Servers als Absenderadresse notwendig.

Bereits bei der Frage der Positionierung des Brokers entsteht allerdings ein Dilemma, denn entgegen der von Kaminsky getroffenen Annahmen ist es bei realen Netzen kaum möglich, eine Position eines Brokers „am äußeren Rand“ des Netzes sicherzustellen. Stattdessen kann maximal eine Positionierung in einem Netz erreicht werden. Ungünstigerweise verändert sich die Aussage einer Messung hierbei deutlich. Zur Veranschaulichung sei auf Abb. 5.1 (S. 89) verweisen. Dort sind mit $x_{1..3}$ drei mögliche Positionen eines Brokers im Netz von Provider 2 eingezeichnet. Die folgenden drei Punkte beschreiben die Aussage einer Messung, wenn sich ein Broker an diesen Positionen befände:

**Position $x_1$.** Der Broker befindet sich (vom Nutzer aus gesehen) am Beginn des Routingpfades im Netz von Provider 2. Messungen mit einem hier installierten Broker können ausschließlich zu Aussagen über das Verhalten vor Provider 1 führen. Diese Position ist für Provider 2 die günstigste – die in seiner Infrastruktur installierte Testmaschine kann nicht zu Aussagen gegen ihn führen; er hat folglich kein Interesse an einer Manipulation von Messungen von diesem Messknoten aus, da ihn betreffende Messungen nicht von diesem Messknoten ausgehen werden.

In Abb. 5.1 würden für Datenpakete, die zwischen dem Nutzer und einem Broker an der Position $x_1$ ausgetauscht werden, nur die vom Provider 1 aufgestellten Regeln (Begrenzung des Durchsatzes auf 1MBit/s und Herausfiltern des Protokolls P) gelten.

**Position $x_2$.** Der Broker ist umgeben von Routern, die unter dem Einfluss von Provider 2 stehen. Im Fall eines so positionierten Brokers vermischen sich die Effekte von Provider 1 mit denen von Provider 2. Bei einer Prüfung können Ergebnisse ermittelt

[20] z. B. bei gleichzeitiger entsprechender Kompromittierung von DNS-Servern

werden, die schlechter sind als nur nach dem Regelsatz von Provider 1 zu erwarten; als obere Grenze sind die Auswirkungen der Regelsätze von Provider 1 und Provider 2 zusammen zu erwarten. Diese Positionierung scheint eine realistische Annahme.

In Abb. 5.1 würden für Datenpakete, die zwischen dem Nutzer und einem Broker an der Position $x_2$ ausgetauscht werden, die vom Provider 1 aufgestellten Regeln (Begrenzung des Durchsatzes auf 1MBit/s und Herausfilter des Protokolls P) gelten. Hinzu käme die Regel der Begrenzung des Durchsatzes von Daten die mit dem Protokoll Q ausgetauscht werden auf 5 MBit/s.

**Position $x_3$.** Eine Messung zu einem Broker in dieser Position wird die aggregierten Einflüsse von Provider 1 und Provider 2 wiederspiegeln.

Ein Broker an dieser Position in der Beispielabbildung würde weiteren Beeinflussungen, die von r2.p2.eu und r3.p2.eu ausgehen würden, unterliegen.

Wie die Positionierung eines Brokers in einem (komplexen) Netzwerk letztendlich geschieht, wird wohl kaum kontrollierbar sein. Daher ist es um so wichtiger, dass diese Unterschiede bei der Bewertung von Testergebnissen beachtet werden. Der Einfluss der unterschiedlichen Positionierungen auf die durch einen Test getroffenen Aussagen werden in Abschnitt 5.4.5 weiter untersucht. Broker sind einheitlich durch Domainnamen wie `broker.domainname.tld` ansprechbar.

Auch auf der Seite des Endnutzers ist eine Anpassung der Netzwerkinfrastruktur notwendig: Um einen Test durchzuführen, muss eine Testsoftware zunächst Datenpakete von den vom Nutzer gewünschten Anwendungen entgegennehmen und diese dann in zur Messung geeigneter Form aufbereiten und die Testdurchführung selbst koordinieren. Dieses Programm wird im Folgenden mit „(Mess-)Client“ bezeichnet. Dieser Client enthält auch die notwendigen öffentlichen Schlüssel, um die Authentizität der Broker beim Verbindungsaufbau zu überprüfen.

Ein Testen der vom Netzbetreiber herausgegebenen Spezifikation des Netzwerkverhaltens kann vom Netzbetreiber als offensives Verhalten aufgefasst werden – entsprechend kann er versuchen, sich hiergegen zur Wehr zu setzen. Das Interesse des Nutzers ist der Nachweis der eines von der Spezifikation abweichenden Verhaltens. Das Interesse des Netzbetreibers ist die Beeinflussung oder das Misslingen entsprechender Tests.

Es kann angenommen werden, dass sich ein Netzbetreiber aller (Manipulations-) Möglichkeiten bedient, die sich aus seinem Netz ergeben: Er kann beliebige Datenpakete manipulieren oder herausfiltern. Der Nutzer kann im Gegenzug den Broker im Betreibernetz verwenden, denn dieser steht per Definition unter Kontrolle einer Regulierungsbehörde (z. B. der BNetzA), die als Garant für die Integrität des Brokers einsteht.

Ein solches Szenario gewinnt der Endnutzer jedoch, zumindest so lange nicht alle Netzbetreiber zusammenarbeiten und wie es dem einzelnen Netzbetreiber nicht möglich ist, den verschlüsselten Kanal zum Broker in einer so kurzen Zeit zu entschlüsseln, dass es

ihm möglich ist, die ausgetauschten Daten mitzulesen und auf diese z. B. mit Policy-Änderungen zu reagieren. So lange wird es dem Nutzer möglich sein, unterschiedliches Verhalten unterschiedliche Netze nachzuweisen – und damit auch eine Einflussnahme der jeweiligen Betreiber.

### 5.4.2. Durchführung einer Prüfung

Um die ermittelten Eigenschaften eines Routingpfades zu prüfen, kommt eine modifizierte Variante des Nooter-Ansatzes zur Neutralitätsprüfung zum Einsatz. Von Nooter wird hierzu eine zentrale Idee importiert: Die Idee eines Messknotens im Providernetz (die für sich kein Alleinstellungsmerkmal von Nooter ist und bereits in [23] diskutiert wurde), zu dem eine verschlüsselte Verbindung aufgebaut wird, um im Anschluss eine Differenzmessung zwischen unverschlüsselt und verschlüsselt übertragenen Daten durchführen zu können.

Zwei Herangehensweisen sind grundsätzlich für eine Messung denkbar: Es kann der in jedem Netz vorhandene Broker einzeln genutzt werden oder aber eine Kaskadierung der Broker der beteiligten Netze. Letzterer Ansatz wird iterativ bezogen auf die ermittelten Teilnetzbetreiber durchgeführt, beginnend z. B. mit dem vom Teilnehmer im Routingpfad entferntesten Netz[21].

Abb. 5.2 visualisiert die Prüfung: Im oberen Bereich ist das Netzwerk skizziert, unterteilt in Knoten, die an einer Messung aktiv beteiligt sind, und Knoten, die an der Messung passiv beteiligt sind (z. B. als eigentliches Anfrageziel oder als Router). Darunter ist ein Sequenzdiagramm für den Datenaustausch zwischen den an der Messung aktiv beteiligten Stellen (sowie das Anfrageziel) angegeben; hierbei werden verschlüsselte Verbindungen von nicht verschlüsselten Verbindungen unterschieden. Als Broker-Position wurde wie bereits bei der Vorstellung von Nooter durch Kaminsky jeweils von einer Position ausgegangen, die einen Test des den Broker enthaltenden Netzes erlaubt.

Die eigentliche Testdurchführung besteht aus einer Phase der Informationsbeschaffung, einer Phase der Testvorbereitung, einer Phase des Testens und schließlich einer Phase der Auswertung. Zunächst werden Informationen über die verwendeten Netzwerke und die in ihnen enthaltenen Broker ermittelt; im Anschluss werden diese auf die Messung mit dem konkreten Testserver eingerichtet. Hiernach kann die Datenübertragung als eigentlicher Test stattfinden. Schließlich können die ermittelten Daten ausgewertet werden.
Im Detail gliedert sich das Verfahren in die folgenden Einzelschritte:

**Routingpfad**

Ermitteln des Routingpfades zum vom Nutzer gewünschten Zielhost durch den Client: $p = (r_0, .., r_n)$

[21] Da sich hier noch keine Effekte durch Maßnahmen unterschiedlicher Netzbetreiber überlagern.

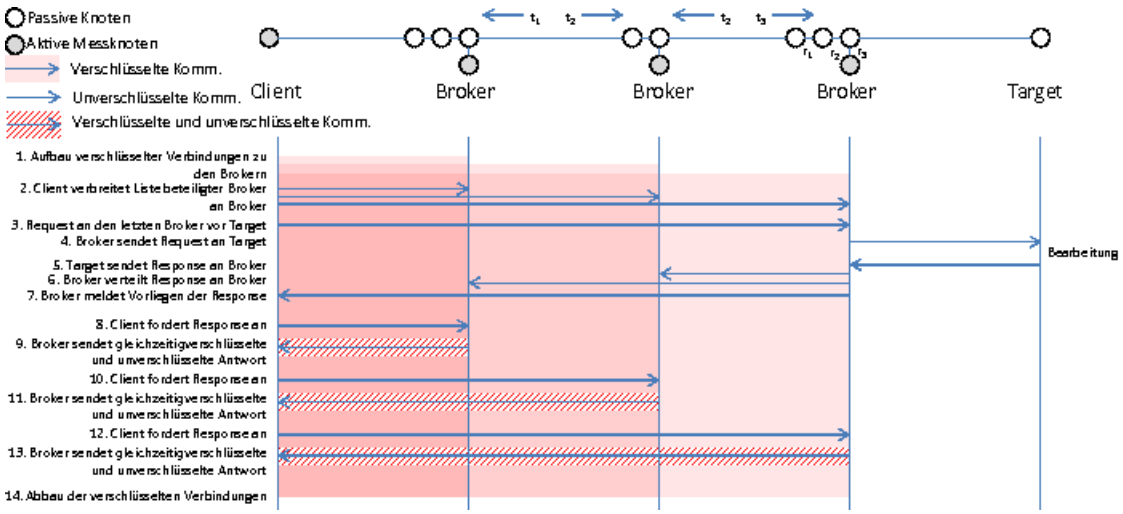


Abbildung 5.2.: Sequenzdiagramm des aktiven Testvorgangs beim Test angegebener Policies unter Benutzung von Brokern für einen Routingpfad durch 3 Teilnetzwerke.

**Boxing**

Zusammenfassen zu einem Pfad von Black Boxes durch den Client, vgl. Abb. 5.1: $t = (t_0, .., t_m)$, mit $t_1 = (r_0, .., r_i), t_2 = (r_{i+1}, .., r_j), ..., t_m = (r_{k+1}, ..r_n)$.

**Informationsbeschaffung**

Für jedes Teilnetz $t_i$:
Beschaffen der Netzwerk-Policies wie in Abschnitt 5.3 beschrieben, zwischenspeichern der Policies als $p_i$ durch den Client; daneben kann zu diesem Zeitpunkt der Hostname des jeweiligen Brokers abgeleitet werden.

**Prognose**

Berechnung der zu erwartenden Werte $\sum_i p_i$, sowohl für eine einfache Verbindung zum Zielhost als für eine Verbindung zu jedem Broker durch den Client. Wenn der Nutzer nur an der Abschätzung der nach Deklaration zu erwartenden Netzqualität interessiert war, kann eine Ausgabe erfolgen und das Verfahren hier abgebrochen werden. Andernfalls wird ab hier mit der tatsächlichen Testdurchführung begonnen.

**Kontaktaufnahme mit Brokern** Abb. 5.2, 1.:

Aufbau von verschlüsselten Verbindungen vom Client mit den Brokern in den Teilnetzen; Einrichtung der Broker für Messung als Einzelbroker oder in Kaskadierung (als Master oder Slave). Beim Aufbau der verschlüsselten Verbindung wird die Authentizität der Broker überprüft – hierzu wird analog der Validierung von HTTPs-Zertifikaten durch Browser vorgegangen: Die korrekten öffentlichen Schlüssel der Broker werden zusammen mit dem Client ausgeliefert.

**Koordination** Abb. 5.2, 2.:
Liste aller beteiligten Broker vom Client an alle beteiligten Broker versenden.

**Anfrage vorbereiten** Abb. 5.2, 3.:
Datenpaket der Nutzeranwendung vom Client an den Broker im Netz $t_m$ (Master) senden. Optional kann das Datenpaket nach einer Manipulation auch durch das unverschlüsselte Betreibernetz an das Testziel geschickt werden. Die Manipulation muss dafür sorgen, dass das Paket vom Zielserver nicht verarbeitet wird[22].

**Zielserver kontaktieren** Abb. 5.2, 4.:
Broker im Netz $t_m$ sendet das Datenpaket vom Client als Proxy an das tatsächliche Ziel.

**Zielserverantwort abwarten** Abb. 5.2, 5.:
Broker im Netz $t_m$ erhält eine die Antwort.

**Antwort verteilen** Abb. 5.2, 6.:
Broker im Netz $t_m$ verbreitet die Antwort (durch verschlüsselte Verbindungen) an die Broker in den Netzen $t_i$ mit $0 < i < m$ (Slaves).

**Bereitschaftsmeldung** Abb. 5.2, 7.:
Broker meldet Vorliegen eines Datenpakets und Erfolg der Verbreitung an den testenden Client.

**Messung** Abb. 5.2, 8ff.:
Der Client kann nun in beliebiger Reihenfolge von den einzelnen Brokern die zuletzt vom Broker im Netz $t_m$ erhaltene Antwort abfordern, die dem Client dann (9.) gleichzeitig (im Rahmen der erreichbaren Gleichzeitigkeit) verschlüsselt wie unverschlüsselt übermittelt wird. Diese nebenläufigen Übertragungen können nun mit den bekannten Verfahren aktiven Testens auf eine unterschiedliche Behandlung der beteiligten Datenpakete untersucht werden.
Dieses Vorgehen wird für alle Broker durchgeführt (Abb. figNTranspTest, 10., 11., 12. und 13.).

**Wiederholungen**
Je nach Protokoll wird nun eine weitere Anfrage an das externe Testziel gesendet (Rücksprung zu Abb. 5.2, 3).

**Verbindungsabbau** Abb. 5.2, 14.:
Nach Abschluss der Datenübertragung zwischen Client und Target werden auch die verschlüsselten Verbindungen abgebaut. Eventuell auf den Brokern zwischengespeicherte Antworten werden zu diesem Zeitpunkt verworfen; während der Messung angefallene Daten werden zum Client übertragen und im Anschluss ebenfalls verworfen.

[22] z. B. falsche TCP-Header-Checksum oder eine TTL, die verhindert, dass das Paket den Zielserver erreicht.

**Auswertung**

Abschließend können die über den Datenaustausch gesammelten Daten statistisch ausgewertet und mit den anhand der deklarierten Policies erwarteten Werte verglichen werden.

Zur Auswertung können prinzipiell beliebige Verfahren genutzt werden; es bietet sich z. B. das von Dischinger et. al [13] genutzte Verfahren an, da es bereits für Testdurchführungen unter Zeitdruck[23] optimiert wurde. Hierbei werden zunächst durch Rauschen verfälschte Daten ausgefiltert. Im nächsten Schritt werden die gemessenen Daten mit den erwarteten Referenzwerte verglichen und der Unterschied festgestellt. Entscheidend ist nun die Festlegung, ab welchem zahlenmäßigen Unterschied von einer Abweichung ausgegangen werden soll. Dischinger et al. schlagen eine relative Abweichung von 20% vor; dies ergebe eine False-Positive-Rate von unter 0.6%.

Offen bleibt zunächst die Frage, wie mit einer nachgewiesenen Abweichung umgegangen wird: Eine vertragliche Beziehung besteht nur zwischen dem Nutzer und seinem ISP; nicht jedoch mit weiteren Netzbetreibern. Sinnvoll erscheint in diesem Zusammenhang primär eine Meldung an die Bundesnetzagentur als zuständiger Regulierungsbehörde, da sich hierdurch ein Überblick über das Verhalten von Netzbetreibern ergibt.

Die Notwendigkeit der Koordination des Absendens der Serverantworten vom Broker zum Client ergibt sich, da ansonsten die unverschlüsselten Antworten der Broker vom Client nicht voneinander zu unterscheiden wären.

Wenn eine Messung nur die Summe der Effekte der Netzwerke $1..k$ betrifft, reicht die Verwendung des Brokers im Netzwerk $k$ und es kann auf das weitere Handshake verzichtet werden. In diesem Fall vereinfacht sich das Vorgehen auf den von Kaminsky vorgestellten Nooter.

Ein Verzicht auf das Verteilen der Antwort des Testservers und die eigenständige Anfrage jedes Brokers beim Testserver scheint zunächst ebenfalls denkbar. Ein solches Vorgehen scheitert allerdings spätestens dann, wenn der getätigte Aufruf eine serverseitige Statusänderung und entsprechend für die folgenden Aufrufe ein verändertes Antwortverhalten bedeutet. Ein Beispiel eines solchen nicht idempotenten Befehls ist beispielsweise das Löschen einer Datei auf einem entfernten Server per CIFS oder HTTP/WebDAV.

Hieraus ergibt sich auch eine weitere Anforderung an die Broker: Im Unterschied zum Kaminsky'schen Nooter müssen Daten vorrätig gehalten werden, da sie erst nach entsprechendem Auftrag zugeschickt werden sollen. Für jeden Client, der mit einem Broker eine

[23] Eine Untersuchung der Geduld der Nutzer mit einem Messverfahren ergab, dass eine Messung von mehr als 6 Minuten Dauer einen großen Teil der Nutzer zum Abbruch der Messung veranlasst.

verschlüsselte Verbindung unterhält, wird jeweils genau ein Datenpaket vorrätig gehalten; es wird überschrieben, wenn vom Broker im Netz $t_m$ das nächste Paket verteilt wird.

Als Variation einer solchen **Live-Messung** können die jeweils gesendeten Testdaten sowohl durch den Client wie auch durch den Broker im Netz $t_m$ zwischengespeichert werden, um nach Verbreitung der Daten von Broker $m$ an die anderen Broker in einer **Replay-Messung** den Datenaustausch mit einem anderen Broker en block zu wiederholen. Dies bietet den Vorteil einer höheren Geschwindigkeit beim Datenaustausch, da der für die Koordination der vom Broker zum Client gesandten Antwortpakete verbundene Aufwand entfallen kann. Dieser Vorteil wird durch einen deutlich erhöhten Speicherbedarf erkauft[24], ferner werden erst nach Ende der Verbindung aus Nutzersicht weitere Messungen durchgeführt[25]. Das oben detailliert beschriebene Verfahren ist für Latenzmessungen optimiert. Grundsätzlich sind Daten nur während eines Tests und keineswegs auf dem Broker zu speichern, da sie (je nach Verwendung durch den Nutzer) personenbeziehbare Daten (IP-Adresse) oder ganz klar personenbezogene Daten (Payload) enthalten können.

Performanceoptimierungen sind möglich, wenn Broker der Funktionalität nach zu vollwertigen Proxies aufgewertet werden können – also wenn Broker das unterliegende Protokoll beherrschen würden anstelle jedes einzelne Antwort-Datenpaket vom Client zu erhalten. Hierdurch können größere Transaktionen eigenständig und ohne Handshake zum Client nach jedem Datenpaket durchführt werden; als Beispiel bietet sich das Protokoll HTTP an.

Neben den bereits diskutierten Fragen der Performance existieren noch weitere mögliche Einschränkungen, die im übernächsten Abschnitt diskutiert werden; im folgenden Abschnitt wird das vorgestellte Verfahren beispielhaft an einer TCP-Verbindung illustriert.

### 5.4.3. Veranschaulichung: Evaluation einer TCP-Verbindung

Zur Illustration des vorgestellten Messverfahrens wird hier bewusst als Beispiel eine TCP-Verbindung gewählt, um die universelle Nutzbarkeit zu unterstreichen.

Als Grundlage soll die in Abb. 5.2 verwendete Topologie dienen: Der Nutzer befindet sich mit seinem Rechner im Teilnetz $t_1$; der Zielserver Target befindet sich im Teilnetz $t_3$. Die Broker seien entgegen der Abbildung an unbekannter Stelle innerhalb der Teilnetze untergebracht. Nach jedem Schritt sollen die Zustände des TCP-Zustandsübergangsmodells jeweils für das Anwenderprogramm App und den Server Serv betrachtet werden. Sie werden nur dann wiedergegeben, wenn es eine Änderung gegeben hat. Da der Fokus der Betrachtung auf dem Messvorgang liegt, wird die Spezifikationsbetrachtung

[24] Jeder Broker muss potentiell alle vom Testziel zum Client versandten Datenpakete zwischenspeichern, eine Obergrenze ist – falls ein Test mit einem Webstream o.ä. durchgeführt wird – ratsam, begrenzt allerdings die Aussagekraft.

[25] ein Problem, wenn verbindungslos kommuniziert wird

verknappt dargestellt.

### Latenzorientierte Messung

Zunächst wird das Messverfahren für eine latenzorientierte Messung skizziert. Als Latenz sei hier der Zeitbetrag zu verstehen, um den sich die Transportzeit für ein Datenpaket bei der Weiterleitung an jedem Netzwerkknoten und jeder Netzwerkkante verlängert. Um diese zu ermitteln, wird die zeitliche Differenz zwischen dem Eingehen von verschlüsselten und unverschlüsselten Datenpaketen beim Messclient betrachtet.

- Grundzustand

| App | Serv |
|---|---|
| CLOSED | LISTEN |

- Ein vom Nutzer verwendetes Programm beginnt, eine TCP-Verbindung zu Target aufzubauen. Der Messclient agiert als Default-Router und empfängt ein ausgehendes TCP-Paket an Target mit gesetztem `SYN`-Byte.

| App | Serv |
|---|---|
| SYN_SENT | LISTEN |

- Der Messclient ermittelt die auf dem Pfad zu Target[26] durchquerten Teilnetze[27] $(t_1, t_2, t_3)$.
- Der Messclient ermittelt die zur Informationsabfrage benötigten Domainnamen `t1.de`, `t2.de`, `t3.eu`; ferner werden die Netzspezifikationen abgerufen[28]. Aus den Domainnamen ergeben sich die Broker `broker.t1.de`, `broker.t2.de` und `broker.t3.eu`.
- Der Messclient baut verschlüsselte VPN-Verbindungen zu den Brokern auf und richtet die Broker für die Messung ein[29]:
  - Einzelpakete, kein Proxying
  - keine Replay-Messung
- Die Broker werden untereinander verbunden[30] und vorbereitet.
- Der Messclient übermittelt das erste Datenpaket dem Broker im Netz $t_3$[31]. Gleichzeitig manipuliert er z. B. die TCP-Header-Prüfsumme des Datenpaketes und sendet es unverschlüsselt durch das Betreibernetzwerk. Das unverschlüsselte Paket

[26] Schritt „Routing“
[27] Schritt „Boxing“
[28] Schritte „Informationsbeschaffung“ und „Prognose“
[29] Schritt „Kontaktaufnahme mit Brokern“
[30] Schritt „Koordination“
[31] Schritt „Anfrage vorbereiten“

wird zwar von Target empfangen, doch auf Grund der fehlerhaften TCP-Header-Prüfsumme verworfen. Entsprechend verändert sich der Verbindungszustand nicht.

| App | Serv |
|---|---|
| SYN_SENT | LISTEN |

- Der Broker im Netz $t_3$ sendet das Datenpaket an Target[32], wobei sich der Broker selbst als Absender des Paketes einträgt und ggf. die ausgehende Portnummer modifiziert[33].

- Target empfängt das Paket und beantwortet den Verbindungsbeginn mit einem Paket mit gesetztem `SYN ACK` an den Broker

| App | Serv |
|---|---|
| SYN_SENT | SYN_RCVD |

- Broker $t_3$ empfängt das Antwortpaket[34]. Er manipuliert die Absenderadresse, so dass das Paket dem Anschein nach von Target stammt und leitet es durch verschlüsselte Verbindungen an die Broker $t_1, t_2$ weiter. Anschließend wird dem Messclient das Vorliegen einer Antwort[35] signalisiert. Eine Kopie des Datenpaketes wird mit dem Messclient als eingetragenem Ziel vorbereitet.

- Messclient fordert die Antwort von den Brokern an[36] (z. B. in $t_3$). Broker sendet – soweit im Rahmen einer Implementation erreichbar – gleichzeitig das Datenpaket durch den verschlüsselten Tunnel und unverschlüsselt an den Messclient oder den Broker.

- Der Messclient empfängt beide Datenpakete und speichert die relative Ankunftszeit für die Auswertung. Analoges Vorgehen für die weiteren Broker.

- Der Messclient stellt das ursprünglich von Target stammende Datenpaket der Anwendung App zu, die daraufhin ein Paket mit gesetztem `ACK`-Bit an Target (den Messclient) schickt, um den Verbindungsaufbau abzuschließen.

| App | Serv |
|---|---|
| ESTABLISHED | SYN_RCVD |

- Das Paket mit gesetztem `ACK`-Bit wird vom Messclient an den Broker weitergeleitet und in modifizierter Form unverschlüsselt durch das Netzwerk verschickt. Der Messclient leitet es mit veränderter Absenderadresse (und ggf. auch angepasstem Port) an Target weiter.

| App | Serv |
|---|---|
| ESTABLISHED | ESTABLISHED |

[32] Schritt „Zielserver kontaktieren“
[33] Vermeidung von Kollisionen mit gleichzeitig laufenden Messungen
[34] Schritt „Zielserverantwort abwarten“
[35] Schritt „Bereitschaftsmeldung“
[36] Messung

Das Verfahren würde entsprechend wieder mit der Annahme eines Datenpaketes durch den Messclient fortgeführt werden.

#### Durchsatzorientierte Messung

Dieser Abschnitt beschreibt ebenfalls eine Messdurchführung anhand einer Anwendung, die sich des Protokolls TCP bedient. Als Durchsatz wird dabei diejenige Datenmenge betrachtet, die in einer bestimmten Zeit ausgetauscht werden kann.

Dabei wird auf eine Wiederholung der im vorangehenden Abschnitt geschilderten Schritte verzichtet. Der wesentliche Unterschied liegt in Messung: Anstelle der zeitlichen Differenz zwischen dem Eingehen einzelner Datenpakete wird die Anzahl innerhalb einer Zeitspanne ausgetauschter Datenpakete betrachtet. Eine wesentliche Anforderung ist hierbei nicht die Gleichzeitigkeit des Absendens der Datenpakete sondern der Durchsatz von Broker und Messclient sowie ein möglichst geringer (Protokoll-)Overhead. Entsprechend ist das Verfahren zu optimieren:

- Während der eigentlichen Kommunikation zwischen App und Target agiert der Broker in $t_3$ vergleichbar einem NAT-Gateway; er speichert allerdings alle ausgetauschten Datenpakete zwischen. Auch der Messclient leitet die Kommunikation lediglich an den Broker in $t_3$ weiter.
- Zur Messung wird die Verbindung durch den Messclient und den Broker in $t_3$ wiederholt – nun unter Aussendung entsprechender unverschlüsselter Datenpakete durch die jeweiligen Netze. App und Target sind an dieser eigentlichen Messung nicht mehr aktiv beteiligt.

Der folgende Abschnitt verlässt Überlegungen auf Protokollebene und wendet sich möglichen Umsetzungen eines Testclients zu.

### 5.4.4. Umsetzungsskizzen

In diesem Abschnitt sollen drei Umsetzungsideen eines Testclients in Vor- und Nachteilen abgewogen werden: Eine Umsetzung mittels Fathom (vgl. Kap. 3; [15]) als In-Browser-Lösung, eine Umsetzung als vom Nutzer ausgeführtes Programm und eine Umsetzung auf einer separaten Hardware.

Eine Umsetzung eines Testclients als Browser-Plugin mit dem Fathom-Framework würde darauf basieren, dass sämtliche Verbindungen des Browsers von einem Plugin gesteuert (und somit auch analysiert) werden können. Dies wäre für den Nutzer sicherlich sehr komfortabel, denn der Einrichtungsaufwand kann minimal gehalten werden. Das Fathom-Framework stellt die zur Umsetzung benötigten Funktionen (insbesondere zum Auffangen ausgehender Netzverbindungen) bereit. Da sich die Netznutzung jenseits abgesetzter Geräte (VoIP, IPTV) zusehends weg von separaten Anwendungen in den Browser verlagert,

ist die Beschränkung auf einen Browser ein vertretbarer Preis für die Einfachheit von Installation und Nutzung. Als Nachteil bliebe die Beschränkung auf einen spezifischen Browser.

Alternativ bietet sich die Entwicklung eines klassischen Programms[37] an, dass sich durch entsprechende Systemaufrufe in sämtliche ausgehenden Netzwerkverbindungen z. B. als virtuelle Default-Gateway einklinkt um diese über Broker zu leiten. Der wesentliche Vorteil wäre das Erfassen sämtlicher von einem Rechner ausgehenden Verbindungen, auch von anderen Anwendungen als einem Browser wie z. B. von Mail-Clients, Spielen oder VoIP-Lösungen wie „Skype“. Nachteil wäre der mit einer entsprechenden Entwicklung einhergehende Aufwand und die Plattformabhängigkeit; eine Anwendung mit dem benötigt feingranularen Zugriff auf Netzwerkfunktionen wird nur schwer so zu entwickeln sein, dass sie sich gleichermaßen auf Microsoft Windows, Apple Mac OS X und Linux einsetzen ließe.

Diese Probleme entfallen mit der Entwicklung einer eigenen Plattform für das Messverfahren, die auf Netzwerkebene physikalisch eingeschleift wird. Eine (ausgediente) Maschine mit zwei Netzwerkkarten könnte als Basis einer derartigen Entwicklung dienen. In diesem Fall würde die Entwicklung die Anpassung eines linuxartigen Betriebssystems umfassen, auf dem dann die Messung stattfinden würde. Wesentlicher Nachteil ist der offensichtlich hiermit verbundene höhere Aufwand; wesentlicher Vorteil ist die gewonnene Unabhängigkeit von konkreten Endgeräten.

Einzig die Untersuchung mobiler Internetanschlüsse, die durch Smartphones genutzt werden, bleibt ein längerfristiges Problem, da Smartphones und die sie kontrollierenden Unternehmen zumeist sehr restriktiv bezüglich der Installation von Anwendungen sind – eine entsprechende Messanwendung bräuchte Systemrechte, die Besitzern von z. B. Apple-Geräten durch den Hersteller nicht gewährt werden.

### 5.4.5. Systematische Grenzen des Messverfahrens

Eine erste Beschränkung ergibt sich aus der Durchführung der Messung selbst: Da ein einzelnes Handshake (Austausch zweier Datenpakete: Ein Datenpaket vom Anwenderprogramm an den Server und ein Antwortpaket vom Server an das Anwenderprogramm) in mindestens sieben[38] Transaktionen[39] aufgeht, sind Messungen des Durchsatzes schwierig – das Verfahren ist in dieser Form zunächst für eine Messung der Latenz optimiert, da sich diese aus der Differenz der Ankunftszeiten ergibt. Für eine Messung des Durchsatzes

[37] und, je nach Betriebssystem, eines entsprechenden Treibers

[38] für einen Broker; je Broker käme ein weiterer Paketaustausch hinzu

[39] Datenpaket von Anwenderprogramm an Messclient, Paket vom Messclient an Broker, Broker an Server, Server an Broker, Broker verteilt Antwort an andere Broker, Broker meldet Bereitschaft an Messclient, Messclient fordert an, Übermittlung von verschlüsseltem und unverschlüsseltem Paket, Messclient liefert Paket an Anwenderprogramm aus, Messclient fordert Antworten anderer Broker ab; erst im Anschluss kann das nächste Datenpaket vom Server an den Nutzer zugestellt werden

ist eine Messung unter Beteiligung eines einzelnen Brokers oder als Replay-Messung effektiver, da hierbei der Overhead für die Koordination der Messungen mit den einzelnen Brokern entfällt, wonach nur noch sechs Transaktionen verbleiben, bei denen nicht auf die Messläufe mit sämtlichen Brokern gewartet werden muss.

Eine weitere systematische Grenze ist dem Verfahren durch den vom Endnutzer bereitgestellten Internetanschluss gezogen: Zur vergleichenden Messung werden – möglichst zeitgleich – die Antwortdaten verschlüsselt wie unverschlüsselt übertragen. Dies bedeutet eine effektive Verdoppelung des übertragenen Datenvolumens. Im Sinne eines Einflusses der Messung auf die Messgröße ist dieser Effekt – in Abhängigkeit vom dem Nutzer zur Verfügung stehenden Anschluss[40] – zu berücksichtigen. Eine solche Lastsituation kann allerdings auch positive Seiteneffekte bewirken: Bestimmte Traffic-Engineering-Methoden in ISP-Netzen, wie beispielsweise Trafficshaping mit Priorisierung bestimmter Datenströme, werden erst ab einer bestimmten Auslastung wirksam; da die Auswertung ausschließlich auf den relativen Unterschieden zwischen den verschlüsselt und unverschlüsselt übertragenen Daten beruht, sind die – bei einem Best-Effort-Ansatz gleichmäßig verteilten – Änderungen bei Priorisierung oder Diskriminierung einzelner Datenströme auftretenden Unterschiede deutlicher ausgeprägt zu erwarten.

Auch der Genauigkeit der Auswertung einer Messung sind durch den Nutzer Grenzen gesetzt: Da sich eine Messung stets an den vom Nutzer ausgetauschten Datenpaketen orientiert, kann nicht garantiert werden, dass bei einer Messung ausreichend viele Testläufe für eine statistisch sichere Auswertung durchgeführt werden. Der Download einer großen Datei wird mehr Rohdaten zur statistischen Auswertung liefern als kurzes E-Mail-Abrufen. Entsprechend kann eine begleitende Messung nur zum Sammeln von Verdachtsmomenten verwendet werden und eine begleitende Überschlagsrechnung, welche Netzqualitäten zu erwarten sind, darf entsprechend ungenau sein. Für eine exakte Auswertung kann auf Initiative des Nutzers hin z. B. nach Vorschlag durch auf Grund bisheriger Messungen im Anschluss eine Replay-Messung durchgeführt werden. Da bei einer Replay-Messung kein Kontakt mit dem eigentlichen Zielserver mehr vonnöten ist, können die angefallenen Daten praktisch unbegrenzt (bis zur für die gewünschte statistische Sicherheit notwendige Menge an Rohdaten) ausgetauscht werden.

Da sich eine Infrastruktur zur Kontrolle der deklarierten Netzwerkregeln innerhalb der Infrastruktur des Netzwerkbetreibers befindet, eröffnen sich diesem natürlich weitgehende Manipulationsmöglichkeiten. Denkbar ist ein Verhindern der Kommunikation mit einem Broker oder auch die Verschlechterung der Qualität der verschlüsselten Verbindung zwischen Client und Broker. Denkbar ist auch eine statistische Analyse der mit dem Broker ausgetauschten Daten, verbunden mit einer Analyse weiterer an den Client gerichteter Daten.

Solche Manipulationsmöglichkeiten sind durch rechtliche Beschränkungen und tech-

[40] als typischerweise durchsatzschwächstem Teil des Routingpfades

nische Umsetzungen auf ein Mindestmaß zu begrenzen, werden aber – aus rein systematischen Gründen – nie mit absoluter Sicherheit zu verhindern sein: Die Absicherung gegen einen Gegner mit physikalischem Zugriff ist Prinzip bedingt schwierig. Im gerade geschilderten Szenario ist der einzige Ausweg, sicherzustellen dass ein Eingriff des Netzbetreibers in eine laufende Messung wenigstens als Eingriff sichtbar wird.

Weiterhin gibt es systematische Begrenzungen, welche qualitativen Aussagen nach einer Messung getroffen werden können; insbesondere die Position der Broker innerhalb der Betreibernetzwerke hat, wie bereits in 5.4.1 geschildert, einen maßgeblichen Einfluss auf die Bedeutung der gemessenen Größe. Bezüglich der Aussage resultiert aus der Positionierung der Broker stets ein „blinder Fleck“: Würden alle Broker (vom Client aus betrachtet) am Eingang der Netze stehen (Position $x_1$ in Abb. 5.1), wäre das letzte Netzwerk nicht mehr zu prüfen, ebenso ergäbe sich ein Problem für das ISP-Netzwerk: Der Broker stünde beim Nutzer. Das Problem löst sich für das Providernetzwerk durch den Broker im nächsten nachfolgenden Netzwerk. Bei konsequenter Positionierung der Broker an der entferntestmöglichen Position ($x_3$ in Abb. 5.1) hingegen ergäbe sich das praktische Problem, dass vor jedem potentiellen Zielserver ein Broker zu installieren wäre – was aus praktischen Gründen ausgeschlossen werden kann. In der Realität können Routing-Traces Aufschluss über die tatsächliche Position geben und sind bei der Aussage eines Tests entsprechend zu berücksichtigen.

Eine weitere systematische Grenze ist durch die räumliche Ausdehnung des Internets über mehrere Staaten gegeben: Einen Broker innerhalb des Netzwerks eines Zugangsproviders wird es nur mit dessen Unterstützung oder durch eine einheitliche rechtliche Regelung geben. Entsprechend würden sich bei einer Prüfung das Verhalten des Targets und der vorgelagerten Netzwerke kumulieren und wären nur zusammenhängend betrachtbar.

Unempfindlich ist das Verfahren hingegen gegenüber Einflüssen, die aus der Performance des angefragten Testservers selbst resultieren, da diese ja nur die Antwort zum anfragenden Broker betreffen – nicht aber die von diesem weitergeleiteten Antworten.

Für eine sinnvolle, aussagekräftige Auswertung darf die Datenübertragung ein bestimmtes Volumen nicht unterschreiten, da ansonsten eine Unterscheidung zwischen kurzzeitigen Lastsituationen und längerfristiger Bevor- oder Benachteiligung nicht mehr unterschieden werden kann. Anschaulich lässt sich dies mit z. B. mit dem von Shaperprobe (vgl. Kap. 3; [36]) gesuchten Einschwingverhalten bestimmter Shapingalgorithmen illustrieren: Wenn ein Datentransfer komplett in der „Burst“-Phase des Einschwingvorgangs abgeschlossen werden kann, wird ein bei größeren Datenvolumina aktiver Shaper nicht bemerkt werden können. Im Ergebnis führt auch die Forderung nach einer statistischen Auswertung mit einer bestimmten Sicherheit zur selben Anforderung.

Schließlich bleibt noch die Frage, wie eine „Best-Effort“-Regel zu validieren ist. Hierbei handelt es sich jedoch um eine Instanz des Urproblems des Neutralitätsnachweises: Ge-

zeigt werden kann lediglich die Abwesenheit von Best-Effort; nicht jedoch die tatsächliche Existenz; zumindest nicht ohne Vorliegen sehr detaillierter Informationen[41]. Diese Fragestellung ist daher aus systematischen Gründen nicht zu beantworten. Abweichungen von Best-Effort-Regelungen hingegen sind nachvollziehbar.

## 5.5. Rechtliche Aspekte

Ein weiteres systematisches Problem ergäbe sich aus der Frage nach der Bereitschaft der Netzwerkbetreiber, Broker in ihren Netzen zu dulden. Dieser Frage wird in diesem Abschnitt nachgegangen; hierbei werden die Regelungen des TKG und der EU-Verordnungsentwurf nach Regelungen untersucht, die die Deklaration von Policies in der in Abschnitt 5.2 geschilderten Form und die Installation von Brokern zur gerade geschilderten Prüfung des Einhaltens dieser Policies unterstützen.

### 5.5.1. Nach TKG

In § 41a TKG zur Regelung der Netzneutralität findet sich kein Hinweis, dass die vorgezeichnete Verordnung eine Transparenzverpflichtung wie hier skizziert enthalten kann; es existieren zwar Bußgeldvorschriften, weitere Instrumente zur Regulierung und Überwachung eines Einhaltens der Diskriminierungsfreiheit werden jedoch nicht benannt [33].

Ergiebiger ist § 43a, der zunächst in Absatz (1), 2. Anbieter öffentlicher Telekommunikationsdienste zur Angabe technischer Leistungsdaten verpflichtet. Diese werden in Absatz (2) weiter ausgeführt. Relevant im Kontext dieses Kapitels ist die Verpflichtung zur Angabe von Informationen über Informationen zu Einschränkungen bezüglich der Nutzung von Diensten und Anwendungen sowie zum Mindestniveau der Dienstqualität nebst möglicher weiterer nach § 41a festgelegter Parameter. Die Regelung nach § 43a (2) geht weiter als der hier vorgestellte Vorschlag, denn nach 4. sind auch Informationen über die zu Messung und Kontrolle des Datenverkehrs bzw. zur Abwehr von Engpässen durch den Betreiber eingerichteten Verfahren anzugeben.

Die Informationen sind nach § 43a (1) dem Verbraucher bzw. anderen Endnutzern in „klarer, umfassender und leicht zugänglicher Form“ zur Verfügung zu stellen – diese Erfordernisse werden von der skizzierten Form der einheitlichen Veröffentlichung von Regelsätzen erfüllt[42]; insbesondere, wenn ergänzend eine Software hinzukommt, die die Verständlichkeit erhöht[43].

Diese Regelung zielt zunächst jedoch nur auf Vertragsparteien – es ist keine Regelung einer allgemeinen Zugänglichmachung dieser Informationen. Hilfreich ist in diesem Kon-

[41] Beispielsweise als Vorliegen sämtlicher aktiver Konfigurationen oder Kenntnis sämtlicher Datenströme, die einen Router beeinflussen.

[42] Nicht zuletzt hier motiviert sich erneut die Forderung nach einer Menschenlesbarkeit der Spezifikation.

[43] Hier wäre gerade die Verwendung von Domainnamen anstelle von IP-Adressen und Protokollen anstelle von Portnummern (soweit fest zugeordnet) zu nennen, da diese Verbrauchern eine deutlich verständlichere Sicht auf Netzwerkbeschränkungen bieten.

text jedoch die Verpflichtung von Betreibern öffentlicher Kommunikationsnetze dazu, Anbietern öffentlicher Kommunikationsdienste die benötigten Informationen bereitzustellen. Hieraus ergäbe sich in der Praxis ein Vorgehen, bei dem die Policies weiterer Netzbetreiber dem ISP zur Verfügung gestellt werden, der diese dann seinen Kunden zugänglich macht. Ein direkter Zugang zu allen Policies auch für Endnutzer wäre in diesem Sinne nur pragmatisch.

§ 43a (3) schließlich ermächtigt die BNetzA dazu, Hilfsmittel zu entwickeln, mit denen Teilnehmer eigenständige Messungen durchführen können.

Ein solches Hilfsmittel stellen die in diesem Kapitel vorgestellten Broker dar; mit dieser Formulierung ist auch die Zulässigkeit der Installation von Brokern nach Sinn und Zweck faktisch geregelt: Die Entwicklung eines Hilfsmittels wäre unsinnig, wenn zur Benutzung des Hilfsmittels notwendige Eingriffe in die Infrastruktur des zu untersuchenden Gegenstandes untersagt bleiben würden.

### 5.5.2. Nach EU-Verordnungsentwurf

Auch der Verordnungsentwurf der EU sieht eine transparente Gestaltung von Netzwerkpolicies vor – dies wird in Art. 25 geregelt. In (1) Abschnitt i. wird eine Verpflichtung zur transparenten Informationsgestaltung für Anbieter elektronischer Kommunikation vorgeschlagen. Diese umfasst neben Angaben zum Betreiber auch die technischen Details des Netzzuganges; ein besonderes Augenmerk liegt hier auf Kontingentbestimmungen. Diese Informationen sollen in „klarer, verständlicher und leicht zugänglicher Form“ publiziert werden. In (3) wird die Bedeutung der Überprüfbarkeit der von den Netzbetreibern angegebenen Richtlinien abgebildet. Allerdings wird hier nicht der Weg der vom Nutzer selbst durchgeführten Messungen eingeschlagen, sondern eine unabhängige Zertifizierungsstelle vorgeschlagen, bei der sich Nutzer informieren sollen.

Vorteil einer solchen Institution wäre sicherlich, dass entsprechende Informationen so aufbereitet werden können, dass sie auch technisch weniger versierten Nutzen zur Verfügung stehen könnten. Nachteil einer solchen Institution wären zunächst die Kosten und eine mit einer solchen Institution einhergehende Vorverarbeitung der Rohdaten über Netzwerke – auch der versierte und interessierte Nutzer wird ggf. keine weitergehenden Informationen erhalten können, da das Angebot auf die mehrheitlich nicht versierten Nutzer abgestimmt ist.

## 5.6. Bewertung

Dieses Kapitel hat das bekannte Thema Netzneutralität zum Themenkomplex Netztransparenz erweitert: Anstelle der Frage, ob es Abweichungen vom Best-Effort-Ansatz gibt, steht die Frage, welche Abweichungen es gibt im Mittelpunkt. Um diese Frage zu beantworten, wurden die Anforderungen an eine Deklaration und eine mögliche technische

Realisation zur Umsetzung einer einheitlichen Spezifikation von Netzwerkpolicies skizziert. Hierauf aufbauend wurde eine hybride Methode zur Überprüfung des Einhaltens deklarierter Netzwerkpolicies entwickelt. Beides ist nach geltendem Recht in Deutschland durch die Bundesnetzagentur umsetzbar; der Entwurf der EU sieht ebenfalls eine weitgehende Veröffentlichung von Netzwerkpolicies vor, scheut jedoch davor zurück, dem Endnutzer Messinstrumente in die Hand zu geben. Stattdessen wird auf eine Zertifizierungsstelle als neutrale Instanz gesetzt.

Das in diesem Kapitel vorgestellte Verfahren zur Überprüfung von Policy-Deklarationen lässt sich in die Gruppe der hybriden Testverfahren einordnen. Entsprechend erbt es deren Vorteile und Herausforderungen:

### Vorteile

- Grundsätzlich sind sämtliche vom Nutzer durchgeführten Datentransfers auch als Testverbindungen nutzbar, die Abdeckung entspricht damit potentiell dem vom Nutzer verwendeten Internet (vom passiven Ansatz geerbter Vorteil).
- Das Testverfahren ist hinsichtlich des Protokollstapels oberhalb des Netzwerklayers des ISO/OSI-Modells für eine Anwendung unsichtbar; daher können beliebige Anwendungen als Datenquelle auf Clientseite verwendet werden (vom passiven Ansatz geerbter Vorteil).
- Da es sich bei den durchgeführten Tests zwischen Brokern und Clients um aktive Tests handelt, können auf beiden Seite detaillierte Aufzeichnungen erfolgen (von aktiven Ansatz geerbter Vorteil).
- Die Datenauswertung erfolgt ausschließlich auf dem Client.
- Manipulationen durch Netzbetreiber fallen stets zugunsten des Endnutzers aus (vom hybriden Ansatz geerbter Vorteil), da sie wenigstens als solche sichtbar werden.

### Herausforderungen

- Aus der Funktion der Broker ergeben sich mögliche Datenschutzprobleme: Nicht nur die IP des eigentlich Testenden sondern auch die während des Tests ausgetauschten Daten werden zwischengespeichert[44].
- Sehr kurze Datenübertragungen sind schwierig auszuwerten und müssen ggf. wiederholt werden[45].

Weitere Forschung könnte sich mit den folgenden – noch offenen – Fragestellungen beschäftigen:

[44] in der Variation sogar als kompletter Datenblock
[45] was allerdings nur in der Variation ohne erneute Kommunikation mit dem Testziel möglich ist

Wie kann mit der Positionierung von Brokern im Providernetzwerk umgegangen werden? Im Rahmen dieser Arbeit wurden obere und untere Schranken beschrieben, lassen sich hier trotz fehlenden Einflusses engere Eingrenzungen finden? Welche Möglichkeiten zur Umgehung einer solchen Architektur gibt es und welche Preise sind die Betreiber von Netzwerken hierfür bereit zu zahlen (spieltheoretische Überlegungen: Wie viele Kunden werden eine solche Messinfrastruktur verwenden – und wie viele Kunden werden sich aus Bequemlichkeit darauf verlassen, dass der Netzwerkbetreiber sich an die Deklaration hält?).

Im Rahmen dieses Kapitels sind keine Fragen der gesellschaftlichen und politischen Umsetzbarkeit diskutiert worden. Die im Sommer 2013 schnell laut gewordene Debatte um einen Vorstoß der „Deutschen Telekom“ zeigt jedoch, dass eine Abkehr von der dogmatischen Neutralität hin zu einer reinen Transparenz eine Reaktion der Öffentlichkeit hervorrufen könnte.

Denkbar sind Einwände derjenigen, die jedes Abweichen von einer Best-Effort-Verpflichtung für sämtliche Datenpakete als Dammbruch betrachten; andere Einwände sind von Seiten der Netzwerkbetreiber zu erwarten: Diese werden mit Aufwand, Kosten und Risiken der Messinfrastruktur argumentieren.

Ebenfalls mit Einwänden ist gegen die Veröffentlichung von gefilterten Inhalten und Adressen zu rechnen; diese ist allerdings nach TKG § 43a und EU-Entwurf ohnehin vorgesehen.

Entgegen der Intention als Hilfsmittel zur Überprüfung von Netzpolicies können die im Netz eingesetzten Broker auch als Teile eines Verschleierungsnetzes verwendet werden:

Der Nutzer verbirgt seine Absender-IP durch die Verwendung eines Brokers gegenüber dem eigentlichen Zielrechner. Ferner lassen sich Broker kaskadieren: Ein Nutzer kann durch einen Broker eine Verbindung zu einem anderen Broker aufzubauen – am Ende steht zwar keine absolute Anonymität, doch der tatsächliche Nutzer kann sich hinter IP-Adressen verbergen, denen er nicht zugeordnet ist. Auch diese Anwendung ist ein „spätes Erbe“ des Nooter-Ansatzes von Dan Kaminsky. Nooter steht bei ihm für „Neutral Router“.

Allerdings ist der Aufwand einer solchen Anwendung relativ hoch – gerade im Vergleich zur Verwendung existenter Verschleierungsdienste mit nutzerfreundlichen und dokumentierten Frontends wie etwa TOR.

*In beiden Einführungsbeispielen (vgl. S. 17) hätten die Protagonisten vom hier vorgestellten Spezifikations- und Testsystem profitiert: Im ersten Einführungsbeispiel hätte bereits die Spezifikation die Reglementierungen bezüglich Verbindungen zur Adresse* `skype.com` *deutlich gemacht. Das von Martin im zweiten Einführungsbeispiel bemerkte Phänomen, dass sich bestimmte Dateien von seinem Arbeitsplatz aus nicht aus dem Internet herunterladen lassen, wäre ebenfalls schnell einem Verursacher zuzuschreiben gewesen: Der gleiche Transfer wäre mit der hier vorgestellten Messtechnik zu den folgenden Ergebnissen gekommen:*

- *Das Durchsuchen der Netzwerkspezifikationen ergäbe nichts: Keines der beteiligten*

*Netzwerke deklarierte ein Filtern der entsprechenden Bytesequenzen.*

- *Die Anfrage an den Server wäre von einem Broker im Netzwerk des Content-Providers gestellt worden. Würde dieser die Datei unvollständig erhalten, läge das Problem entweder im Content-Provider-Netzwerk oder aber beim Server selbst.*
- *Alle unverschlüsselten Transfers bis auf den vom Broker im Universitätsnetz würden bei der kritischen Bytesequenz abgebrochen werden*[46]*.*
- *Alle verschlüsselten Transfers würden die Datei erfolgreich ausliefern.*

*Somit ließe sich schnell feststellen, dass es ein Problem im Bereich des Rostocker Universitätsnetzes gibt; selbst ohne Blick in eine Spezifikation der Netzwerkeigenschaften. Als angenehmen Nebeneffekt hätte Martin die entsprechenden Dateien bei seinen Testläufen erfolgreich herunterladen (und so eine Inhaltsfilterung umgehen) können.*

Nachdem in diesem Kapitel die Transparenz von Netzwerken beschrieben wurde, werden im folgenden Kapitel Anwendungen mit Netzwerkbezug unter ähnlichen Fragestellungen untersucht.

[46] Wenn der unverschlüsselte Datenstrom des Brokers zum Client nicht der Kontrolle und dem Einfluss des IDS unterläge – die Problematik der konkreten Position eines Brokers im Netz.

# 6. Transparenz von Anwendungen

Die Öffentlichkeit hat eine unersättliche Neugier, alles zu wissen,
nur nicht das Wissenswerte.
—*Oscar Wilde*

*Warum hat Martin im zweiten Einführungsbeispiel versucht, die fragliche Daten zunächst mit unterschiedlichen Browsern herunterzuladen? Ganz offensichtlich, weil er nicht ausschließen konnte, dass das beobachtete Phänomen seinem primären Browser entsprang. Hätte er sich anders verhalten, wenn er den Quellcode des Browsers bis in die letzte Zeile gelesen und verstanden – oder gar selbst geschrieben – hätte? Vielleicht nicht. Das beobachtete Phänomen ist so überraschend, dass ein Entwickler wohl an komplexe Ursachen wie Race-Conditions gedacht hätte und dass der Test mit einem anderen Browser die trivialste Lösung darstellte.*

Wie oft sollten Anwendungen verglichen werden um nachzuprüfen, ob eine andere Anwendung die gleichen Ergebnisse liefert – und wie oft ist ein Anwender praktisch dazu in der Lage? Wieviel weiß der Anwender eigentlich über eine Anwendung?

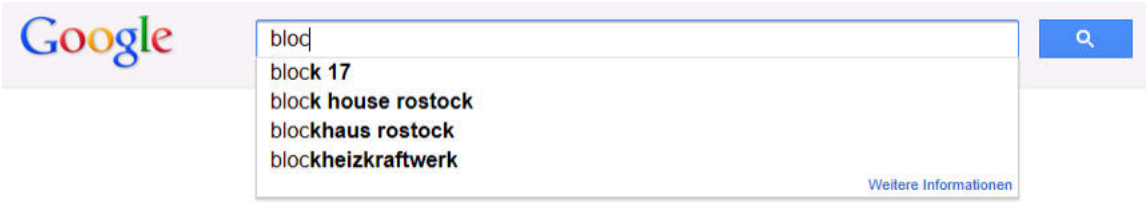


Abbildung 6.1.: Der Nutzer erfährt nicht, weshalb ihm exakt diese Vorschläge gemacht werden – und weshalb andere Begriffe („**Bloc**ksberg“) nicht enthalten sind.

Weiß ein Anwender, dem sich eine Suchmaschine wie in Abb. 6.1 mit einer Liste von Vorschlägen präsentiert, weshalb diese Vorschläge gemacht werden – und keine anderen? Die Untersuchung dieser Fragen wurde in [10] prominent, wenn auch in gegenüber diesem Kapitel gekürzter Form, publiziert.

Oftmals wird gerade im Zusammenhang mit Suchmaschinen die Forderung nach einer „Suchmaschinenneutralität“ als eine Gleichbehandlung aller Eingaben gefordert – diese

wäre ein Spezialfall von Anwendungsneutralität.

Neutralität von Anwendungen kann es jedoch nur in Ausnahmefällen geben: Neutral wäre diejenige Anwendung, deren Ausgabe stets nach einer feststehenden und nicht von der Eingabe abhängigen Regel erzeugt wird. Wichtig ist die Abgrenzung zum erwartungs- oder spezifikationsgemäßen Verhalten, das intuitiv oft als „Neutralität“ bezeichnet wird. Ein Programm wie `cat`, welches die Eingabe direkt zur Ausgabe macht oder ein Programm, dass jede Zeile der Eingabe bei der Ausgabe um ein Zeilennummer ergänzt, wären beispielsweise neutral. Bereits ein Programm wie `grep`, dass Zeilen ausfiltert, ist per Definition nicht mehr neutral, denn es behandelt Eingaben (je nach ihrem Inhalt) unterschiedlich. Eine Anwendungsneutralität lässt sich also definieren, doch sie wird in den wenigsten Fällen ernsthaftes Ziel sein können; in der Praxis sind es eher Fragen der Erwartungs- und Spezifikationserfüllung, wenn Verhalten und „Fairness“ von Programmen diskutiert[1] wird.

Transparenz von Anwendungen ist – wie auch bei Netzwerken – eine relative Eigenschaft, die sich erst aus der Betrachtung durch den Anwender und seinem Wissen über eine Anwendung ergibt. Maschinelle Vergleiche der Transparenz von Anwendungen sind kaum möglich, wie bereits der Blick auf die Terminologie nahelegt. Die Transparenz einer Anwendung kann in Abhängigkeit vom Anwender subjektiv stark unterschiedlich wahrgenommen werden. Hieraus folgt die Unmöglichkeit einer objektiven Metrik jenseits der Unterscheidung zwischen Umständen, die der Anwender erfahren kann und solchen, die ihm verschwiegen werden, mithin ist noch nicht einmal eine klare duale Klassifikation möglich.

Im Folgenden wird zunächst der Forschungsstand zur Transparenz von Anwendungen aufgearbeitet; hiernach wird die Existenz von intransparenten Anwendungen in allen Bereichen und nicht nur in einzelnen Nischen ausgeführt und auf mögliche Motive hin untersucht. Schließlich wird ein Vorschlag zur transparenteren Gestaltung von Anwendungen skizziert.

## 6.1. Verhalten von Anwendungen

In diesem Abschnitt werden bestehende Vorarbeiten zum Verhalten von Anwendungen betrachtet.

Es gibt primär zwei Arten von Anwendungen, deren Verhalten gegenüber dem Nutzer hinsichtlich eines Transparenzkriteriums bislang in der Literatur mit besonderer Aufmerksamkeit verfolgt wurde: Suchmaschinen und soziale Netzwerke [58, 7, 9, 8, 5].

[1] oder Suchmaschinenneutralität gefordert

Beide Arten von Anwendungen sind gerade deshalb besonders effektiv, weil sie eine planerische Lücke des Internets schließen. Aufgrund seines dezentralen Aufbaus als Netzwerk ohne zentrale Inhaltsverwaltung oder -verzeichnisse[2], kommt Suchfunktionen eine besondere Rolle zu. Ihre Dienste sind für viele Nutzer die einzige Möglichkeit, Informationen oder Personen im Internet zu finden. Aus diesem Grund wird insbesondere die Rolle der Suchmaschinen auch als die eines „Gatekeepers“ beschrieben.

### 6.1.1. Beteiligte Parteien

Bevor auf die Beobachtungen im Detail eingegangen wird, sollen die an der Verwendung von Anwendungen beteiligten Parteien kurz betrachtet werden.

In untersuchten Konstellationen des Anwendungsverhaltens treffen regelmäßig drei Gruppen mit unterschiedlichen Interessen aufeinander, was hier am Beispiel einer Suchmaschine verdeutlicht wird:

**Nutzer.** Die Nutzer haben an eine Suchmaschine die Erwartung, die „objektiv besten[3]“ Treffer in Abhängigkeit von den von ihnen eingegebenen Suchbegriffen zu erhalten und verhalten sich entsprechend [58].

**Suchmaschinenbetreiber.** Der Betreiber einer Suchmaschine wäre idealerweise in der Position eines interessenlosen Vermittlers. In der Praxis sind Suchmaschinen ein Geschäftsmodell und Suchmaschinenbetreiber mittlerweile oft vertikal integrierte Unternehmen, das Angebot einer Suche ist also nur einzelner Unternehmenszweig etwa neben eigenen Inhaltsangeboten (Nachrichtenseiten, Freemail, etc). Ein Interesse an der Stärkung der anderen Geschäftsbereiche erscheint plausibel. Suchmaschinenbetreiber haben – so sie sich z. B. durch Werbung finanzieren (und daher wie soziale Netze an Nutzerbindung interessiert sind, vgl. [5]) – ferner das Interesse, die Erwartungshaltung der Nutzer zu erfüllen.

**Inhalteanbieter.** Inhalteanbieter haben das wesentliche Interesse, von Nutzern in bestimmten Kontexten von Suchtermen gefunden zu werden[4] – Inhalteanbieter sind an einer möglichst optimalen Platzierung unter den Suchergebnissen interessiert, was sowohl hohe Rankings in positiven Zusammenhängen wie auch niedrige Rankings in negativen Zusammenhängen bedeuten kann.

Zwischen diesen drei Interessensgruppen im Umfeld von Suchmaschinen gibt es – jenseits besonderer Konstellationen – keinerlei vertraglich geregelte Beziehungen. Anträge auf Aufnahme in einen Suchindex sind unüblich; Suchmaschinen erarbeiten sich ihre Kataloge autonom durch automatisiertes Durchsuchen des Internets (durch das Verfolgen

[2] Frühe Bemühungen, redaktionell betreute Verzeichnisse zu erstellen und zu pflegen, wurden mittlerweile wieder eingestellt.

[3] in der subjektiven Wahrnehmung des Nutzers

[4] Und in anderen, negativen, Kontexten nicht. Fälle, bei denen Dritte versuchen negative Begriffe mit einer Person oder einem Unternehmen zu verbinden sind als „Google-Bombing“ bekannt [78].

von Verlinkungen und automatisierte Bewertung der gefundenen Seiten und Beziehungen zwischen Seiten). Dennoch bestehende Geschäftsbeziehungen – etwa zur Erlangung prominenterer Platzierungen [79] – sollen hier nicht näher betrachtet werden, da angenommen werden kann, dass sie ohnehin der Erwartungshaltung der Nutzer zuwiderlaufen.

Die Erwartungshaltung der Nutzer zusammen mit der Gatekeeper-Position der Suchmaschinen hat dazu geführt, dass das ganz reale Leben von Nutzern durch Algorithmen beeinflusst wird. Diese Kombination wird problematisch, wenn Nutzer die sie taxierenden Algorithmen nicht kennen und nicht verstehen. Ein alltägliches Beispiel ist die Sortierung von Artikeln bei Online-Händlern wie Amazon nach einer Suche: Werden diejenigen Artikel angezeigt, die am besten auf die Suchanfrage passen – oder diejenigen, die am besten zur Suchanfrage im Kontext des Benutzerprofils passen[5]?

Das Interesse der Inhalteanbieter an einem bestimmten – ihnen genehmen – Verhalten der Suchmaschinenbetreiber ergibt sich aus den Auswirkungen der Suchergebnisse auf Konsumentscheidungen; Suchmaschinenbetreiber verwendeten in der Vergangenheit mit Erfolg den Ausschluss von den angezeigten Ergebnissen als „Zwangsmaßnahme“, um ein bestimmtes Verhalten der Inhalteanbieter (z. B. den Verzicht auf bestimmte Techniken zur Verbesserung der Platzierung in Suchergebnissen) zu erzwingen (vgl. [80]) und demonstrierten so ihre Macht[6] bei der Einflussnahme auf Kaufentscheidungen.

In diesem Kontext wird gelegentlich von „Suchmaschinenneutralität“ als denkbarer Forderung an Suchmaschinenbetreiber gesprochen (z. B. [81, 82]). Diese Forderung bleibt jedoch angesichts der Bedeutung von „Anwendungsneutralität“ kaum haltbar; selbst eine Abmilderung auf eine Forderung nach „Gleichbehandlung“ geht an der Kernidee der Suchmaschine vorbei: Kernbaustein jeder Suchmaschine ist schließlich die Erstellung gewichteter Listen – ein Vorgang, der per Definition eine Ungleichbehandlung der Eingabedaten bedeutet, bei dem die Eingabedaten das Verhalten des Algorithmus beeinflussen müssen.

### 6.1.2. Bestehende Forschungsarbeiten

Wesentliche Beachtung fand die Untersuchung von Suchfunktionen durch Eli Pariser [5]. Pariser kam zu dem Ergebnis, dass viele Suchfunktionen – insbesondere solche, die in sozialen Netzen eingebettet sind – den Nutzer in einer geschlossenen Welt halten. Die die Suchfunktion anbietenden Systeme haben durch andere Eingaben (z. B. Nutzer-Tracking via „Like-Button“) weit über die Eingabe hinausgehendes Wissen. Daneben besteht von Seite der Betreiber ein Interesse daran, dem Nutzer nur solche Inhalte zu zeigen, von denen bekannt ist, dass sie ihm gefallen werden. Das Ergebnis beschreibt Pariser als eine

[5] Wegen des den Nutzern zugerechneten Vermögens, ihres Markenbewusstseins oder ihrer letzten Einkäufe aus dem gleichen Warensegment.

[6] oder zumindest die Angst der Marketingabteilungen vor dieser angenommenen Macht

die Nutzer umgebende „Filter Bubble“, die ihn von weiteren Inhalten gleichsam abschirmt.

Der rechtswissenschaftliche Diskurs und Aktivitäten der Europäischen Union beziehen sich im wesentlichen auf den Suchmaschinenbetreiber „Google“, der die Suchmaschinenlandschaft aktuell mit Abstand dominiert; die Suchmaschine „Google“ erreichte 2012 Marktanteile von mehr als 90% (nach [9]).

Die rechtswissenschaftliche Betrachtung von Suchmaschinen hat sich (neben den Themen Urheber- und Persönlichkeitsrechte) mit der Aufnahme von Webseiten in den Suchindex befasst. In der Praxis wird von Unternehmen bei Fragen der Aufnahme in Suchmaschinenindizes primär der Weg der Suchmaschinenoptimierung oder der Unterwerfung unter die Regeln des Suchmaschinenanbieters, nicht aber der Weg einer rechtlichen Klärung gewählt [9].

Bei der Frage nach einer eventuellen Pflicht zur Aufnahme von Seiten in den Suchindex ist nach Suchmaschinenanbietern zu unterscheiden, wie aus einer Untersuchung von Ott 2007 [8] für die Aufnahme von Unternehmen in den Suchindex hervorgeht. Der Autor kam zu dem Schluss, dass sich aus dem Verbot der unbilligen Behinderung oder ungleichen Behandlung ungleicher Unternehmen nach § 20 GWB ein Recht auf Unterlassung des Ausschlusses und ggf. Schadenersatz ergeben könnte. Voraussetzung hierfür ist allerdings das Vorliegen einer Marktbeherrschenden Stellung i. S. des § 19(2) GWB. Als problematisch erwies sich die Bestimmung des relevanten Marktes, da es zwischen Suchmaschinenbetreiber und Inhalteanbietern regelmäßig nicht zu expliziten Vertragsverhältnissen kommt; das Vorliegen eines Marktes würde von weiten Teilen der Literatur daher verneint, sei jedoch nach teleologischer Auslegung denkbar. Eine richterliche Klärung der Fragestellung steht weiterhin aus.

Während sich die Betrachtungen von Ott auf die ökonomische Bedeutung von Suchmaschinen mit erheblicher Marktmacht beziehen, wurden von Danckert und Mayer 2010 die Auswirkungen von Suchmaschinen mit erheblicher Marktmacht auf die Meinungsvielfalt untersucht [7]. Die Autoren kommen zu dem Schluss, dass ein Eingreifen des Gesetzgebers notwendig sei: Dieser solle „Google“ zu den für Rundfunkanbieter obligaten Pflichten zu Erhalt und Sicherung der Meinungsvielfalt zwingen. „Google“ erfülle – so die Autoren – die Merkmale der Aktualität, Breitenwirkung und Suggestivkraft und sei somit durch eine Änderung des Rundfunkstaatsvertrages adressierbar. Ohne eine solche Regelung sehen Danckert und Mayer die Gefahr des Ausschlusses bestimmter Meinungen und die des erpresserischen Missbrauchs der Machtstellung eines einzelnen Unternehmens.

Im Mai 2012 wurde von Seiten der EU Kommission erklärt, dass im Rahmen einer nach eingegangenen Beschwerden gegen „Google“ eingeleiteten (weiter andauernden) kartellrechtlichen Prüfung vier Punkte zu beanstanden seien [83][7]:

**Bevorzugung eigener Angebote** Das vertikal integrierte Unternehmen „Google“ zeigt in

[7] Erläuterungen ergänzt durch den Autor

seiner Suchmaschine Ergebnisse, die auf seine anderen Unternehmenszweige verweisen, bevorzugt an.

**Nutzung fremder Inhalte** „Google“ bindet auf Ergebnisseiten Inhalte originär konkurrierender Anbieter ein, z. B. bei Nutzerbewertungen von Diensten oder Produkten[8].

**Suchbegriffsverwandte Werbeeinblendungen** „Google“ bietet die Platzierung von Werbeeinblendungen im Rahmen seiner Suchmaschine an, die anhand der vom Nutzer eingegebenen Suchbegriffe ausgewählt werden. Diese Praktik unterdrückt andere Anbieter entsprechender Dienste.

**Adwords** „Google“ versucht Nutzer der auktionsbasierten Werbebannerplattform „AdWords“ durch vertragliche und programmatische Regelungen an sich zu binden und den Wechsel zu konkurrierenden Anbietern zu unterbinden.

Auffällig ist, dass sich unter den von der EU Kommission bemängelten Punkten nur ein Punkt mit einer Thematik befasst, die dem öffentlich thematisierten Begriff der „Suchmaschinenneutralität“ nahe kommt. Hinsichtlich der Aufnahme oder Nichtaufnahme oder des Ausschlusses von Angeboten aus dem Suchindex finden sich keine Kritikpunkte.

## 6.2. Anwendungstransparenz

Die neben den in der Literatur untersuchten Anwendungen, bei denen die genauen Vorgänge innerhalb der Anwendung dem Nutzer verborgen bleiben obwohl sie sein Verhalten direkt oder indirekt beeinflussen, umfassen einen weiten Bereich, der sowohl Netzwerke wie auch klassische Anwendungen und Webanwendungen einschließt – so dass es kaum einem Anwender möglich ist, den Auswirkungen dieser Phänomene zu entgehen und das Verhalten von Anwendungen wirklich zu verstehen, und nicht nur zu erahnen (unter der Annahme, dass Programme wohlmeinend geschrieben wären).

Dass dies ein Problem ist, wird spätestens angesichts der Auswirkungen der Ausgabe von Anwendungen auf lebensverändernde Entscheidungen deutlich: Sei es der einfache Einkauf oder die Suche nach einem Arbeitsplatz; die Suche nach einem Partner oder nach einer Versicherung – wann immer sich der Nutzer auf eine Such- und Filterfunktion verlässt, muss er dies notgedrungen im Vertrauen darauf tun, dass die Suche oder der Filter sich erwartungsgemäß verhält. Hierbei ist eine „Erwartung“ nur schwer objektiv festzuhalten, da sie stets auch von der Vorstellung des einzelnen Nutzers über das Verhalten eines Programms abhängig ist[9].

[8] Dies ist nur scheinbar ein Widerspruch zum vorangehenden Punkt, denn kritisiert wird das Einbinden fremder Inhalte auf den „Google“-eigenen Seiten; pointiert formuliert also das **Zu-eigen-machen** fremden Inhalts.

[9] So mag ein Nutzer erwarten, dass die Ergebnisse einer Suchmaschine vom Vorkommen des Suchterms auf der Seite abhängen; ein Anderer mag erwarten, dass die Ergebnisse von der Zahl der verlinkten Seiten abhängt.

Im Folgenden wird die Allgegenwärtigkeit und damit die Bedeutung mangelnder Anwendungstransparenz anhand von Beispielen aufgezeigt. Dem schließt sich eine Gruppierung und Analyse der Beispiele an. Der folgende Unterabschnitt betrachtet die technischen Hintergründe, bevor nach denkbaren Ursachen und Motiven für die aufgezeigten Intransparenzen gesucht wird.

### 6.2.1. Praktische Beispiele

Im Rahmen dieser Arbeit wurden Beispiele für Anwendungen mit mangelnder Transparenz aus verschiedensten Gebieten des Internets recherchiert, womit gezeigt werden kann, dass Anwendungstransparenz auch jenseits von Suchmaschinen und sozialen Netzwerken die gleiche Bedeutung haben sollte. Die Ergebnisse wurden in verkürzter Form in [10] publiziert. Die Bandbreite beobachtbarer Phänomene bewegt sich von der Manipulation von Datenpaketen auf den unteren Netzwerkschichten bis hin zum Verhalten von Anwendungen, also Layer 7. Dementsprechend – und weil die Zugehörigkeit zu einem bestimmten Layer des ISO/OSI-Modells das einzig objektiv feststellbare Kriterium ist (etwa im Vergleich zur Motivation, die höchstens als mit allen bekannten und vermuteten Umständen stimmig erscheinen kann) – werden die Beispiele zunächst analog zu Netzwerklayern sortiert, beginnend mit Netzwerk- und Transport- bis zu den Anwendungsphänomenen. Insgesamt werden 9 Beispiele intransparenter Anwendungen geschildert; die detaillierte gemeinsame Analyse der Beispiele findet im darauffolgenden Abschnitt statt.

#### Beispiel 1: Zurücksetzen der Verbindung

Wenn ein Nutzer im April 2011 von Deutschland aus mit der chinesischen Suchmaschine „Baidu“ nach dem Term „Falun Gong“ zu suchen versuchte, erhielt er von seinem Browser eine Fehlermeldung, dass die Verbindung auf TCP-Ebene zurückgesetzt wurde – anstelle der erwarteten Webseite (mit den Suchergebnissen). Die detaillierten Vorgänge auf Netzwerkebene sind bereits im Abschnitt 2.3 dokumentiert.

Der Nutzer wird allein aus der Fehlermeldung nicht herauslesen können, ob es sich bei dem beobachteten Fehler um

- einen technischen Fehler auf der Seite der Suchmaschine handelt und ob die Fehlermeldung das ist, was sie zu sein scheint – die Meldung eines technischen Fehlers, der einen Verbindungsabbruch zur Folge hatte und der (zufällig) bei dieser Suche aufgetreten ist;
- eine Reaktion der Suchmaschine auf diesen Suchbegriff oder
- eine Reaktion einer Infrastrukturkomponente inner- oder außerhalb des Netzes des Suchmaschinenbetreibers auf den eingegebenen Suchbegriff handelt.

Der Nutzer kann lediglich mit Hilfe seines Kontextwissens eine Vermutung aufstellen. Wenn diese Vermutung die letztgenannte These ist, wird er gegebenenfalls sein Nutzungsverhalten anpassen, da er sich überwacht fühlt [59].

### Beispiel 2: DNS-Server fügt Einträge hinzu

Einige DNS-Server antworten nicht mit einer (im Protokoll als reguläre Antwort vorgesehenen) Fehlermeldung, wenn diese nach nicht existierenden Einträgen (nicht registrierten Domainnamen) gefragt wurde. Dies betrifft insbesondere Einträge für A-Records, also die Auflösung von Namen zu IP-Adressen. Dies soll hier am Beispiel eines von „OpenDNS“ betriebenen DNS-Servers nachvollzogen werden: Angefragt werden Daten über die Domain „inter7.jp“.
Zunächst wurde ein von „Google“ betriebener DNS-Server angefragt, dieser lieferte Daten für Mailserver, zuständige DNS-Server sowie einen SOA-Record. Im Unterscheid hierzu lieferte der von „OpenDNS“ betriebene DNS-Server auch einen Eintrag für den A-Record, also eine dem Domainnamen zugeordnete IP-Adresse. Diese IP-Adresse ließ sich selbst wiederum zu `hit-nxdomain.opendns.com` auflösen. Der Test kann im Detail in Anhang B nachvollzogen werden.

Bedeutung erlangt diese Manipulation im Alltag beispielsweise im Fall eines Vertippens im Domainnamen: Die Domain `www.deutsche-vank.de` ist nicht vergeben. Versucht ein Nutzer nun, diese Seite in seinem Browser anzeigen zu lassen, so wird der Browser beim Versuch, die Domain aufzulösen, im Fall eines nicht manipulierenden DNS-Servers einen Fehler feststellen und den Nutzer hierüber informieren. Im Falle eines manipulierenden DNS-Servers, wird der Browser eine erfolgreiche Namensauflösung vornehmen und die Anfrage an den entsprechenden – nicht vom Nutzer beabsichtigten – Server verschicken. In der Praxis wird dies regelmäßig zu einer Seite des Providers führen, die beispielsweise eine Suche nach dem nicht existierenden Domainnamen anbietet.

Denkbar wäre aber auch ein bösartiges Verhalten: Im Falle eines Tippfehlers ist es oft einfach möglich, den tatsächlich gemeinten Domainnamen zu ermitteln; anschließend könnte der falsche Domainname zu einer IP einer Phishingseite aufgelöst werden, die dem Nutzer schließlich angezeigt wird.

### Beispiel 3: Browser zeigt eine sichere Verbindung an, obwohl die Verbindung eigentlich unsicher ist

Die meisten Browser unterstützen die nachträgliche Installation von Serverzertifikaten als vertrauenswürdig; einige wie beispielsweise der „Microsoft Internet Explorer“ erlauben auch, dass Zertifikate von Administratoren durch Gruppenrichtlinien[10] (vom Nutzer unbemerkt) installiert werden [84].

[10] System- und Programmeinstellungen, die zentral verwaltet werden können

Aus der Reaktion des Browser-UIs auf eine per SSL gesicherte Verbindung lässt sich somit für den Nutzer ohne weitere Recherche[11] nicht mehr ableiten, ob eine Seite als vertrauenswürdig angezeigt wird,

- weil das Zertifikat von einer dem Browserhersteller als vertrauenswürdig bekannten Root-CA (oder einer hiervon abgeleiteten Stelle) ausgestellt wurde,
- weil das Zertifikat bereits vorher einmal durch den Nutzer als dauerhaft vertrauenswürdig akzeptiert wurde[12],
- weil das Zertifikat durch eine vom Nutzer nicht bemerkte Installation per Gruppenrichtlinie als vertrauenswürdig eingestuft wurden.

Ein solches Browserverhalten ist ungünstig, da dem Nutzer die Entscheidung über die Validität des Zertifikats (und damit auch die Vertrauenswürdigkeit der Verbindung) noch weiter erschwert wird. Bereits im „Normalzustand“ ist der Nutzer blind dem Urteil seines Browsers ausgeliefert – die Entscheidung, welche Root-Zertifikate vom Browser anerkannt werden und welche nicht, ist bereits nicht oder nur schwer nachvollziehbar.

Dennoch wird dem Nutzer vielfach (z. B. beim Online-Einkauf) suggeriert, dass eine als valide angezeigte HTTPs-Verbindung eine sichere Verbindung sei. Entsprechend vertrauensvoll wird der Nutzer mit einer entsprechend als sicher markierten Verbindung umgehen. Genau dieses Vertrauen kann z. B. von einem bösartigen Systemadministrator ausgenutzt werden, der ein eigenes Zertifikat per Gruppenrichtlinie als vertrauenswürdig installiert und im Anschluss – z. B. per Manipulation eines lokalen DNS-Servers oder per IP-Spoofing – Verbindungen zu Banken auf seine gefälschte Seite umleitet.

### Beispiel 4: E-Mail

Im Herbst 2011 wurde von Nutzern berichtet, dass E-Mails, die den Domainnamen `occupywallst.org` enthalten vom Freemail-Anbieter „Yahoo!“ nicht transportiert würden [85]. Dies wurde nach einiger Zeit durch einen „Yahoo!“-Sprecher in Form einer „Twitter“-Meldung bestätigt und mit einem Fehler im Spamfilter-System begründet [86].

Der deutsche Freemail-Anbieter „GMX“ bietet seinen Nutzern einen Spam-Filter an. Regeln, die ungewünschte Werbemails von „GMX“ auszuschließen versuchen, sind jedoch unwirksam.

[11] die entsprechende Kenntnisse über die Details von HTTPs voraussetzt, die nicht jeder Nutzer mitbringen wird

[12] Default-Einstellung z. B. in „Mozilla Firefox“

### Beispiel 5: Vorschläge von Suchbegriffen

Viele Anwendungen (offline wie online) unterstützen den Nutzer durch Anzeige von Dropdownlisten mit vorgeschlagenen Werten für eine Eingabezeile, so auch viele Suchmaschinen. Abb. 6.1 zeigte bereits beispielhaft, wie dies von der Suchmaschine „Google“ bei Eingabe des Fragments „bloc“ umgesetzt wird.

Der Nutzer erhält bei einem Klick auf den mit „Weitere Informationen“ betitelten Link einige Informationen, allerdings keine konkrete Beschreibung, weshalb exakt die angezeigten Begriffe vorgeschlagen werden. Es finden sich lediglich Verweise auf „Richtlinienverstöße“, die neben einigen offensichtlichen Kriterien (z. B. dass ein vorgeschlagener Suchterm eine gewisse Trefferzahl ergeben soll; Ausschluss von rechtswidrigen Inhalten) gelten. Diese Richtlinien werden allerdings nicht weiter ausgeführt oder verlinkt [87].

Für Aufmerksamkeit im rechtswissenschaftlichen Schrifttum [60] und in der breiten Öffentlichkeit [88, 89] sorgte der Fall Bettina Wulff, die sich mit Rechtsmitteln teilweise erfolgreich gegen Vorschläge zur automatischen Vervollständigung (die sie mit dem Rotlichtmilieu in Verbindung brachten) wehrte.

### Beispiel 6: Unklare Auswahl von Einträgen in Karten

Ähnlich obskur wie Vorschläge von Suchbegriffen ist für Nutzer die Auswahl von in Karten angezeigten Points of Interest (POI), also z. B. Geschäften. In unterschiedlichen Zoomstufen werden hierbei z. B. in „Google Maps“ unterschiedliche Mengen von POIs angezeigt. Die Notwendigkeit einer Selektion liegt angesichts des begrenzten Bildschirmplatzes auf der Hand, allerdings bleiben die vom Anbieter hierzu verwendeten Kriterien unklar. Ein Beispiel findet sich in Abb. 6.2. Im Beispiel ist jeweils ein Ausschnitt des Stadtplans von New York in der Version von „Google-Maps“ in jeweils unterschiedlichen Auflösungen zu sehen. Die geringer aufgelöste Version zeigt nur zwei Restaurants an – die höher aufgelöste Fassung hingegen deutlich mehr („PJ Clarke's“ liegt außerhalb des Ausschnitts des zweiten Bildes). Eine offensichtliche Erklärung, weshalb „Le Cirque“ bereits bei der Auflösungsstufe 1:200 und nicht erst bei einer Kartenauflösung von 1:100 eingeblendet wird, findet sich nicht.

### Beispiel 7: Länderspezifische Verfügbarkeit von Inhalten

Der Kurznachrichtendienst „Twitter“ reagierte im Frühjahr 2012 auf externe Bitten um Entfernung einzelner Nachrichten mit dem Einblenden von Informationstexten anstelle der eigentlich vom Nutzer angeforderten Mitteilungstexte. In diesen Informationstexten wird der Nutzer darauf hingewiesen, dass der jeweilige Inhalt in seinem Land nicht verfügbar sei. Die Zuordnung der Nutzer zu Ländern erfolgt durch den Nutzer selbst; im Frühjahr 2012 ließ sich das Vorenthalten von Kurzmitteilungen durch den Nutzer mittels Änderung der Landeseinstellungen umgehen.


Abbildung 6.2.: Je nach Auflösung werden von „Google Maps“ unterschiedliche POIs angezeigt; in den Ausschnitten sind jeweils alle angezeigten Restaurants hervorgehoben. Ein Auswahlkriterium ist nicht erkennbar; Bilder Stand Frühjahr 2012.

### Beispiel 8: Angabe von rechtlichen Vorgaben für Transaktionen

Online-Handelsplattformen legen oft besonderen Wert auf eine klare Darstellung des rechtlichen Status von Transaktionen gegenüber dem Nutzer (z. B. bei der Feststellung, wann ein Vertrag geschlossen wird) oder rechtlicher Voraussetzungen und Begrenzungen. Dies ist durchaus verständlich, profitieren doch alle Beteiligten (z. B. bei einer Auktion Plattformbetreiber, Käufer und Verkäufer) von einer klaren Abgrenzung. Mehr noch: Es minimiert das Risiko zivilrechtlicher Streitigkeiten für den Betreiber. Eine klare Darstellung wird gerade auf diesem Gebiet oft erreicht.

So ist es kaum verwunderlich, dass Nutzer z. B. bei der Auktionsplattform „eBay“ sehr schnell und direkt zu Antworten auch auf Fragen wie etwa „Welche Gegenstände dürfen bei eBay zum Verkauf angeboten werden?“ geleitet werden.

### Beispiel 9: Suchfunktionen

Zum Abschluss der praktischen Beispiele soll hier noch kurz mit einem atypischen und einem typischen Beispiel auf die eigentliche Kernfunktion von Suchmaschinen eingegangen werden.

Hier kann zunächst der als Beispiel 1 geschilderte Sachverhalt als atypisches Verhalten einer Suchmaschine in Abhängigkeit von der Nutzereingabe wiederholt werden: Im Frühjahr 2011 führte eine Suche nach dem Suchbegriff „falun gong“ mit der chinesischen Suchmaschine `baidu.cn` zu einem sofortigen Abbruch der Netzwerkverbindung zwischen

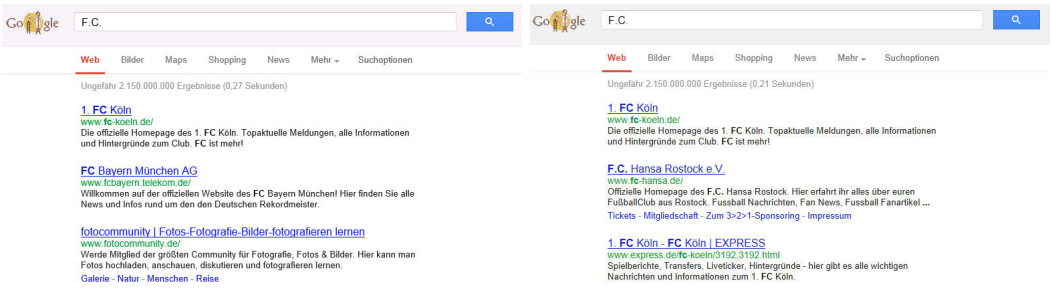


Abbildung 6.3.: Ergebnisse einer Suche nach dem Begriff „F.C.“ – links durchgeführt an einem PC in Frankfurt am Main, rechts aus dem Netz der Universität Rostock am 2013-03-25

Browser und Server.

Andere, klassischere Beispiele für die Intransparenz von Suchmaschinen lassen sich beliebig finden: Kaum ein Nutzer kann das Ranking einer Suchmaschine zu einem Suchbegriff objektiv nachvollziehen. Stattdessen können die Ergebnisse von einer Vielzahl von Faktoren abhängen:

- Dem angenommenen geographischen Ort des Anfragenden; zugeordnet auf Grund seiner öffentlichen IP-Adresse,
- seiner Suchhistorie sowie den jeweils verfolgten Ergebniseinträgen (dem Nutzerverhalten),
- weiteren über diesen Nutzer bekannten Informationen (E-Mail-Accounts, Betriebssystem, Browser, weitere bekannte Daten z. B. aus sozialen Netzen...),
- den eingegebenen Suchbegriffen.

Praktische Beispiele sind einfach aufzuzeigen – insbesondere durch Eingabe von nur mit Bezug zu einem der oben genannten Themenfeldern eindeutigen Suchbegriffen. Als Anschauungsbeispiel sei auf Abb. 6.3 verwiesen:

Gesucht wurde mit der Suchmaschine „Google“ jeweils mit einem Browser[13], der im Netzwerk der Universität Rostock angebunden und einem Browser, der in einem Rechenzentrum in Frankfurt am Main angebunden war nach dem Suchterm „F.C.“, welcher von diversen Sportvereinen als Namenspräfix genutzt wird und daher ortsbezogen unterschiedliche Ergebnisse liefern kann. In der Stichprobe zeigte sich dann ein offensichtlicher Ortsbezug: Die aus dem Rostocker Universitätsnetz durchgeführte Suche lieferte als ersten Treffer den „1. FC Köln“, als zweiten Treffer den „1. F.C. Hansa Rostock e.V.“; die Suche aus Frankfurt am Main lieferte als zweiten Treffer stattdessen „FC Bayern München

[13] in beiden Fällen dem „Microsoft Internet Explorer“

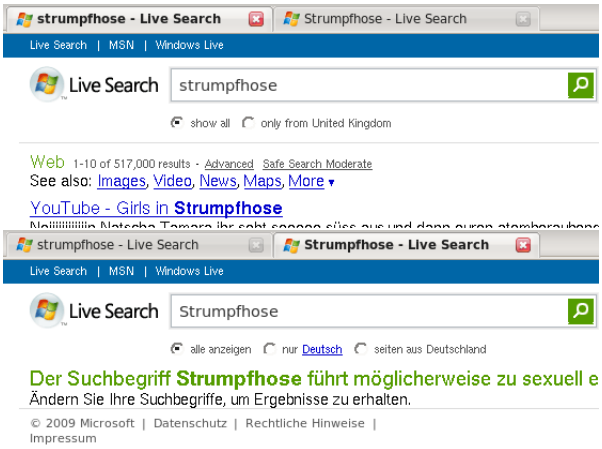


Abbildung 6.4.: Suche mit der Suchmaschine „MSN Livesearch“ nach dem Suchbegriff „Strumpfhose“ am 2009-05-06.

AG“. Interessant ist die offenbar in dieser Stichprobe ortsunabhängige Positionierung des „1. FC Köln“ als erstem Treffer.

Als abschließendes Beispiel zur Intransparenz von Suchmaschinen sei hier das Verhalten der Suchmaschine „MSN Livesearch“ (heute „Bing“) im Jahr 2009 beschrieben: Eine Suche nach dem Begriff „Strumpfhose“ endete in der deutschsprachigen Variante mit einem Warnhinweis, dass diese Suche zu sexuell eindeutigen Ergebnissen führen könnte – in der englischsprachigen Variante wurde die Suche hingegen ausgeführt, wie in Abb. 6.4 gezeigt.

### 6.2.2. Analyse und Gruppierung der Beispiele

In diesem Unterabschnitt werden die vorangegangenen Beispiele analysiert und klassifiziert. Hierzu wird ein deduktives Vorgehen gewählt. Entsprechend wird mit der Vorstellung der Klassifikation begonnen.

Die beobachteten Phänomene lassen sich in solche aufteilen, die ihren Ursprung im Aufbau des Internets haben und solche, die aus Filter- und Selektionskriterien – also der verwendeten Anwendung selbst – entstammen.

#### Netzwerkbasierte Intransparenz

Mit dem Netzwerk verbundene Intransparenzen ergeben sich aus Diensten, die vom Nutzer nicht aktiv und bewusst genutzt werden, wie etwa der Weiterleitung von Datenpa-

keten oder der Auflösung von Domainnamen. Erstere ist dabei vom Nutzer nicht aktiv beeinflussbar, da ein Source Routing (die Vorgabe eines bestimmten Routingpfades durch den Absender des Datenpaketes) aus Sicherheitsgründen nicht mehr unterstützt wird. Stattdessen werden die von Netzwerkbetreibern vorgesehenen Routen und Router verwendet. Diese allerdings können Datenpakete jederzeit blockieren oder manipulieren – im Falle des unter Abschnitt 6.2.1 beschriebenen Beispiels der chinesischen Suchmaschine Baidu lässt sich vom Standpunkt eines Benutzers in Deutschland aus nicht unterscheiden, ob die Verbindung durch einen der auf dem Weg liegenden Router oder den Anwendungsserver selbst unterbrochen wurde. Diese Nichtunterscheidbarkeit zwischen einem tatsächlich von der Gegenstelle ausgesandten Paket und einem von einer Routingstation zusätzlich eingespeisten Datenpaket folgt direkt aus der Architektur der genutzten Protokolle, die vom Wohlwollen aller beteiligten Knotenpunkte ausgehen (vgl. 2.3). Tatsächlich kann aber jeder Knoten des Routingpfades die weitergeleiteten Datenpakete beliebig manipulieren.

Die Nicht-Nachvollziehbarkeit von Routingpfaden hat sich mit der Einführung verschiedener Lastverteilungsverfahren wesentlich verschärft, da die traditionellen Testtools für Routing nicht mehr unbedingt korrekte Ergebnisse liefern. Stattdessen können bedingt durch Optimierungsentscheidungen in den Ergebnissen Lücken oder Zyklen auftreten [26].

Im Gegensatz zur Weiterleitung von Datenpaketen kann die Auflösung von Domainnamen getestet und der vom Nutzer verwendete DNS-Server gewechselt werden. Hierbei muss allerdings in Betracht gezogen werden, dass die Namensauflösung als Dienst nur den sachkundigen Nutzern bewusst ist. Der weite Teil der Internetbenutzer würde es nicht bemerken, wenn eine Domain falsch aufgelöst würde, solange der Inhalt annähernd der Erwartung entspricht [22].

Alle Phänomene, die nicht direkten Netzwerkeinflüssen entstammen, lassen sich auf Filter- und Selektionsvorgänge in den verwendeten Anwendungen zurückführen, wie im folgenden Abschnitt beschrieben.

#### Aus Filter- und Selektionsvorgängen stammende Intransparenz

Das Filtern und Selektieren von Daten ist heute ein allgegenwärtiger und unverzichtbarer Vorgang geworden. Genannt seien Suchmaschinen und Spamfilter, ohne die das Internet respektive der virtuelle Posteingang kaum noch beherrschbar wären. Andere Filter leiten Nutzer zu Produkten oder Personen, die für sie von Interesse sein könnten. Viele der heute verwendeten Filter benutzen im Hintergrund Verfahren des maschinellen Lernens.

Den oben genannten Beispielen (die nicht auf Netzwerkphänomene zurückzuführen sind) ist gemeinsam, dass Nutzer auf Filter- und Selektionssysteme vertrauen, ohne deren exakte Funktionsweise zu kennen. Sei es, weil sie keinen Zugriff auf haben oder sei es, weil sie ihnen unverständlich sind.

Das Problem wird weiter dadurch verschärft, dass viele Webanwendungen dazu führen, dass der Anwendungsentwickler oder -anbieter nicht nur die Hoheit über das ausgeführte Programm sondern auch die vom Nutzer eingegebenen Daten er- oder behält; der Nutzer somit keinerlei Kontrollmöglichkeit besitzt.

In diesem Sinne lassen sich die oben angegeben Beispiele leicht auf Filtersysteme (im Fall von Spamfiltern und der landesabhängigen Verfügbarkeit von Inhalten) und Selektionsmechanismen (im Fall von Suchbegriffsvorschlägen, der Auswahl von POIs in Karten, Artikelvorschlägen und Suchergebnissen) zurückführen. Einzig das Szenario als glaubwürdig eingestufter unglaubwürdiger HTTP-Verbindungen scheint auf den ersten Blick mehr Netzwerk- als Selektions- und Filterbasiert. Im Detail betrachtet handelt es sich jedoch um einen Fall einer Filterintransparenz: Der im Browser befindliche Filter, welcher Zertifikate als vertrauenswürdig – oder nicht – einstuft, ist intransparent. Diese Intransparenz verschärft sich weiter, wenn es einem (Remote-) Administrator ohne Wissen des Nutzers möglich ist, weitere Zertifikate als vertrauenswürdig zu installieren und so Warnungen vor potentiell unsicheren Verbindungen zu unterdrücken.

### 6.2.3. Ursachen und Motivation

In diesem Abschnitt werden einige Überlegungen dargelegt, weshalb sich Anbieter von Anwendungen für die intransparente Gestaltung von Anwendungen entscheiden könnten. Dabei sollen ökonomische, politische und rechtliche wie auch soziale Gründe beleuchtet werden. Den oben genannten Beispielen lassen sich intuitiv plausibel erscheinende Motive zuordnen. Da jedoch keine objektiven Belege vorliegen, soll im Rahmen dieser Arbeit von einer konkreten Zuordnung Abstand genommen werden.

#### Ökonomische Gründe

Ein großer Teil der öffentlich zugänglichen Webanwendungen finanziert sich aus Werbeeinnahmen. Werbeeinnahmen ergeben sich aus der (ggf. indirekten, vermittelten) Kooperation mit Unternehmen. Eine solche Kooperation schafft allerdings weitere Interessen: Die Unternehmen werden in negativen Kontexten nicht genannt werden wollen.
Werbende Unternehmen zahlen für Nutzeraufmerksamkeit – gemessen entweder in Seitenaufrufen oder in Klicks auf Anzeigen. Hieraus erwächst für den Seitenbetreiber wiederum das Interesse, Nutzer auf seiner Seite zu halten. Hierzu werden oftmals Algorithmen verwandt, die aus dem bisherigen Nutzerverhalten auf dem Nutzer gefällige Inhalte schließen, um ihm stets vermeintlich attraktive, weil als beliebt bekannten Inhalten ähnliche, Angebote unterbreiten zu können [5]. Dies wiederum hält den Nutzer – so die Betreiberhoffnung – auf der Seite und führt so wiederum zu Einnahmen durch das Anzeigen von Webeeinblendungen.

Nicht zuletzt kann sich das Verhältnis zwischen Inhalteanbieter und Inserent in der Form verschieben, dass der Inhalteanbieter von den Zahlungen des Werbekunden abhängig ist. Im Print-Bereich werden die Anzeigen von Unternehmen durchaus zur Einflussnahme auf die Berichterstattung genutzt – beispielsweise, um die Veröffentlichung kritischer Artikel zu unterdrücken [90].

In vertikal integrierten Unternehmen gibt es eine weitere Motivation für die Beeinflussung von Selektions- und Filtervorgängen: Die Hervorhebung eigener Dienste, um Nutzer zu diesen Diensten – und nicht zur Konkurrenz – zu führen. Ein Beispiel wäre die prominente Platzierung von vom Suchmaschinenbetreiber angebotenen Diensten in Ergebnissen seiner Suchmaschine. Ähnliche Konstrukte sind auch bezüglich der (präventiven) Unterdrückung von (existenten) Konkurrenzangeboten geplanter Dienste denkbar.

### Politische und rechtliche Gründe

Weitere Gründe ergeben sich aus politischen und rechtlichen Gegebenheiten: Bestimmte Gesetzgebungen reglementieren die Anzeige bestimmter Inhalte – in der Bundesrepublik Deutschland ist z. B. die Abbildung von militärischem Gelände durch Luftaufnahmen mit Strafe bedroht (§ 109g StGB). Derartige gesetzliche Regelungen äußern sich in entsprechendem Verhalten von Anwendungen, wie z. B. in Abb. 6.5 dargestellt.

Weitere Regelungen bestehen in der Bundesrepublik Deutschland seit dem 1. August 2012 bezüglich des Abschlusses von (Kauf-) Verträgen. Mit § 312g (3) BGB wurde eine gewisse Transparenz von Webanwendungen bezüglich der Kennzeichnung von Buttons gesetzlich vorgeschrieben: Der Vertragsschluss muss durch Betätigung eines Buttons mit der Aufschrift „Zahlungspflichtig bestellen“ (gesetzlicher Vorschlag) oder vergleichbar eindeutiger Formulierung erfolgen.

Weitere Vorgaben können sich aus kulturellen oder systemischen Vorgaben ergeben – in einem Ein-Parteien-System müssen sich Anbieter anderen Regelungen unterwerfen als in pluralistischen Demokratien. [91] beschreibt das Vorgehen des Suchmaschinenbetreibers „Google“ bei seiner Expansion nach China.

Anbieter werden sich gegebenenfalls im Vorfeld einer Expansion anpassen, indem Algorithmen so verändert werden, dass sie keine problematischen Ausgaben mehr produzieren können.

### Soziale Gründe

Die letzte Klasse von Gründen für Intransparenzen von Anwendungen liegen im sozialen Bereich. Die meisten Inhalteanbieter streben danach, ein gutes Image zu bewahren. Kein Anbieter hat das Verlangen, dass seine Suchmaschine öffentlich als „Bestes Werkzeug für

Abbildung 6.5.: Beispiel für rechtliche Gründe als Hintergrund eines Anwendungsverhaltens: Teilweise nur grob aufgelöste Anzeige von Satellitenaufnahmen, hier am Beispiel einer Bundeswehrliegenschaft in Eckernförde, Schleswig-Holstein, in „Bing Maps“ (Stand Winter 2014; Koordinaten 54.48546, 9.80096).

die Suche nach Kinderpornographie“ bewertet wird. Entsprechend werden Algorithmen beeinflusst, um bestimmte – gesellschaftlich als anstößig wahrgenommene – Ergebnisse zu unterdrücken.

Aus sozialen Gründen getriebene Eingriffe scheinen von Betreibern oftmals als Balanceakt ausgeführt: Einerseits sollen möglichst alle unerwünschten Inhalte oder Funktionen unterdrückt werden; andererseits werden empörte Nutzerreaktionen bei einem zu invasiven Vorgehen (ein Verhalten wie in Abb. 6.4 dokumentiert kann leicht als solches empfunden werden) gefürchtet.

### 6.2.4. Gestaltung transparenter Anwendungen

Als Beitrag zur Lösung der aus Intransparenz von Anwendungen entstehenden Probleme wird im folgenden Unterabschnitt eine für den Anwender transparentere Gestaltung von Anwendungsoberflächen vorgeschlagen. Den aus netzwerkbasierter Intransparenz folgenden Probleme kann mit den im Kapitel 5 beschriebenen Mitteln begegnet werden. Die hier vorgeschlagenen Methoden adressieren entsprechend die aus Such- und Filtervorgängen resultierende Intransparenz. Den Ausgangspunkt bildet die reine Information, am Ende wird die Beeinflussung des Anwendungsverhaltens durch den Nutzer stehen. Zur Anschauung werden Mockups (also rein graphische Gestaltungsvorschläge) eines Kartendienstes verwendet.

Die Schritte sind aufeinander aufbauend konzipiert und bedingen einander. Eine Einflussnahme des Nutzers ohne Information, worauf er Einfluss nimmt, ist im besten Fall sinnlos.

#### Information

Der erste Schritt auf dem Weg zu einer transparenten Anwendung besteht aus der Information der Nutzer, dass z. B. eine angezeigte Ergebnismenge überhaupt einer Beschränkung unterliegt. Ein klassisches Beispiel sind Suchergebnisse: Es wird zwar nur eine bestimmte Zahl von Ergebnissen auf der ersten Seite angezeigt, aber bereits die Angabe der totalen Anzahl der Ergebnisse zusammen mit der Verlinkung weiterer Ergebnisseiten verdeutlicht, dass es weitere Ergebnisse gibt. Über die Gründe für die Beschränkung der angezeigten Ergebnisse wird unterdessen keine Angabe gemacht. Diese würde im Fall einer Suchmaschine einen Hinweis auf die vom Rankingalgorithmus verwendeten Gewichtungen bedeuten.

Die in diesem Schritt angeregte reine Information über die Tatsache der Beschränkung der Ausgabemenge kann beispielsweise in Form einer Einblendung erfolgen.

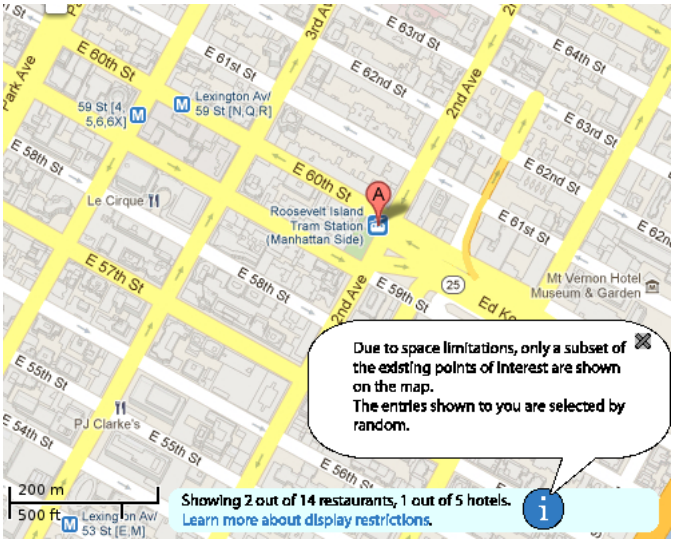


Abbildung 6.6.: Vorschlag einer modustransparenten Oberflächengestaltung.

### Modus

Der zweite Schritt auf dem Weg zur transparenten Anwendung ist die Information über die Umstände einer Beschränkung der Ausgabe. Für einen Kartendienst kann eine solche Ausgabe z. B. wie in Abb. 6.6 dargestellt realisiert werden. Der Nutzer erhält die Information, dass eine Beschränkung der Ausgabe geschieht (Forderung aus Schritt 1), ferner ist der Grund angegeben und es besteht die Option, weitere Details hierzu in einer gesonderten Anzeigeform (eigener Seite) zu erhalten. Eine solche Umsetzung ist ebenfalls mit geringem Aufwand realisierbar, da es sich – aus Programmsicht – nur um weitere Ausgaben handelt.

### Einflussnahme des Nutzers

Der letzte Schritt auf dem Weg zur Transparenz ist die Einbeziehung des Nutzers in den Filterprozess. In Abb. 6.7 ist eine denkbare Umsetzung für das Beispiel einer Kartenanwendung vorgestellt. In einer idealen Umsetzung kann der Nutzer selbst bestimmen, welche Parameter welchen Einfluss auf einen Auswahlprozess haben. Zwar kann ein Nutzer nun immer noch durch voreingestellte Default-Werte beeinflusst werden. Doch bereits die in Schritt 1 und 2 vorgeschlagenen Maßnahmen sorgen dafür, dass dem Nutzer diese Voreinstellungen wie auch die Möglichkeit, ihrer Veränderung bewusst wird. Ein weiterer Zugewinn bestünde aus der hinzukommenden weiteren Verfeinerung von Suchanfragen: Aus der abstrakten Suche z. B. nach „Restaurant“ wird so eine deutlich präzisere Suche nach „Restaurant mit Raucherbereich, 7±3 Tische, klassisch italienische Küche“. Oder aber „Restaurants, die mehr als 300 Euro an den Kartendienstanbieter bezahlt haben,

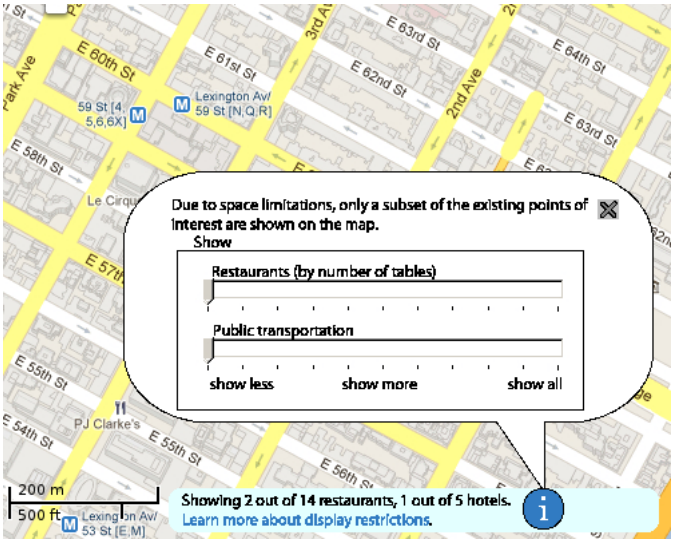


Abbildung 6.7.: Vorschlag für eine Oberfläche, in der Nutzer die Anzeigefilter beeinflussen können.

um eine prominente Platzierung zu erhalten“.

Eine Umsetzung würde deutlich stärkere Eingriffe in die zu Grunde liegende Anwendung bedeuten: Zuvor von Redaktions- oder gar Kommandozeilenumgebungen aus zugängliche ebenso wie nur in Algorithmen eingebettete Parameter müssten von der Nutzeroberfläche her zugreif- und für die einzelne Nutzersitzung änderbar gestaltet werden.

## 6.3. Bewertung des Vorschlags hinsichtlich der Umsetzbarkeit

Die Grenzen transparenter Anwendungen aus Anwendersicht sind leicht zu finden: Man stelle sich Abb. 6.7 mit weiteren Schiebereglern oder Checkboxen für öffentlichen Personennahverkehr, Bibliotheken, Einkaufszentren, Clubs, Bars, Kultureinrichtungen (nach Sparte), Sportclubs, medizinischen Einrichtungen (nach Fachrichtung), Verwaltungseinrichtungen, Tierhandlungen, Spielwaren- und Buchläden vor, inklusive deren Ausprägungen. Somit ergibt sich aus der Benutzbarkeit der Anwendung eine natürliche Gegenposition, da sich aus der Forderung nach Einstellbarkeit aller Parameter leicht eine beliebig unübersichtliche Oberfläche ergeben kann. Dennoch bleibt die Anwendersicht wichtigstes Kriterium für die Transparenz einer Anwendung. Wenn eine Anwendung in ihrem Verhalten transparent ist, dann ist der Benutzer jederzeit in der Lage herauszufinden, welche Eingaben die aktuelle Ausgabe beeinflussen.

Der Einfluss einer übersichtlichen Nutzeroberfläche einer Anwendung auf ihre Benutzbarkeit kann hierbei indes nicht überschätzt werden – ist dies doch Teil des Geschäftsgeheimnisses hinter vielen Webanwendungen. Zu einer Zeit, als Suchmaschinen dem Nutzer noch mit komplexen Formularen zur Zusammenstellung von Datenbankanfragen begegneten, war die Einfachheit der Oberfläche von „Google“ revolutionär: Eine Eingabezeile, zwei Buttons. Dabei blieben die weiteren Details des Suchvorganges unklar. Hiermit scheint der Nutzer jedoch einverstanden, so lange die Ergebnisse zufriedenstellend sind.

Ebenfalls darf der Aufwand bei der Modifikation bestehender Anwendungen nicht unterschätzt werden – kaum ein Unternehmen wird viele Mannstunden investieren, um ein Nutzerinterface zu erzeugen, welches später vielleicht weniger Zuspruch unter den Nutzern finden wird.

Insgesamt ergibt sich also die Einschätzung, dass eine Umsetzung der ersten beiden Schritte (Information des Nutzers über Anzeigebeschränkungen und deren Gründe) im Kontext der meisten Anwendungen erreichbar ist. Eine Umsetzung des dritten Schrittes, also der Beeinflussung der Ergebnisfilter durch den Nutzer selbst hingegen erscheint aus Gründen der Interfacegestaltung und der notwendigen Veränderungen am Programm kaum durchführbar.

Aus juristischer Sicht sind die hier hervorgehobenen Eigenschaften von Webanwendungen jenseits eindeutiger Rechtsverstöße nicht fassbar, da es formal an einer Rechtsgrundlage mangelt, die den Anbieter einer Webanwendung zur Transparenz verpflichtete.
Weiterhin zeigen sowohl der Blick auf das Vorgehen von „eBay“ (hier werden rechtliche Bedingungen dem Nutzer leicht erreichbar zur Verfügung gestellt) oder von Anbietern von Satellitenbildern (welche den Blick auf militärisches Sperrgebiet gemäß gesetzlicher Vorgabe verhindern), dass auch nicht transparent gemachtes Verhalten einer Anwendung gesetzlichen Vorgaben entspringen kann.

Im vergangenen Kapitel konnte aufgezeigt werden, dass viele internetbasierte Anwendungen in ihrem Verhalten für den Nutzer nicht transparent sind. Dieses Verhalten begründet sich aus (nicht ohne tiefgreifende Eingriffe in die bestehende Infrastruktur änderbaren) Eigenschaften und Implementationsdetails von Computernetzwerken sowie programmimmanenten Intransparenzen. Letzteren kann mit dem hier vorgestellten Verfahren begegnet werden.

Eine genauere Untersuchung, wie transparente Anwendungen zu gestalten sind, damit Nutzer wie auch die Funktionalität der Anwendung nicht zu kurz kommen soll hier als Ausblick nur skizziert werden: Eine solche Untersuchung hätte die schwierige Aufgabe, das Wissen eines Anwenders über die Abläufe innerhalb einer Anwendung zu beurteilen. Hierzu böten sich zwei Wege an: Zum einen die „klassische“ Befragung mit Hilfe von Fragebögen; zum anderen die Gestaltung einer Testanwendung, deren Benutzung ein möglichst umfassendes Verständnis innerer Vorgänge voraussetzt. Eine Anwendung, deren innere Funktionen bekannt sein muss, ist in jedem Fall notwendig.

Ebenso unsicher ist eine Betrachtung der Frage, wie Anwendungsanbieter und -gestalter überzeugt werden können. Trivial mutet die Verpflichtung durch eine gesetzliche Regelung an; jedoch bleibt die Frage, welche Rechtsgrundlage eine solche Verpflichtung haben sollte. Die Verpflichtung im Zuge eines Vertragsabschlusses ist vergleichsweise einfach zu begründen; bei einer Vielzahl von Online-Anwendungen (Suchmaschinen, Kartendienste, ...) hingegen gibt es hingegen zunächst einmal gar keine vertragliche Bindung zwischen Anwendungsanbieter und Nutzer. Realistischer Anreiz könnte hingegen das Nutzerverhalten sein: Würden Nutzer transparente Angebote (also solche, bei denen Werbeeinblendungen unter Ergebnissen eindeutig gekennzeichnet sind; bei denen Algorithmen und Filter nachvollziehbar sind) honorieren und intransparente Angebote durch Missachtung abstrafen, ergäbe sich ein ökonomisches Motiv, transparent zu werden.

Das zweite Einführungsbeispiel macht allerdings deutlich, dass auch eine ideale und komplette Transparenz von Anwendungen nicht in jedem Fall helfen kann: Wenn es ein Problem auf einer für die Anwendung nicht sichtbaren Ebene gibt (wie hier eben ein Netzwerkproblem, dessen einziges Symptom das „Stocken“ der Verbindung ist), wird eine Anwendung zunächst keine Hilfe leisten können; selbst wenn jedes eingehende Datenpaket dem Nutzer visualisiert würde – das Ausbleiben weiterer Datenpakete nach bestimmten Bytesequenzen ist einfach zu ungewöhnlich; zumindest in Netzen in der Bundesrepublik Deutschland.

# 7. Abschluss

Statt zu klagen, dass wir nicht alles haben, was wir wollen,
sollten wir dankbar sein, dass wir nicht alles bekommen, was wir verdienen.
—*Dieter Hildebrand*

In diesem Kapitel werden zunächst die Ergebnisse der Arbeit zusammengefasst; im zweiten Abschnitt findet sich eine Bewertung dieser Ergebnisse; schließlich werden ausgewählte weitergehende Forschungsfragen erörtert.

## 7.1. Ergebnisse

Zunächst wurde das Phänomen der Netzneutralität als Eigenschaft eines durch die alleinige Anwendung von Best-Effort gekennzeichneten Netzwerks untersucht. In diesem Zusammenhang wurden in Kapitel 4 die sich aus dieser Definition ergebenden Probleme analysiert und systematisch untersucht. Diese Untersuchung kam zu dem Ergebnis, dass durch technische Neuerungen und eine veränderte Nutzung des Internetanschlusses eine klare Unterscheidung zwischen neutralen und nichtneutralen Netzen zunehmend erschwert wird und nur durch Anpassung der Definition möglich ist, wenn „Netzneutralität“ weiterhin jeden Eingriff des Netzbetreibers beschreiben soll. Ferner wurden der in der Bundesrepublik Deutschland bestehende rechtliche Rahmen des § 41a TKG sowie der aktuell vorliegende Entwurf einer Verordnung der EU unter technischen Gesichtspunkten untersucht. Beidem liegt nicht die Idee einer klassischen Netzneutralität sondern die Akzeptanz existierender „Managed Services“ zugrunde; geregelt und geschützt werden soll allerdings die weitere Existenz eines „Best-Effort-Bereichs“ neben privilegierten Diensten. Schließlich wurden bestehende Verfahren zum Nachweis von Neutralitätsverletzungen beschrieben, klassifiziert und bewertet. Hierbei wurde bewusst zugunsten eines qualitativen Vergleichs unterschiedlicher Messmethoden auf einen quantitativen Vergleich verzichtet; Ergebnisse eines quantitativen Vergleichs wären ohnehin auf Grund der stark unterschiedlichen Aussagen der jeweiligen Testverfahren (die in Kap. 4 ausführlich erörtert werden) nur von geringem Nutzen.

Nach der Betrachtung des Forschungsstandes wurde im ersten Teil von Kapitel 5 der Blick auf die Netzneutralität aus der dualen neutral-nichtneutral-Klassifikation auf eine andere Ebene verschoben, indem der Fokus von der Neutralität auf die Transparenz als übergeordnetes Konzept des Wissens um Neutralität gewechselt wurde. Hierzu wurden die Anforderungen an eine über Netzbetreibergrenzen hinweg konsistente Deklaration von Regeln[1] zur Behandlung einzelner Datenpakete oder -ströme detailliert ausgeführt. Diese

[1] im Sinne eines „Beipackzettels“ zum Netzwerk

Anforderungen berücksichtigen, dass sich wohl kaum ein Netzbetreiber in „seine Karten“ – in diesem Fall die konkrete Routerkonfiguration – schauen lassen wird. Stattdessen sind die Anforderungen so gestaltet, dass einzelne, nach Betreiber zusammengefasste Teilnetze, gleich einer Black-Box, ausschließlich in Form der von außen beobachtbaren Phänomene umschrieben werden. Zusammen mit dieser Beschreibung des Netzverhaltens (im Sinne eines vom Netzbetreiber behaupteten Soll-Zustands hinsichtlich der Auswirkungen auf den Nutzer) würden auch weitere – für den Nutzer interessante – Angaben über Ursachen beobachtbaren Verhaltens übermittelt; solche Angaben betreffen beispielsweise Kontingente. Anhand dieses Dualismus zeigen sich bereits unterschiedliche Forderungen an eine entsprechende Deklaration: Einerseits sollen technische Verhalten überprüfbar deklariert werden; andererseits soll dem Nutzer auch die Motivation der jeweiligen Regel – so vorhanden – angegeben werden. Solche Angaben können ferner gesetzliche Auflagen zur Weitergabe von Gründen für bestimmte Netzbeschränkungen erfüllen.

Anhand der in der Spezifikation enthaltenen Beschreibung kann in einem nächsten Schritt die Einhaltung dieser Deklaration getestet werden, indem versucht wird, eine Abweichung des beobachtbaren vom spezifizierten Verhalten aufzuzeigen. Hierzu wurde im zweiten Teil von Kapitel 5 ein Verfahren auf Basis von im untersuchten Netzwerk enthaltenen Messstationen vorgestellt. Dieses Verfahren erlaubt die Verwendung beliebiger Protokolle zu beliebigen Zielservern; mithin die Untersuchung aller Verbindungen während der „normalen“ Verwendung des Internetanschlusses durch den Nutzer. Dadurch ergibt sich eine Aussage, die aus Sicht des Nutzers mit einer Beschreibung „des gesamten Internets“ identisch ist. Hierzu sind Broker notwendig, die in den an einem Datenaustausch beteiligten Teilnetzen Pakete mit beliebiger Absenderadresse verschicken können müssen.

Kapitel 6 betrachtet schließlich auch Anwendungen aus dem Blickwinkel der Transparenz. Hierzu wurden über einen langen Zeitraum unterschiedlichste Artefakte mangelnder Transparenz zu einem Gesamtbild zusammengetragen. Als Ergebnis lässt sich festhalten, dass fehlende Anwendungstransparenz ein – gerade unter Webanwendungen – weit verbreitetes Phänomen ist. Drei Motive erscheinen für die Intransparenz von Anwendungen plausibel: Ökonomische Motive, soziale Motive und rechtliche Motive. Als Ergebnis der Untersuchung wurden Vorschläge zur transparenten Entwicklung von Anwendungen vorgestellt, die hinsichtlich ihrer Umsetzbarkeit bewertet wurden.

## 7.2. Bewertung

Die Frage nach einer praxistauglichen Definition der Netzneutralität ist ein kontroverses Thema: Auch der in Abschnitt 4 unterbreitete Vorschlag einer Neutralitätsdefinition, die Dienste mit erhöhten Netzanforderungen wie IP-TV oder VoIP berücksichtigt, wird dem Puristen zu weit gehen. Ein Netz ist neutral, wenn alle Komponenten nach dem Best-Effort-Prinzip arbeiten – das ist die Position des Puristen; für ihn ist jede Abweichung

eine Verletzung der Netzneutralität.

Die Schaffung einheitlicher Spezifikationen des Netzwerkverhaltens erzwingt eine regelmäßige Anpassung der veröffentlichten Spezifikation an die tatsächliche Netzwerkkonfiguration und ist mit einem entsprechenden Mehraufwand für den Netzbetreiber verbunden. Wesentlicher Nachteil einer solchen Umschreibung ist eine gewisse Unschärfe, die allerdings auch mit den von außen durchführbaren Beobachtungen korreliert, denn einem Nutzer wird der Einblick in die Details eines Netzwerks üblicherweise ebenfalls verborgen bleiben. Darüber hinaus werden Netzbetreiber mit Aufstellung und Betrieb der Broker belastet – die Fragen eines finanziellen Ausgleichs für diesen Aufwand wurden im Rahmen dieser Arbeit nicht betrachtet; dies sind bei einer Umsetzung absehbare Streitpunkte. Auch die Frage, wie exakt mit einer mit der vorgestellten Infrastruktur festgestellten Verletzung der eigenen Spezifikation umgegangen werden soll, ist nur partiell zu beantworten: Zwar findet sich leicht die Aussage, dass der Netzanbieter in diesem Moment ja vertragsbrüchig sei, da er seine Zusicherung nicht einhalte. Hierbei bleiben jedoch zwei blinde Flecken: Zum einen sind es nur die ISP, die mit einem Endnutzer einen Vertrag geschlossen haben. Zum anderen kann kaum sichergestellt werden, dass sich ein Broker innerhalb der Netzwerktopologie an einer bestimmten Position befindet; die genaue Position hat jedoch erheblichen Einfluss auf die getroffene Aussage. Schlimmstenfalls bezieht sich eine Messung mithilfe eines Brokers im Netzwerk $n$ auf Netzwerkeffekte von Teilen des Netzes $n-1$ und Teilen des Netzes $n$. Eine Aussage wird sich bei ungünstiger Broker-Positionierung entsprechend auf einen Netzabschnitt beziehen, der von zwei Betreibern anteilig verwaltet wird.

Die als Konsequenz aus der Untersuchung der Transparenz von Anwendungen vorgeschlagenen Designempfehlungen sind für sich betrachtet plausibel; eine Umsetzung wird in der Praxis eines Unternehmens jedoch aus mehreren Gründen schwerfallen: Zunächst sind gewisse Intransparenzen auch Geschäftsmodell. Wie genau der der Ranking-Algorithmus einer Suchmaschine funktioniert, ist wesentlicher Erfolgsfaktor des betreibenden Unternehmens. Entsprechend unwillig werden Unternehmen sein, wenn es darum geht, ebendiese Mechanismen offenzulegen. Das Sichtbarmachen erkaufter guter Ergebnisse würde der Intention des Kaufs guter Positionierungen in Suchergebnissen negieren; der „Kunde“ zahlt ja gerade für den Eindruck, in einem scheinbar ausgewogenen und offenen Prozess als objektiv Bester gewonnen zu haben.

Schwierig umsetzbar ist auch die Forderung nach absolut transparenter Gestaltung von Anwendungsoberflächen, droht eine absolut transparente Oberfläche doch schnell auch zu einer (auf Grund ihrer Überfrachtung mit Eingabemöglichkeiten) absolut unbedienbaren Oberfläche zu werden.

## 7.3. Weiterführende Forschungsfragen

In der Einführung wurde ausgeführt, in welch aktiver und anregender interdisziplinärer Forschungslandschaft sich das Thema Netzneutralität bewegt. Entsprechend wäre es

leicht, Dutzende weiterführender Forschungsfragen zu formulieren. Dennoch soll sich dieser Abschnitt auf vier Forschungsfragen begrenzen, die im Folgenden kurz andiskutiert werden und die sich thematisch jeweils ein Stück weiter von dieser Arbeit hin zur Interdisziplinarität bewegen.

### Eindeutige Zuordnung von Tests zu Providern

Das in Kap. 5 beschriebene Verfahren kann einen möglichen Regelverstoß nicht eindeutig einem Netzbetreiber zuordnen, da hierzu die Position der Broker in den jeweiligen Netzen sehr exakt vorgegeben (und bei der Auswertung entsprechend bekannt) sein müsste. Diese Einschränkung zu umgehen wäre wünschenswert, weil sich damit eine Beobachtung exakter einem einzelnen Teilnetz zuordnen lassen würde.

Verwirft man die triviale Lösung (bei der die Netzbetreiber zu einer bestimmten Anbindung eines Brokers verpflichtet würden), ergibt sich ein für einen einzelnen angebundenen Nutzer nicht lösbares Problem. Kern des Problems ist die Unfähigkeit des Nutzers, den Routingpfad eines von ihm versandten Datenpaketes zu beeinflussen, in diesem Fall mit dem Ziel, Eigenschaften unterschiedlicher Routen miteinander vergleichen zu können.

Eine – als Ansatz eines Forschungsvorhabens geeignet scheinende Lösung – könnte jedoch in der gemeinsamen Auswertung der Messergebnisse mehrerer Messclient liegen. In diesem Fall könnten die disjunkten und gemeinsamen Teile der Routinggraphen mehrerer Messclients genutzt werden, um Beobachtungen näher einzugrenzen.

Auf die Beschreibung eines solchen Ansatzes wurde in dieser Arbeit verzichtet – aus Gründen der Übersichtlichkeit aber auch aus Datenschutzgründen, denn in der Umsetzung eines solchen Ansatzes würden sich die Datenschutzprobleme stellen, die bereits bei der Bewertung des passiven Ansatzes zum Nachweis von Neutralitätsverletzungen in Kap. 4 beschrieben wurden.

### Neutralitätszusicherung per Protokoll

Als zweite weiterführende Forschungsfrage soll die Idee eines neutralitätssichernden Netzwerkprotokolls skizziert werden.

Erfolgversprechendster Ansatz für eine solche Lösung wäre das Verhindern einer erfolgreichen Datenstromidentifikation. Verschlüsselung des übertragenen Inhalts und verschleierndes Multi-Hop-Routing sind Ausgangspunkte. Da Pakete jedoch immer dem Endnutzer auf Grund seiner IP-Adresse zugestellt werden, sind sie spätestens an diesem Punkt auch einer Shallow-/Deep-Packet-Inspection und nachfolgenden Neutralitätsverstößen auf Grund wenigstens der erkennbaren Kommunikationsumstände (Absender, Empfänger, ggf. Layer-4-Protokoll/–Adressierung) zugänglich.

Durch eine für den Netzbetreiber unkenntliche Adressierung würde ihm die Erfolgsaussicht für eine Shallow-/Deep-Packet-Inspection genommen (wenn alle weiteren übertragenen Informationen verschlüsselt sind) und ihm bliebe nur noch die statistische Pro-

tokollidentifikation. Dieser müsste durch Einstreuen von mit zufälligem Inhalt gefüllten Datenpaketen in einen Strom verschlüsselter Datenpakete begegnet werden, so dass der Provider lediglich einen ständig konstanten Datenstrom beobachtet, der es ihm unmöglich macht, sinnlose von sinntragenden Paketen zu unterscheiden.

Ergebnis einer Umsetzung wäre die Einführung einer neuen Adressierung auf Ebene 3 des ISO-/OSI-Modells (also in Konkurrenz zum Internet Protocol). Eine solche Umsetzung würde nur geringe Veränderungen an bestehenden Anwendungen[2] erfordern.

Als konkrete Vorstellung soll eine Variation des Token-Ring-Netzwerks dienen: Dem im Netzwerk kreisenden Token werden beliebige Datenpakete angefügt. Die Datenpakete sind komplett verschlüsselt – der Empfänger erkennt an ihm gerichtete Pakete dadurch, dass er sie erfolgreich entschlüsseln kann (Prüfsumme). Um einer statistischen Identifikation zuvorzukommen, wird dafür gesorgt, dass dem Token stets $n$ Datenpakete folgen – wenn keine echten Pakete enthalten sind, werden Zufallsdaten verwendet. Einzig notwendige Sicherungsmaßnahme wäre das regelmäßige Entfernen von Datenpaketen, die eine „Runde mitgefahren“ sind, wahlweise im Tausch gegen ein „echtes“ oder ein „auffüllendes“ Paket.

### Kontraproduktive Anwendungstransparenz

Auch aus der Anwendungstransparenz ergeben sich spannende weiterführende Fragestellungen, zum Beispiel auf dem Gebiet der Mensch-Computer-Interaktion. Besonders interessant hinsichtlich der Entwicklung von Anwendungen scheint die Suche nach einem Optimum bei der Oberflächengestaltung zwischen den Extrema der unbedienbaren Transparenz (auf Grund der Komplexität von Eingabemasken) und intuitiver Intransparenz. Wegweisend könnte hier eine genaue Analyse der Entwicklung von Suchmaschinenoberflächen sein. Mitte der 1990er Jahre kamen Suchmaschinen als Hilfsmittel zur Navigation im World Wide Web auf, sie entwickelten sich aus Datenbankabfragen und hatten Nutzerinterfaces, die nicht weit von einem SQL-Assistenten entfernt waren. Nutzer konnten viele Parameter der Suche beeinflussen und sehr präzise Ausdrücke eingeben. Doch den Durchbruch machte eine Suchmaschine, die ohne diese Komplexität dem Nutzer sinnvoll erscheinende Ergebnisse lieferte.

Um zu untersuchen, welcher Grad der Transparenz vom Nutzer honoriert wird – und ab welchem Grad der Nutzer verschreckt wird – ließe sich eine entsprechende Studie mit Mock-Ups von gestalten. Ebenso aufschlussreich könnten Experteninterviews mit Entwicklern von Oberflächen bekannter und verbreiteter Programme sein, da sie einen Einblick in (vielleicht rein intuitive) Designentscheidungen bieten würden.

### Wen interessiert Netzneutralität?

Als letzte weiterführende Forschungsfrage soll hier zum Schluss das defätistisch erscheinende „Wen interessiert das überhaupt?“ erwähnt werden. Wie viele Nutzer interessieren

[2] sondern nur an Betriebssystemen

sich für Netzneutralität? Weshalb? Diese Fragestellung geht weit über den Rahmen dieser Arbeit hinaus, kratzt sie doch zugleich auch an der Frage, welche Relevanz das Internet mit seinen „Digital Natives“ für die gesellschaftliche Realität besitzt. Einer Facebook-Gruppe beizutreten ist einfach; eine Online-Petition schnell unterschrieben. Auf eine Demonstration – unter gleißender Sonne – zu gehen, kostet hingegen Überwindung und Zeit. Online und anonym für eine unbequeme Meinung einzustehen oder eine unbequeme Frage zu stellen ist leicht; sie aggressiv zu vermarkten ebenso. Für eine unbequeme Meinung eine Konfrontation mit Mitbürgern oder dem Staat zu riskieren, ist es nicht.

Wen also interessiert die Netzneutralität? Diejenigen, die das Netz als Quelle stets der neuesten Filme aus Hollywood betrachten? Wenige Idealisten, die die öffentliche Debatte nur deshalb beeinflussen können, weil sie die Klaviatur der neuen Medien beherrschen? Oder gibt es einen tatsächlichen gesellschaftlichen Diskurs über den Wert eines Netzes in dem jedes Datenpaket die gleichen Chancen hat – und einen Konsens dass es so sein sollte? Verwandt ist auch die Frage, welche Rolle in diesem Zusammenhang selbsterklärte Vertreter einer postulierten schweigenden Mehrheit spielen – und ob sie tatsächlich für eine Mehrheit der Nutzer sprechen.

Antworten könnten Soziologen und Kommunikationswissenschaftler liefern – sie bringen entsprechende Methoden in die Debatte ein und können vielleicht auch die Frage beantworten, welche Relevanz ein Messinstrument wie es in dieser Arbeit vorgeschlagen wird in der Praxis überhaupt hätte – und welcher Nutzerkreis sich dafür interessieren würde.

# A. Verbindungsabbruch bei Suche nach „Falun Gong“

Angefragt wird die chinesische Suchmaschine „Baidu“; diese wird zunächst mit einem (scheinbar) unproblematischen Begriff („germany“), dann mit einem politisch missliebigen Begriff („falun gong“) angefragt.

```
14:44:28.966833 IP wlan033084.uni-rostock.de.11479 > 220.181.111.147.http:
 Flags [P.], ack 1, win 65535, length 236
E....3@.@.....!T..o.,..P.J;..BO.P.......GET /s?wd=germany HTTP/1.1
Host: www.baidu.com
User-Agent: ELinks/0.11.7 (textmode; FreeBSD 8.0-BETA2 i386; 197x67-2)
Referer: http://www.baidu.com/
Accept: */*
Accept-Encoding: gzip
Accept-Language: en
Connection: Keep-Alive
```

```
14:44:29.273641 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [.], ack 237, win 6432, length 0
E..(.|@.,.....o...!T.P,..BO..J<.P.. Sb..
14:44:29.309478 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [P.], ack 237, win 6432, length 369
E....~@.,..%..o...!T.P,..BO..J<.P.. .x..HTTP/1.1 200 OK
Date: Mon, 04 Apr 2011 12:44:28 GMT
Server: BWS/1.0
Content-Length: 10697
Content-Type: text/html;charset=gbk
Cache-Control: private
Content-Encoding: gzip
Set-Cookie: BAIDUID=704188E60C2467232B6E99245E32E26C:FG=1; expires=Mon,
 04-Apr-41 12:44:28 GMT; path=/; domain=.baidu.com
P3P: CP=" OTI DSP COR IVA OUR IND COM "
Connection: Keep-Alive
```

Es folgt also auf die Anfrage der Beginn der Antwort. Nicht so beim anderen Stichwort:

```
14:45:11.197146 IP wlan033084.uni-rostock.de.11479 > 220.181.111.147.http:
Flags [P.], ack 2789374741, win 65535, length 241
E.....@.@.....!T..o.,..P.J<..B{.P...sZ..GET /s?wd=falun%20gong HTTP/1.1
Host: www.baidu.com
User-Agent: ELinks/0.11.7 (textmode; FreeBSD 8.0-BETA2 i386; 197x67-2)
Referer: http://www.baidu.com/
Accept: */*
Accept-Encoding: gzip
Accept-Language: en
Connection: Keep-Alive


14:45:11.502060 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [.], ack 241, win 7504, length 0
E..(..@.,..z..o...!T.P,..B{..J<.P..P#...
14:45:11.502466 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [R.], seq 1, ack 241, win 2390, length 0
E..(..@.|.yh..o...!T.P,..B{..J<.P.      V6...
14:45:11.503090 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [R.], seq 1461, ack 241, win 2391, length 0
E..(..@.}.xk..o...!T.P,..B...J<.P.      W1H..
14:45:11.503432 IP 220.181.111.147.http > wlan033084.uni-rostock.de.11479:
Flags [R.], seq 4381, ack 241, win 2392, length 0
E..(..@.~.u...o...!T.P,..B.1.J<.P.      X%...
```

Anstelle einer Antwort wird die TCP-Verbindung durch Pakete mit gesetztem `RST`-Flag zurückgesetzt. Dem Nutzer wird durch den Browser in diesem Fall die Meldung ausgegeben, dass die Verbindung zurückgesetzt wurde.

Der Test wurde am 04.04.2011 durchgeführt.

# B. Ergänzung von DNS-Einträgen

Im durchgeführten Test wurde eine DNS-Abfrage zunächst an den von OpenDNS betriebenen DNS-Server unter der IP-Adresse 208.67.222.222 und anschließend an den von Google betriebenen DNS-Server unter der IP-Adresse 8.8.8.8 gestellt. Hierbei ergaben sich unterschiedliche Antworten: Die erste Antwort enthielt einen weiteren Eintrag, der auf einen von OpenDNS angebotenen Dienst verwies.

```
[ad001@glas /usr/home/ad001]$ host inter7.jp 208.67.222.222
Using domain server:
Name: 208.67.222.222
Address: 208.67.222.222#53
Aliases:

inter7.jp has address 67.215.65.132
inter7.jp mail is handled by 15 hana.inter7.jp.
inter7.jp mail is handled by 5 aya.inter7.jp.
[ad001@glas /usr/home/ad001]$ host 67.215.65.132
132.65.215.67.in-addr.arpa domain name pointer hit-nxdomain.opendns.com.
[ad001@glas /usr/home/ad001]$ host inter7.jp 8.8.8.8
Using domain server:
Name: 8.8.8.8
Address: 8.8.8.8#53
Aliases:

inter7.jp mail is handled by 15 hana.inter7.jp.
inter7.jp mail is handled by 5 aya.inter7.jp.
```

Der Test wurde am 17.03.2013 durchgeführt.

# C. Ermittlung von Schätzwerten für die Netzqualitäten Durchsatz und Latenz mit einfachen Mitteln

Die in Kap. 5 beschriebene Testmethode soll an dieser Stelle um einen Vorschlag zur Abschätzung zu erwartender Netzqualitäten ergänzt werden. Die hier vorgeschlagene Methode ist mit einer relativ hohen Ungenauigkeit behaftet – dies ist allerdings nicht störend angesichts des Einsatzzwecks. Denn sie soll während einer „Live-Messung“ Verwendung finden, also dann, wenn die gerade vom Nutzer ausgelösten Datentransfers analysiert werden. Da die Aussagekraft einer solchen Messung auf Grund der (erwartbar) geringen Testdaten[1] ohnehin nur ein Indiz für eine Abweichung und Auslöser einer tiefgehenden Untersuchung (z. B. mit einer Replay-Messung) sein kann, muss auch eine hierzu durchgeführte Schätzung erwarteter Qualitätsparameter keine besonders hohe Genauigkeit aufweisen. Eine Abweichung von bis zu 10% scheint daher tolerierbar.

Im ersten Unterabschnitt wird eine minimalistische Überschlagsberechnung für die nach Passage mehrerer Netzwerke zu erwartenden Werte von Latenz und Durchsatz beschrieben. Die vorgeschlagenen Berechnungsvorschriften wurden im Rahmen eines Experiments empirisch untersucht; das Experiment und dessen Ergebnisse werden im zweiten Unterabschnitt beschrieben.

## Zusammenwirken von Netzwerkeffekten

In diesem Abschnitt werden für Latenz und Durchsatz zwei aus der Anschauung abgeleitete Vorschläge zur überschlagsweisen Berechnung eingeführt.

### Definitionen

Im folgenden sei „das Netzwerk“ das gesamte Internet; ein „Teilnetzwerk“ bezeichne ein AS. Aktive Technik (Router) befindet sich ausschließlich innerhalb von AS; Einflüsse der Verbindung zwischen AS werden implizit jeweils einem AS zugeschlagen. Die AS werden als Black Boxes behandelt, von denen nur das Verhalten nach außen hin beobachtet werden kann.

[1] nicht jede HTTP-Verbindung ist ein großer Download, der ein paar tausend Datenpakete als Messobjekt zur Verfügung stellt; viele Transfers bewegen sich im Bereich einiger Kilobytes

### Latenz

Anschaulich handelt es sich bei der Latenz um einen Zeitbetrag, um den sich die Transportzeit für ein Datenpaket bei der Weiterleitung an jedem Netzwerkknoten und jeder Netzwerkkante[2] verlängert. Da jedes Datenpaket stets individuell behandelt wird, hängt die konkrete Latenz eines Datenpaketes nicht nur vom Router sondern auch vom Zeitpunkt des Eintreffens ab. Aus dieser Nichtkonstantheit bei der Weiterleitung von Datenpaketen ergibt sich einen Jitter.

Als Überschlagsberechnung für ein Datenpaket, dass die Teilnetzwerke $N_0, ..., N_n$ passiert, wird hier vorgeschlagen, dass

$$\mathrm{lat}^*(N_0, N_n) \approx \sum_{i=0}^{n} \mathrm{lat}(N_i) + \mathrm{k}_{\mathrm{lat}}(n, \text{Messverfahren})$$

wobei k ein von der Anzahl passierte Netzwerke und dem verwendeten Messverfahren abhängiger Korrekturfaktor ist, um ggf. auftretende systematische Fehler zu dämpfen. Die Bestimmung von k wird in einem folgenden Abschnitt im Detail erörtert. Die Unschärfe der Formel ist angesichts des Einsatzzwecks bewusst gewählt.

### Durchsatz

Der Durchsatz wird anschaulich durch den Durchsatz des Teilnetzwerks mit dem geringsten Durchsatz begrenzt, im Falle des Passierens der Netzwerke $N_0, .., N_n$ ergäbe sich also der Gesamtdurchsatz zu

$$\mathrm{tp}^*(N_0, N_n) \approx \min_{i=0}^{n} \mathrm{tp}(N_i).$$

Auf die Einführung eines von der Messung abhängigen Korrekturfaktors soll hier verzichtet werden, da sich bei der Berechnung des Minimums die Fehler einzelner Messungen nicht zu einem größeren Gesamtfehler aufsummieren.

## Empirische Untersuchung

Da eine Verwendung realer WAN-Verbindungen zur Überprüfung der im vergangenen Abschnitt vorgestellten Methodik nicht in Frage kam (reale Zugangsnetze würden eine Kooperation mit einem Netzbetreiber bedeuten, bei der dieser fürchten muss, dass ihn belastende Geschäftsinterna Forschungsergebnis werden; die Nutzung von Overlay-WAN wie „PlanetLab“ ist ebenfalls kein gangbarer Ersatz da in diesem Fall keine Kontrolle über die unterliegenden realen Netze bestünde und eine Messung unwissentlich Artefakte unterliegender Netze eintrüge), wurde eine Abbildung eines typischen Internetszenarios auf ein umsetzbares Labormodell vorgenommen. Diese Umsetzung wird im ersten Teil dieses Abschnitts beschrieben. Darauffolgend werden die verwendeten Messmethoden für

[2] z. B. Satelliten-Uplink oder Ethernet

Durchsatz bzw. Latenz inklusive der empirischen Ermittlung des Korrekturfaktors k für die Latenzmessung beschrieben. Schließlich werden die Ergebnisse der Durchführung des Experiments zur Validierung der angenommenen Zusammenhänge verwendet.

## Labormodell

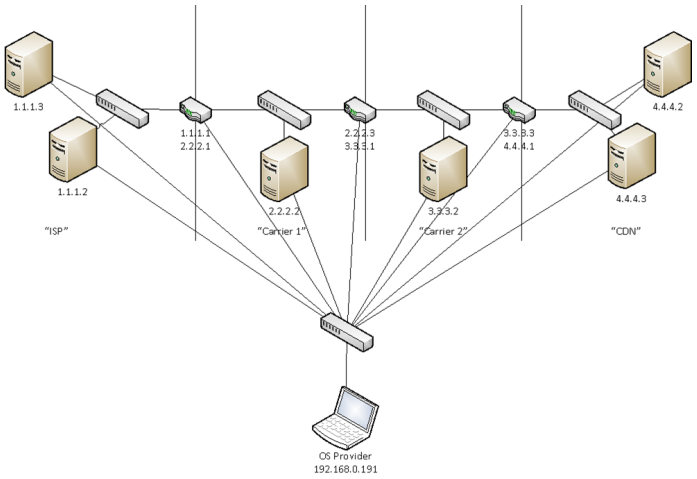


Abbildung C.1.: Topologie

Eine kurze Stichprobe mit einigen populären Nachrichten- und Medienportalen legt nahe, das eine Verbindung zwischen Endnutzer und Inhalteanbieter typischerweise zwei weitere Teilnetzwerke (wie in 5.3 eingeführt) passiert.

Hieran orientiert sich das Labormodell, dass ein ISP-, zwei Carrier- und ein CD-Netzwerk nachbildet. Für die Durchführung dieses Experiments wurde der Netzwerkpraktikumspool 301 im Institut für Informatik verwendet, so dass für sämtliche Router eine identische Hardware zur Verfügung stand. Die auf den Routern verwendete Software wurde zusammen mit dem Betriebssystem per PXE von einem zentralen Rechner zur Verfügung gestellt. Um eine Beeinflussung der Messung durch die per NFS zur Verfügung gestellten Betriebssysteme auszuschließen, wurde dieses Netzwerk mit abgetrennter Hardware (dedizierter Gigabit-Switch, dedizierte Netzwerkhardware) realisiert. Jeder IP-Adresse wurde mit einer eigenständigen Netzwerkkarte verbunden. Die zur Messung verwendete Hardware ist hiervon unabhängig.
Die Konfiguration der Teilnetze gibt Tabelle C.1 wieder; Abb. C.1 zeigt die verwendete Topologie.

Im ISP-Netz übernahmen die Maschinen 1.1.1.1 und 1.1.1.2 die Rolle von Endnutzer-Clients; im CDN-Netz übernahmen die Maschinen 4.4.4.2 und 4.4.4.3 die Rolle von Webservern. Jedes Netz erhielt einen eigenen Hardware-Switch, um gegenseitige Beeinflussungen zu minimieren; die zur Messung verwendeten Knoten wurden mit 100MBit/s

| Netz | Maschinen-IPs | Netmask |
|---|---|---|
| „ISP“ | 1.1.1.{1,2,3} | 255.255.255.0 |
| „Carrier 1“ | 2.2.2.{1,2,3} | 255.255.255.0 |
| „Carrier 2“ | 3.3.3.{1,2,3} | 255.255.255.0 |
| „CDN“ | 4.4.4.{1,2,3} | 255.255.255.0 |

Tabelle C.1.: Konfiguration der Teilnetze

verbunden; die Betriebssystemversorgung wurde mit 1GBit/s durchgeführt.

Bei den verwendeten Rechnern handelte es sich um Maschinen mit einem Intel Core 2 Duo Prozessor betrieben bei 2.9 GHz und 4 GB RAM; bei den Switches handelte es sich um je zwei Summit 200-24 und zwei Summit X250e-24t des Herstellers Extreme Networks, die entsprechend konfiguriert wurden.

## Messmethoden

In diesem Unterabschnitt werden die zur Ermittlung von Latenz und Durchsatz verwendeten Methoden erläutert. Zunächst wird die Latenzmessung mit dem Hilfsprogramm `ping` motiviert, anschließend die gewählte Form der Durchsatzmessung.

### Latenz

Da eine präzise Ermittlung von Latenzwerten anhand eines Datentransfers hohe Anforderungen an den Gleichlauf von Uhren auf den betrachteten Systemen stellt und dieser im Rahmen dieser Arbeit nicht sichergestellt werden konnte[3], wurde die Messung der Latenz mithilfe des bewährten Hilfsprogramms `ping` durchgeführt. `ping` versendet an sein Ziel ein ICMP-Echo-Request-Paket, welches von diesem mit einem (vom Kernel generierten) ICMP-Echo-Paket beantwortet wird. Durch die Wahl dieses Werkzeuges verändert sich auch der Charakter der Messung: Anstelle der unidirektionalen Latenz wird die Roundtripzeit (RTT) gemessen. Die RTT setzt sich hierbei bei einer Anfrage von `a` an `b` aus den folgenden Einzelkomponenten zusammen:

- Kernel von `a`: ICMP-Paket aussenden
- Transportzeit im Netzwerk
- Kernel passierter Router: Weiterleitung des Datenpaketes.
- Kernel von `b`: Generierung eines Antwortpaketes und dessen Aussendung
- Kernel von `a`: Antwortpaket an `ping`-Prozess zustellen

[3] Quartzabweichungen, schwer nachvollziehbarer nichtlinearer Temperaturdrift durch Mikroklima innerhalb des einzelnen Rechnergehäuses

Die Verwendbarkeit der RTT anstelle der realen Latenz bei der Übertragung eines Datenstroms ergibt sich aus den Vorgaben der Laborsituation: In einer realen Umgebung ist nicht absehbar, ob ICMP-Pakete differenziert von anderen Datenpaketen behandelt (privilegiert oder diskriminiert) werden. Da die verwendeten Systeme nicht derart konfiguriert wurden, beschränken sich unterschiedliche Behandlungen unterschiedlicher Pakete auf die vom Betriebssystem vorgegebenen (vernachlässigbar geringen) Effekte. Entsprechend reduziert sich der Unterschied zwischen RTT und Latenz bei der Weiterleitung eines Datenstroms auf die Bidirektionalität und ein ggf. anderes Verhalten der Router. Die Bidirektionalität ist unproblematisch, ja sogar hilfreich, da ausschließlich unterschiedliches Verhalten zwischen Teilnetzen von Interesse ist und das doppelte Passieren ebendiese mit einem Faktor 2 betont. Unterschiede im Verhalten der Router bestehen weiterhin zwischen dem Weiterleiten von ICMP-Echo (und -Reply)-Paketen und anderen gerouteten IP-Paketen. Entsprechende Pakete werden von unterschiedlichen Kernelteilen weitergeleitet; für ICMP-Pakete ist die Tiefe des Callstacks leicht geringer, dies kann zu einer minimalen Bevorzugung von ICMP-Paketen führen[4]. Dieser Effekt ist im vorgestellten Messdesign jedoch unerheblich, da ausschließlich ICMP-Pakete verwendet und miteinander verglichen werden. Für eine Messung wurden jeweils 3.000 ICMP-Pakete in einem zeitlichen Abstand von 0.01 Sekunden versandt. Die Paketzahl wurde als Kompromiss zwischen Dauer der Testdurchführung (bei einer kompletten Abdeckung ergeben sich 36 Tests; ein einzelner Test dauert bei den gegebenen Parametern 300 Sekunden; somit sind für die Durchführung der Latenzmessung ca. 3 Stunden anzunehmen) und der Umsetzbarkeit des Tests innerhalb eines gegebenen zeitlichen Rahmens (das Experiment sollte an einem Wochenende durchführbar sein) gewählt. In jedem Testdurchlauf wurden die Latenzen von jeder zu jeder Maschine ohne Rücksicht auf eventuell bestehende Symmetrieeffekte ermittelt.

### Durchsatz

Zur Bestimmung des Durchsatzes wurde ein Datenblock definierter Größe (s.u.) übertragen und die hierzu benötigte Zeit bestimmt. Die Verbindung wurde unidirektional ausgeführt: Es gab stets eine Daten sendende und eine Daten empfangende Station. Zunächst wird die Daten sendende, anschließend die Daten empfangende Station beschrieben.

Um den Einfluss weiterer Betriebsmittel auf die Übertragung (und damit auf die Messung) auszuschließen, wurden die zu übertragenden Daten vom Kernel generiert – konkret wurde mit Hilfe des Programms `dd` ein Datenblock von 75 MB aus dem virtuellen Device `/dev/zero` gelesen. Hier stellt der Betriebssystemkern einen konstanten Strom von Nullbytes zur Verfügung – nur von der CPU-Geschwindigkeit begrenzt[5]. Dieser Datenblock wurde durch das Hilfsprogramm `nc` per TCP zum messenden Rechner übertragen. Nach Abschluss einer Übertragung wurde die Nächste vorbereitet.

Auf der Seite des messenden Clients wurde eine TCP-Verbindung zum jeweiligen Mess-Peer aufgebaut, der Datenblock entgegengenommen (und in das virtuelle Device

[4] FreeBSD

[5] Auf einem aktuellen mit FreeBSD 9 betriebenen System ca. 6GByte/s

`/dev/null` geschrieben) und auf das Verbindungsende gewartet. Gemessen wurde die Laufzeit des Clients. Somit umfasst die gemessene Zeitspanne die folgenden Prozesse:

- Start des Prozesses `nc`
- Aufbau einer TCP-Verbindung mit dem jeweiligen Mess-Peer
- Übertragung des Datenblocks
- Abbau der TCP-Verbindung
- Eventuelle Aufräumprozesse von `nc` inklusive Prozessende

Die verstrichene Zeit wurde mit dem Hilfsprogramm `time` erfasst, dessen Ungenauigkeit durch eine entsprechend große Übertragung und Übertragungsdauer ausgeglichen wurde, so dass die relative Ungenauigkeit sank. Auf die weitere Auswirkungen wird bei den Ergebnissen näher eingegangen. Als Messergebnis ergaben sich die Übertragungszeiten für einen Datenblock von konstanter Größe und Herkunft zwischen den verschieden Rechnern des Versuchsaufbaus. Der Einfluss von neben der eigentlichen Datenübertragung vorhandenen weiteren Faktoren ist im Vergleich zu dieser vernachlässigbar, da es im vermessenen Netz keine konkurrierenden Datenübertragungen gab. Von Interesse sind grundsätzlich nur Differenzen, die sich zwischen unterschiedlichen Verbindungen ergeben. Programmstart und -finalisation sind jedoch auf Grund der Homogenität von Plattformen und Betriebssystem über alle eingesetzten Maschinen hinweg annähernd konstant. Mögliche Einflüsse durch Optimierungsmechanismen des TCP wurden in diesem Test nicht weiter separat betrachtet, da alle Messungen den gleichen Bedingungen unterliegen und somit auch Optimierungsverfahren zu vergleichbaren Ergebnissen führen werden. Die Größe des übertragenen Datenblocks motiviert sich ebenhieraus: Sie wurde so gewählt, dass der Einfluss der Datenübertragung gegenüber eventuell verbleibenden Einflüsse im Ergebnis deutlich dominanter ausfällt. Der Datenblock wies eine Größe von 75 MB auf; hieraus resultierten Übertragungszeiten im Bereich von 6 Sekunden. Zwischen zwei Übertragungen wird eine Pause von mindestens 1 Sekunde gewartet, in der die sendende Seite einen neuen Datenblock zur Verfügung stellen kann. In jedem Testlauf wird die Geschwindigkeit zwischen je zwei Maschinen in zehn einzelnen Durchläufen bestimmt; es werden Messungen zwischen allen Maschinen ohne Berücksichtigung eventueller Symmetrien durchgeführt.

### Bestimmung des Korrekturfaktors für die Latenzmessung

In der vorgeschlagenen Überschlagsberechnung der Latenz ist der Korrekturfaktor k($n$, Messverfahren) enthalten, der hier für das oben vorgestellte Messverfahren beschrieben werden soll. Dieser Korrekturfaktor gleicht systematische Fehler der Messung aus. Die folgende Abbildung illustriert, dass jede einzelne Latenzmessung zwangsläufig einen Messfehler einbringt. In einer einzelnen Messung kann dieser Effekt vernachlässigt werden; bei mehreren Messungen hingegen summieren sich diese Fehler und können das Ergebnis systematisch beeinflussen. Ein geeigneter Korrekturfaktor dämpft diesen Einfluss; im Idealfall unterdrückt er ihn komplett.

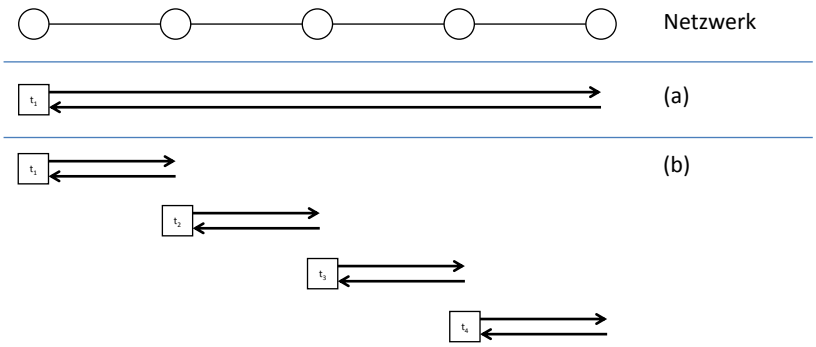


Abbildung C.2.: Akkumulation systematischer Fehler bei der Latenzmessung: (a) Einfacher Overhead bei der Messung der RTT zwischen den Endknoten, (b) Akkumulation der Overheads $t_{1,..,4}$.

Der systematische Messfehler der Latenz ist hier die Zeitspanne, die vom Programm `ping` zwar zur RTT gezählt wird, obwohl es sich nicht um die Zeit handelt, die ausschließlich für die Weiterleitung der Pakete verbraucht wird und der bei jedem Messvorgag anfällt. Der Fehler beinhaltet somit sowohl eine (von der Messstation aus betrachtet) lokale[6] wie auch eine Remote[7]-Komponente.

k kann durch die Ermittlung der Differenz aus der Laufzeit des Programms `ping` und der Summe der RTTs abgeleitet werden. Für eine hinreichend hohe Anzahl von Paketen ergibt sich die für Paketgenerierung und -Empfang benötigte Zeit. Hierbei bleibt die für Erzeugen und Versenden des Antwortpaketes benötigte Zeit unberücksichtigt. Eine alternative Methode wäre die Verwendung von aus den im Experiment gewonnenen Daten und einer entsprechenden Auswertung sowie einer anschließenden Nutzung zum Vergleich von Messungen zwischen Netzwerkknoten mit unterschiedlicher Entfernung im Routinggraphen. Bei Verfolgen dieses Ansatzes würde jedoch letztlich dasselbe Experiment zur Bestimmung einer freien Variablen wie auch zum Test des hierdurch gewonnenen Verfahrens dienen – schwerlich seriös.

Im Rahmen der folgenden Auswertung wird k(3000, Latenzmessung mit Ping im Laborversuch ) = 0 angenommen, also davon ausgegangen dass die systematischen Fehler vernachlässigbar klein sind. Ein solches Vorgehen ist hier vertretbar, da die – als Fehler in Frage kommende – Verarbeitungszeit hinreichend klein ist, da die Netzwerkknoten weder anderweitig ausgelastet werden noch durch weitere Datenübertragungen oder schwergewichtige Prozesse belastet sind.

[6] Erzeugen des ICMP-Echo-Request und Empfangen des Reply
[7] generieren des Echo-Paketes durch die angefragte Maschine

| Hosts (von, nach) | | $\overline{x}\pm$ sd [ms] |
|---|---|---|
| | Teilnetze | |
| 1.1.1.3 | 1.1.1.1 | $0.087 \pm 0.0839$ |
| 2.2.2.1 | 2.2.2.3 | $0.087 \pm 0.0263$ |
| 3.3.3.1 | 3.3.3.3 | $0.087 \pm 0.0622$ |
| 4.4.4.1 | 4.4.4.3 | $0.086 \pm 2.1057$ |
| | Testpfad | |
| 1.1.1.3 | 4.4.4.3 | $0.337 \pm 0.01825$ |

Tabelle C.2.: Messergebnisse, n=3000. Die Topologie ist in Abb. C.1 angegeben.

## Ergebnisse

In diesem Unterabschnitt sollen die vorgeschlagenen Überschlagsberechnungen einem Hypothesentest unterzogen werden. Hierzu wird als Beispiel eine Verbindung zwischen den Endknoten 1.1.1.3 und 4.4.4.1 dienen.

### Latenz

Der gemessene Wert der Latenz (Mittelwert) wird als Vergleichswert gegenüber der Überschlagsschätzung verwendet. In der Überschlagsschätzung werden die Daten der jeweiligen Teilnetze verwendet (diese würden von den Netzbetreibern veröffentlicht).

Im Experiment ergeben sich die in Tabelle C angegebenen Daten.

Getestet wird zwischen den Hosts 1.1.1.1 und 4.4.4.3. Verglichen werden jeweils die gemessene Latenz und der nach C ermittelte Latenzschätzer. Bezüglich der Latenz wird eine Normalverteilung angenommen; dies ergibt sich argumentativ aus der Ursache des Jitters im Labormodell: Jitter entsteht aus unterschiedlichen Reaktionszeiten einzelner Router beim Bearbeiten einzelner Datenpakete.

Als erwartete Latenz ergibt sich

$$\text{lat}^* \approx 0.347$$

Da es sich um einen überschlagsmäßigen Schätzer handelt, soll eine Irrtumswahrscheinlichkeit von $\alpha = 0.1$ erlaubt sein. Es ergeben sich die kritischen Werte von $-1,64$ und $1,64$. Die Testgröße ergib sich zu

$$T = \frac{\overline{X} - \mu}{\sigma}\sqrt{n} = 0.124$$

Im Rahmen der hier geforderten Sicherheit kann also von einer Übereinstimmung von Schätzer und tatsächlichem Wert ausgegangen werden.

| Hosts (von, nach) | | | $\overline{x}\pm$ sd [s] |
|---|---|---|---|
| | | Teilnetze | |
| 1.1.1.3 | 1.1.1.1 | | $6.689 \pm 0.00316$ |
| 2.2.2.1 | 2.2.2.3 | | $6.69 \pm 0$ |
| 3.3.3.1 | 3.3.3.3 | | $6.69 \pm 0$ |
| 4.4.4.1 | 4.4.4.3 | | $6.69 \pm 0$ |
| | | Testpfad | |
| 1.1.1.3 | 4.4.4.3 | tp= | $6.741 \pm 0.10236$ |

Tabelle C.3.: Messergebnisse, n=10. Die Topologie ist in Abb. C.1 angegeben.

### Durchsatz

Der gemessene Wert des Durchsatzes als wird als diejenige Zeit, die zum Transport eines bestimmten Datenvolumens gemessen wird festgehalten. Der Mittelwert der Messung zwischen den Endknoten wird als Vergleichswert für die Überschlagsschätzung verwendet. In der Überschlagsschätzung werden die Daten der jeweiligen Teilnetze verwendet (diese würden von den Netzbetreibern veröffentlicht).

Im Experiment ergeben sich die in Tabelle C angegebenen Daten.

Auch bezüglich des Durchsatzes wird eine Normalverteilung angenommen.

Als erwarteter Schätzer ergibt sich (hier als Maximalwert; Durchsatz entspräche einem normalisierten Kehrwert).

$$\mathrm{tp}^* \approx 6.689.$$

Da es sich um einen überschlagsmäßigen Schätzer handelt, soll eine Irrtumswahrscheinlichkeit von $\alpha = 0.1$ erlaubt sein. Es ergeben sich die kritischen Werte von $-1,64$ und $1,64$. Die Testgröße ergib sich zu

$$T = \frac{\overline{X} - \mu}{\sigma}\sqrt{n} = 0.774$$

Im Rahmen der hier geforderten Sicherheit kann also von einer Übereinstimmung von Schätzer und Erwartungswert für einen Schätzer ausgegangen werden.

## Zusammenfassung und Bewertung

Im Rahmen dieses Abschnitts konnte aufgezeigt werden, dass hinreichend genaue Abschätzungen wie für eine Prüfung der von Providern angegebenen Leistungsdaten mit den hier vorgestellten simplen Berechnungsvorschriften durchführbar sind. Die vorgestellte Methode kommt mit einfachen Mitteln zu Ergebnissen, die für einen Schätzer wie

in Kap. 5 im Rahmen einer Messung neben der tatsächlichen Verwendung benötigt geeignet erscheinen.

Die praktische Umsetzung wird ggf. nach anderen Mechanismen verlangen: Die Verwendung von ICMP wird in der Praxis nicht zur Prüfung der eigentlichen vom Provider deklarierten Policies geeignet sein, da es ein leichtes wäre, diese speziellen Datenpakete zu privilegieren und so Ergebnisse zu erzielen, die nicht für Nutzdaten gelten. Stattdessen müsste auch die Latenz an tatsächlichen Datenströmen gemessen werden.

# D. Test von Shaperprobe mit Dummynet

Um die Funktion von Shaperprobe zu testen, wurde der Arbeitsplatzrechner des Autors (Intel Core2Duo bei 3.17 GHz, 4 GB Ram, Microsoft Windows 7, Netzwerkanbindung durch das Institut für Informatik, sashimi.informatik.uni-rostock.de, für die Topologie vgl. Abb 1 auf S. 18) als Versuchsrechner eingesetzt. Die Topologie wurde für den Test um einen Shaper direkt zwischen dem Arbeitsplatzrechner und der nächsten Infrastrukturkomponente ergänzt, so dass sich aus Sicht des Programms Shaperprobe die in Abb. D.1 dargestellte Topologie in der Variante „Mit Shaper“ ergibt.

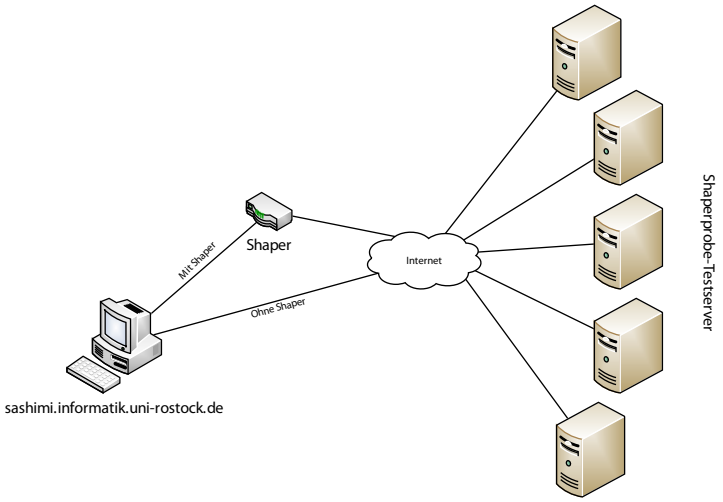


Abbildung D.1.: Infrastruktur des Experiments

Zunächst wurde ein Shaperprobe-Lauf durchgeführt, bei dem die Topologie in der Variante „Ohne Shaper“ aktiv war. Die Ausgabe von Shaperprobe ist in Abb. D.2 dargestellt: Es wurde kein Shaper gefunden.

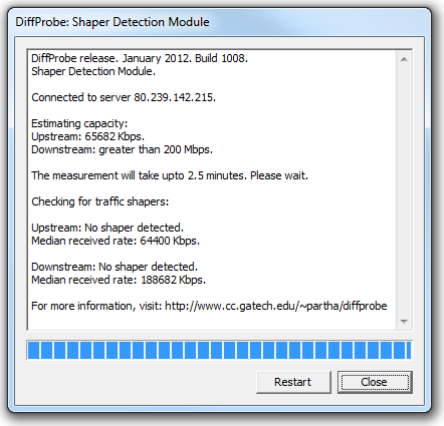


Abbildung D.2.: Ausgabe von Shaperprobe nach einem Lauf ohne Shaper

Im Anschluss wurde ein auf FreeBSD 9 und dem mittels `ipfw` angesprochenen Dummynet ein Layer-2-Shaper mit dem Regelsatz

```
add pipe 100 ip from any to any
pipe 100 config bw 1MBit/s
```

entsprechend der Variante „Mit Shaper“ in die Verbindung eingeschleift. Dieser Shaping-Regelsatz begrenzt sämtliche den Shaper passierenden Daten auf einen maximalen Durchsatz von 1 MBit/s. Die Verwendung dieser Kombination aus FreeBSD und Dummynet wurde gewählt, da sie sich bereits im praktischen Einsatz für sämtliche Studentenwohnheime der Rostocker Südstadt (ca. 1.000 Anschlüsse) bewährt hat [61].

Nach Installation des Shapers wurde ein erneuter Shaperprobe-Testlauf durchgeführt, dessen Ausgabe Abb. D.3 zeigt.

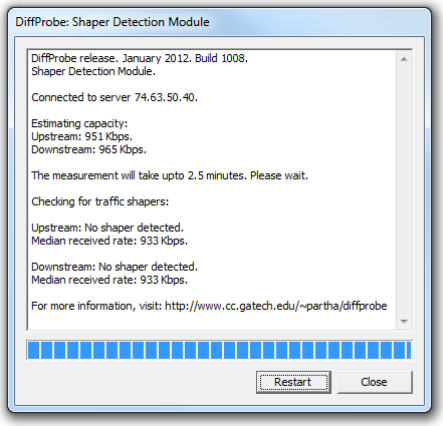


Abbildung D.3.: Ausgabe von Shaperprobe nach einem Lauf mit Shaper

Offensichtlich wurde der Shaper nicht detektiert.

Für den Test wurde ein trivialer Regelsatz verwendet; die Entwicklung eines Regelsatzes der zu Shaperprobe zugehörige Daten ungehindert passieren lässt, gestaltet sich einfach, da Testdaten von Shaperprobe leicht erkennbar sind und es nur einen begrenzten Satz an Testpeers gibt.

Durch dieses Experiment konnte gezeigt werden, dass die Beschränkung von Nachweisverfahren auf bestimmte Artefakte, die durch Shapingalgorithmen entstehen können, keine Sicherheit gegeben ist, Shaping in jedem Fall sicher zu erkennen.

# E. Überblick Regelungen zur Netzneutralität in verschiedenen Ländern

Dieser Anhang bietet eine kurze Übersicht über existierende Regelungen zur Netzneutralität. Entnommen aus [2].

**Chile** Festschreibung der Netzneutralität, Ausnahmen für „Sicherheit und Integrität des Netzes (Viren-, Jugend- und Datenschutz)“ [2]

**England** Keine Regelung, da von der OFCOM kein Regulierungsbedarf gesehen; Wettbewerb und Transparenz wurde als ausreichend angesehen. Die britischen Anbieter haben jedoch ein freiwilliges Abkommen unterzeichnet, in dem sie sich selbst zu einem „offenen und voll zugänglichen Netz“ verpflichten [2]

**Kanada** Transparenzforderung, Trafficengineering subsidiär ggü. Netzausbau, Blockaden nur nach Zustimmung der Regulierungsbehörde, Information und Rechtfertigung ggü. Kunden über alle Punkte des Netzwerkmanagements

**Niederlande** Festschreibung der Netzneutralität, Ausnahmen zur Sicherstellung der Funktionsfähigkeit; Regelung gilt nur für „Internet-Access“, nicht für „Managed Services“

**Norwegen** Absicherung der Netzneutralität in Form eines „Best-Effort-Basiskanals“, der eine ausreichende Kapazität aufweisen muss. Andere Angebote sind zulässig, hier gilt eine Transparenzpflicht

**Slowenien** Festschreibung der Netzneutralität, Ausnahmen bei akute Engpässen

**USA** Leitlinien der FCC von 2010[1], Offenlegung von Engineering u. Qualitäten, Verbot des Blockierens rechtmäßiger Inhalte.

[1] `http://www.fcc.gov/document/preserving-open-internet-broadband-industry-practices-1`

# F. Gesetzesauszüge

## BDSG

### § 3 Weitere Begriffsbestimmungen

(1) Personenbezogene Daten sind Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person (Betroffener).

(2) Automatisierte Verarbeitung ist die Erhebung, Verarbeitung oder Nutzung personenbezogener Daten unter Einsatz von Datenverarbeitungsanlagen. Eine nicht automatisierte Datei ist jede nicht automatisierte Sammlung personenbezogener Daten, die gleichartig aufgebaut ist und nach bestimmten Merkmalen zugänglich ist und ausgewertet werden kann.

(3) Erheben ist das Beschaffen von Daten über den Betroffenen.

(4) Verarbeiten ist das Speichern, Verändern, Übermitteln, Sperren und Löschen personenbezogener Daten. Im Einzelnen ist, ungeachtet der dabei angewendeten Verfahren:

1. Speichern das Erfassen, Aufnehmen oder Aufbewahren personenbezogener Daten auf einem Datenträger zum Zweck ihrer weiteren Verarbeitung oder Nutzung,
2. Verändern das inhaltliche Umgestalten gespeicherter personenbezogener Daten,
3. Übermitteln das Bekanntgeben gespeicherter oder durch Datenverarbeitung gewonnener personenbezogener Daten an einen Dritten in der Weise, dass
   a) die Daten an den Dritten weitergegeben werden oder
   b) der Dritte zur Einsicht oder zum Abruf bereitgehaltene Daten einsieht oder abruft,
4. Sperren das Kennzeichnen gespeicherter personenbezogener Daten, um ihre weitere Verarbeitung oder Nutzung einzuschränken,
5. Löschen das Unkenntlichmachen gespeicherter personenbezogener Daten.

(5) Nutzen ist jede Verwendung personenbezogener Daten, soweit es sich nicht um Verarbeitung handelt.

(6) Anonymisieren ist das Verändern personenbezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können.

(6a) Pseudonymisieren ist das Ersetzen des Namens und anderer Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren.

(7) Verantwortliche Stelle ist jede Person oder Stelle, die personenbezogene Daten für sich selbst erhebt, verarbeitet oder nutzt oder dies durch andere im Auftrag vornehmen lässt.

(8) Empfänger ist jede Person oder Stelle, die Daten erhält. Dritter ist jede Person oder Stelle außerhalb der verantwortlichen Stelle. Dritte sind nicht der Betroffene sowie Personen und Stellen, die im Inland, in einem anderen Mitgliedstaat der Europäischen Union oder in einem anderen Vertragsstaat des Abkommens über den Europäischen Wirtschaftsraum personenbezogene Daten im Auftrag erheben, verarbeiten oder nutzen.

(9) Besondere Arten personenbezogener Daten sind Angaben über die rassische und ethnische Herkunft, politische Meinungen, religiöse oder philosophische Überzeugungen, Gewerkschaftszugehörigkeit, Gesundheit oder Sexualleben.

(10) Mobile personenbezogene Speicher- und Verarbeitungsmedien sind Datenträger,

1. die an den Betroffenen ausgegeben werden,
2. auf denen personenbezogene Daten über die Speicherung hinaus durch die ausgebende oder eine andere Stelle automatisiert verarbeitet werden können und
3. bei denen der Betroffene diese Verarbeitung nur durch den Gebrauch des Mediums beeinflussen kann.

(11) Beschäftigte sind:

1. Arbeitnehmerinnen und Arbeitnehmer,
2. zu ihrer Berufsbildung Beschäftigte,
3. Teilnehmerinnen und Teilnehmer an Leistungen zur Teilhabe am Arbeitsleben sowie an Abklärungen der beruflichen Eignung oder Arbeitserprobung (Rehabilitandinnen und Rehabilitanden),
4. in anerkannten Werkstätten für behinderte Menschen Beschäftigte,
5. nach dem Jugendfreiwilligendienstegesetz Beschäftigte,
6. Personen, die wegen ihrer wirtschaftlichen Unselbständigkeit als arbeitnehmerähnliche Personen anzusehen sind; zu diesen gehören auch die in Heimarbeit Beschäftigten und die ihnen Gleichgestellten,
7. Bewerberinnen und Bewerber für ein Beschäftigungsverhältnis sowie Personen, deren Beschäftigungsverhältnis beendet ist,
8. Beamtinnen, Beamte, Richterinnen und Richter des Bundes, Soldatinnen und Soldaten sowie Zivildienstleistende.

# BGB

## § 312g Pflichten im elektronischen Geschäftsverkehr

(1) Bedient sich ein Unternehmer zum Zwecke des Abschlusses eines Vertrags über die Lieferung von Waren oder über die Erbringung von Dienstleistungen der Telemedien (Vertrag im elektronischen Geschäftsverkehr), hat er dem Kunden

1. angemessene, wirksame und zugängliche technische Mittel zur Verfügung zu stellen, mit deren Hilfe der Kunde Eingabefehler vor Abgabe seiner Bestellung erkennen und berichtigen kann,

2. die in Artikel 246 ğ 3 des Einführungsgesetzes zum Bürgerlichen Gesetzbuche bestimmten Informationen rechtzeitig vor Abgabe von dessen Bestellung klar und verständlich mitzuteilen,

3. den Zugang von dessen Bestellung unverzüglich auf elektronischem Wege zu bestätigen und

4. die Möglichkeit zu verschaffen, die Vertragsbestimmungen einschlieSSlich der Allgemeinen Geschäftsbedingungen bei Vertragsschluss abzurufen und in wiedergabefähiger Form zu speichern.

Bestellung und Empfangsbestätigung im Sinne von Satz 1 Nr. 3 gelten als zugegangen, wenn die Parteien, für die sie bestimmt sind, sie unter gewöhnlichen Umständen abrufen können.

(2) Bei einem Vertrag im elektronischen Geschäftsverkehr zwischen einem Unternehmer und einem Verbraucher, der eine entgeltliche Leistung des Unternehmers zum Gegenstand hat, muss der Unternehmer dem Verbraucher die Informationen gemäSS Artikel 246 ğ 1 Absatz 1 Nummer 4 erster Halbsatz und Nummer 5, 7 und 8 des Einführungsgesetzes zum Bürgerlichen Gesetzbuche, unmittelbar bevor der Verbraucher seine Bestellung abgibt, klar und verständlich in hervorgehobener Weise zur Verfügung stellen. Diese Pflicht gilt nicht für Verträge über die in ğ 312b Absatz 1 Satz 2 genannten Finanzdienstleistungen.

(3) Der Unternehmer hat die Bestellsituation bei einem Vertrag nach Absatz 2 Satz 1 so zu gestalten, dass der Verbraucher mit seiner Bestellung ausdrücklich bestätigt, dass er sich zu einer Zahlung verpflichtet. Erfolgt die Bestellung über eine Schaltfläche, ist die Pflicht des Unternehmers aus Satz 1 nur erfüllt, wenn diese Schaltfläche gut lesbar mit nichts anderem als den Wörtern ĎzahlungspflichtigŞ bestellen oder mit einer entsprechenden eindeutigen Formulierung beschriftet ist.

(4) Ein Vertrag nach Absatz 2 Satz 1 kommt nur zustande, wenn der Unternehmer seine Pflicht aus Absatz 3 erfüllt.

(5) Absatz 1 Satz 1 Nr. 1 bis 3 und die Absätze 2 bis 4 finden keine Anwendung, wenn der Vertrag ausschlieSSlich durch individuelle Kommunikation geschlossen wird. Absatz 1 Satz 1 Nr. 1 bis 3 und Satz 2 findet keine Anwendung, wenn zwischen Vertragsparteien, die nicht Verbraucher sind, etwas anderes vereinbart wird.

(6) Weitergehende Informationspflichten auf Grund anderer Vorschriften bleiben unberührt. Steht dem Kunden ein Widerrufsrecht gemäSS ğ 355 zu, beginnt die Widerrufsfrist

abweichend von § 355 Abs. 3 Satz 1 nicht vor Erfüllung der in Absatz 1 Satz 1 geregelten Pflichten.

# GWB

## § 19 Verbotenes Verhalten von marktbeherrschenden Unternehmen

(1) Die missbräuchliche Ausnutzung einer marktbeherrschenden Stellung durch ein oder mehrere Unternehmen ist verboten.

(2) Ein Missbrauch liegt insbesondere vor, wenn ein marktbeherrschendes Unternehmen als Anbieter oder Nachfrager einer bestimmten Art von Waren oder gewerblichen Leistungen

1. ein anderes Unternehmen unmittelbar oder mittelbar unbillig behindert oder ohne sachlich gerechtfertigten Grund unmittelbar oder mittelbar anders behandelt als gleichartige Unternehmen;

2. Entgelte oder sonstige Geschäftsbedingungen fordert, die von denjenigen abweichen, die sich bei wirksamem Wettbewerb mit hoher Wahrscheinlichkeit ergeben würden; hierbei sind insbesondere die Verhaltensweisen von Unternehmen auf vergleichbaren Märkten mit wirksamem Wettbewerb zu berücksichtigen;

3. ungünstigere Entgelte oder sonstige Geschäftsbedingungen fordert, als sie das marktbeherrschende Unternehmen selbst auf vergleichbaren Märkten von gleichartigen Abnehmern fordert, es sei denn, dass der Unterschied sachlich gerechtfertigt ist;

4. sich weigert, einem anderen Unternehmen gegen angemessenes Entgelt Zugang zu den eigenen Netzen oder anderen Infrastruktureinrichtungen zu gewähren, wenn es dem anderen Unternehmen aus rechtlichen oder tatsächlichen Gründen ohne die Mitbenutzung nicht möglich ist, auf dem vor- oder nachgelagerten Markt als Wettbewerber des marktbeherrschenden Unternehmens tätig zu werden; dies gilt nicht, wenn das marktbeherrschende Unternehmen nachweist, dass die Mitbenutzung aus betriebsbedingten oder sonstigen Gründen nicht möglich oder nicht zumutbar ist;

5. seine Marktstellung dazu ausnutzt, andere Unternehmen dazu aufzufordern oder zu veranlassen, ihm ohne sachlich gerechtfertigten Grund Vorteile zu gewähren.

(3) Absatz 1 in Verbindung mit Absatz 2 Nummer 1 und Nummer 5 gilt auch für Vereinigungen von miteinander im Wettbewerb stehenden Unternehmen im Sinne der §§ 2, 3 und 28 Absatz 1, § 30 Absatz 2a und ğ 31 Absatz 1 Nummer 1, 2 und 4. Absatz 1 in Verbindung mit Absatz 2 Nummer 1 gilt auch für Unternehmen, die Preise nach § 28 Absatz 2 oder § 30 Absatz 1 Satz 1 oder § 31 Absatz 1 Nummer 3 binden.

## § 20 Verbotenes Verhalten von Unternehmen mit relativer oder überlegener Marktmacht

(1) § 19 Absatz 1 in Verbindung mit Absatz 2 Nummer 1 gilt auch fÜr Unternehmen und Vereinigungen von Unternehmen, soweit von ihnen kleine oder mittlere Unternehmen als Anbieter oder Nachfrager einer bestimmten Art von Waren oder gewerblichen Leistungen in der Weise abhängig sind, dass ausreichende und zumutbare Möglichkeiten, auf andere Unternehmen auszuweichen, nicht bestehen (relative Marktmacht). Es wird vermutet, dass ein Anbieter einer bestimmten Art von Waren oder gewerblichen Leistungen von einem Nachfrager abhängig im Sinne des Satzes 1 ist, wenn dieser Nachfrager bei ihm zusätzlich zu den verkehrsüblichen Preisnachlässen oder sonstigen Leistungsentgelten regelmäßig besondere Vergünstigungen erlangt, die gleichartigen Nachfragern nicht gewährt werden.

(2) § 19 Absatz 1 in Verbindung mit Absatz 2 Nummer 5 gilt auch für Unternehmen und Vereinigungen von Unternehmen im Verhältnis zu den von ihnen abhängigen Unternehmen.

(3) Unternehmen mit gegenüber kleinen und mittleren Wettbewerbern überlegener Marktmacht dürfen ihre Marktmacht nicht dazu ausnutzen, solche Wettbewerber unmittelbar oder mittelbar unbillig zu behindern. Eine unbillige Behinderung im Sinne des Satzes 1 liegt insbesondere vor, wenn ein Unternehmen

1. Lebensmittel im Sinne des § 2 Absatz 2 des Lebensmittel- und Futtermittelgesetzbuches unter Einstandspreis oder

2. andere Waren oder gewerbliche Leistungen nicht nur gelegentlich unter Einstandspreis oder

3. von kleinen oder mittleren Unternehmen, mit denen es auf dem nachgelagerten Markt beim Vertrieb von Waren oder gewerblichen Leistungen im Wettbewerb steht, für deren Lieferung einen höheren Preis fordert, als es selbst auf diesem Markt anbietet,

es sei denn, dies ist jeweils sachlich gerechtfertigt. Das Anbieten von Lebensmitteln unter Einstandspreis ist sachlich gerechtfertigt, wenn es geeignet ist, den Verderb oder die drohende Unverkäuflichkeit der Waren beim Händler durch rechtzeitigen Verkauf zu verhindern sowie in vergleichbar schwerwiegenden Fällen. Werden Lebensmittel an gemeinnützige Einrichtungen zur Verwendung im Rahmen ihrer Aufgaben abgegeben, liegt keine unbillige Behinderung vor.[1]

(4) Ergibt sich auf Grund bestimmter Tatsachen nach allgemeiner Erfahrung der Anschein, dass ein Unternehmen seine Marktmacht im Sinne des Absatzes 3 ausgenutzt hat, so obliegt es diesem Unternehmen, den Anschein zu widerlegen und solche anspruchsbegründenden Umstände aus seinem Geschäftsbereich aufzuklären, deren Aufklärung dem betroffenen Wettbewerber oder einem Verband nach § 33 Absatz 2 nicht möglich, dem in Anspruch genommenen Unternehmen aber leicht möglich und zumutbar ist.

(5) Wirtschafts- und Berufsvereinigungen sowie Gütezeichengemeinschaften dürfen die Aufnahme eines Unternehmens nicht ablehnen, wenn die Ablehnung eine sachlich nicht

gerechtfertigte ungleiche Behandlung darstellen und zu einer unbilligen Benachteiligung des Unternehmens im Wettbewerb führen würde.

[1] § 20 Absatz 3 gilt gemäß Artikel 2 in Verbindung mit Artikel 7 Satz 2 des Gesetzes vom 26. Juni 2013 (BGBl. I S. 1738) ab 1. Januar 2018 in folgender Fassung:

„(3) Unternehmen mit gegenüber kleinen und mittleren Wettbewerbern überlegener Marktmacht dürfen ihre Marktmacht nicht dazu ausnutzen, solche Wettbewerber unmittelbar oder mittelbar unbillig zu behindern. Eine unbillige Behinderung im Sinne des Satzes 1 liegt insbesondere vor, wenn ein Unternehmen

1. Waren oder gewerbliche Leistungen nicht nur gelegentlich unter Einstandspreis anbietet oder

2. von kleinen oder mittleren Unternehmen, mit denen es auf dem nachgelagerten Markt beim Vertrieb von Waren oder gewerblichen Leistungen im Wettbewerb steht, für deren Lieferung einen höheren Preis fordert, als es selbst auf diesem Markt anbietet,

es sei denn, dies ist jeweils sachlich gerechtfertigt.

# StGB

## § 109g Sicherheitsgefährdendes Abbilden

(1) Wer von einem Wehrmittel, einer militärischen Einrichtung oder Anlage oder einem militärischen Vorgang eine Abbildung oder Beschreibung anfertigt oder eine solche Abbildung oder Beschreibung an einen anderen gelangen läßt und dadurch wissentlich die Sicherheit der Bundesrepublik Deutschland oder die Schlagkraft der Truppe gefährdet, wird mit Freiheitsstrafe bis zu fünf Jahren oder mit Geldstrafe bestraft.

(2) Wer von einem Luftfahrzeug aus eine Lichtbildaufnahme von einem Gebiet oder Gegenstand im räumlichen Geltungsbereich dieses Gesetzes anfertigt oder eine solche Aufnahme oder eine danach hergestellte Abbildung an einen anderen gelangen läßt und dadurch wissentlich die Sicherheit der Bundesrepublik Deutschland oder die Schlagkraft der Truppe gefährdet, wird mit Freiheitsstrafe bis zu zwei Jahren oder mit Geldstrafe bestraft, wenn die Tat nicht in Absatz 1 mit Strafe bedroht ist.

(3) Der Versuch ist strafbar.

(4) Wer in den Fällen des Absatzes 1 die Abbildung oder Beschreibung an einen anderen gelangen läßt und dadurch die Gefahr nicht wissentlich, aber vorsätzlich oder leichtfertig herbeiführt, wird mit Freiheitsstrafe bis zu zwei Jahren oder mit Geldstrafe bestraft. Die Tat ist jedoch nicht strafbar, wenn der Täter mit Erlaubnis der zuständigen Dienststelle gehandelt hat.

# TKG

## § 41a Netzneutralität

(1) Die Bundesregierung wird ermächtigt, in einer Rechtsverordnung mit Zustimmung des Bundestages und des Bundesrates gegenüber Unternehmen, die Telekommunikationsnetze betreiben, die grundsätzlichen Anforderungen an eine diskriminierungsfreie Datenübermittlung und den diskriminierungsfreien Zugang zu Inhalten und Anwendungen festzulegen, um eine willkürliche Verschlechterung von Diensten und eine ungerechtfertigte Behinderung oder Verlangsamung des Datenverkehrs in den Netzen zu verhindern; sie berücksichtigt hierbei die europäischen Vorgaben sowie die Ziele und Grundsätze des § 2.

(2) Die Bundesnetzagentur kann in einer Technischen Richtlinie Einzelheiten über die Mindestanforderungen an die Dienstqualität durch Verfügung festlegen. Bevor die Mindestanforderungen festgelegt werden, sind die Gründe für ein Tätigwerden, die geplanten Anforderungen und die vorgeschlagene Vorgehensweise zusammenfassend darzustellen; diese Darstellung ist der Kommission und dem GEREK rechtzeitig zu übermitteln. Den Kommentaren oder Empfehlungen der Kommission ist bei der Festlegung der Anforderungen weitestgehend Rechnung zu tragen.

## § 43a Verträge

(1) Anbieter von öffentlich zugänglichen Telekommunikationsdiensten müssen dem Verbraucher und auf Verlangen anderen Endnutzern im Vertrag in klarer, umfassender und leicht zugänglicher Form folgende Informationen zur Verfügung stellen:

1. den Namen und die ladungsfähige Anschrift; ist der Anbieter eine juristische Person auch die Rechtsform, den Sitz und das zuständige Registergericht,
2. die Art und die wichtigsten technischen Leistungsdaten der angebotenen Telekommunikationsdienste, insbesondere diejenigen gemäß Absatz 2 und Absatz 3 Satz 1,
3. die voraussichtliche Dauer bis zur Bereitstellung eines Anschlusses,
4. die angebotenen Wartungs- und Kundendienste sowie die Möglichkeiten zur Kontaktaufnahme mit diesen Diensten,
5. Einzelheiten zu den Preisen der angebotenen Telekommunikationsdienste,
6. die Fundstelle eines allgemein zugänglichen, vollständigen und gültigen Preisverzeichnisses des Anbieters von öffentlich zugänglichen Telekommunikationsdiensten,
7. die Vertragslaufzeit, einschließlich des Mindestumfangs und der Mindestdauer der Nutzung, die gegebenenfalls erforderlich sind, um Angebote im Rahmen von Werbemaßnahmen nutzen zu können,

8. die Voraussetzungen für die Verlängerung und Beendigung des Bezuges einzelner Dienste und des gesamten Vertragsverhältnisses, einschließlich der Voraussetzungen für einen Anbieterwechsel nach § 46, die Entgelte für die Übertragung von Nummern und anderen Teilnehmerkennungen sowie die bei Beendigung des Vertragsverhältnisses fälligen Entgelte einschließlich einer Kostenanlastung für Endeinrichtungen,

9. etwaige Entschädigungs- und Erstattungsregelungen für den Fall, dass der Anbieter die wichtigsten technischen Leistungsdaten der zu erbringenden Dienste nicht eingehalten hat,

10. die erforderlichen Schritte zur Einleitung eines außergerichtlichen Streitbeilegungsverfahrens nach § 47a,

11. den Anspruch des Teilnehmers auf Aufnahme seiner Daten in ein öffentliches Teilnehmerverzeichnis nach § 45m,

12. die Arten von Maßnahmen, mit denen das Unternehmen auf Sicherheits- oder Integritätsverletzungen oder auf Bedrohungen und Schwachstellen reagieren kann,

13. den Anspruch auf Sperrung bestimmter Rufnummernbereiche nach § 45d Absatz 2 Satz 1 und

14. den Anspruch auf Sperrung der Inanspruchnahme und Abrechnung von neben der Verbindung erbrachten Leistungen über den Mobilfunkanschluss nach § 45d Absatz 3.

Anbieter öffentlicher Telekommunikationsnetze sind dazu verpflichtet, Anbietern öffentlich zugänglicher Telekommunikationsdienste die für die Sicherstellung der in Satz 1 genannten Informationspflichten benötigten Informationen zur Verfügung zu stellen, wenn ausschließlich die Anbieter von öffentlichen Telekommunikationsnetzen darüber verfügen.

(2) Zu den Informationen nach Absatz 1 Nummer 2 gehören

1. Informationen darüber, ob der Zugang zu Notdiensten mit Angaben zum Anruferstandort besteht oder nicht, und über alle Beschränkungen von Notdiensten,

2. Informationen über alle Einschränkungen im Hinblick auf den Zugang zu und die Nutzung von Diensten und Anwendungen,

3. das angebotene Mindestniveau der Dienstqualität und gegebenenfalls anderer nach § 41a festgelegter Parameter für die Dienstqualität,

4. Informationen über alle vom Unternehmen zur Messung und Kontrolle des Datenverkehrs eingerichteten Verfahren, um eine Kapazitätsauslastung oder Überlastung einer Netzverbindung zu vermeiden, und Informationen über die möglichen Auswirkungen dieser Verfahren auf die Dienstqualität und

5. alle vom Anbieter auferlegten Beschränkungen für die Nutzung der von ihm zur Verfügung gestellten Endeinrichtungen.

(3) Die Einzelheiten darüber, welche Angaben in der Regel mindestens nach Absatz 2 erforderlich sind, kann die Bundesnetzagentur nach Beteiligung der betroffenen Verbände und der Unternehmen durch Verfügung im Amtsblatt festlegen. Hierzu kann die Bundesnetzagentur die Anbieter öffentlich zugänglicher Telekommunikationsdienste oder die Anbieter öffentlicher Telekommunikationsnetze verpflichten, Erhebungen zum tatsächlichen Mindestniveau der Dienstqualität anzustellen, eigene Messungen anstellen oder Hilfsmittel entwickeln, die es dem Teilnehmer ermöglichen, eigenständige Messungen anzustellen. Ferner kann die Bundesnetzagentur das Format der Mitteilung über Vertragsänderungen und die anzugebende Information über das Widerrufsrecht festlegen, soweit nicht bereits vergleichbare Regelungen bestehen.

# Abbildungsverzeichnis

# Wissenschaftliche Quellen

Viele in dieser Arbeit diskutierte Aspekte der Netzneutralität sind noch nicht in Form wissenschaftlicher Publikationen bearbeitet worden. Diese sollten aber dennoch zitiert werden, da die Arbeit ohne diese Quellen unvollständig wäre. Um dem Leser dennoch die Einschätzung der Quellen zu erleichtern, wurde die Bibliographie in zwei Teile geteilt: Die „wissenschaftlichen“ und die „weiteren Quellen“. Die Einordnung in die ein- oder andere Kategorie richtete sich nach Präsentation und Fundstelle einzelner Zitate.


[1] J. Crowcroft, “Net Neutrality: The Technical Side of the Debate: A White Paper.” in SIGCOMM Computer Communication Review, Volume 37, Issue 1, ACM, Seiten 49ff., New York, 2007.

[2] K. Mengering, “§ 41a Netzneutralität.” in F. J. Säcker, TKG (Kommentar), Seiten 1431ff., dtv, 2013.

[3] S. Schlauri, “Network Neutrality: Netzneutralität als neues Regulierungsprinzip des Telekommunikationsrechts.” Habilitationsschrift, erschienen bei Nomos/Dike Verlag (ISBN 978-3-0375-1261-6), Baden-Baden/Zürich, 2010.

[4] D. Baecker, “Form und Formen der Kommunikation.” erschienen bei Suhrkamp (ISBN 978-3518584392), Frankfurt a.M., 2007.

[5] E. Pariser, “The Filter Bubble: What the Internet is Hiding from You.” erschienen bei Viking (ISBN 978-0-6709-2038-9), London, 2011.

[6] E. Bozdag, “Bias in Algorithmic Filtering and Personalization.” in Ethics and Information Technology, Volume 15, Issue 3, Seiten 209ff., Springer, Dodrecht, 2013.

[7] B. Danckert and F. J. Mayer, “Die vorherrschende Meinungsmacht von Google.” in Multimedia und Recht MMR, Volume 14, Issue 4, Seiten 219ff., C.H. Beck, München, 2010.

[8] S. Ott, “Marktbeherrschende und öffentlich-rechtliche Suchmaschinen.” in Kommunikation und Recht K&R, Volume 10, Issue 7/8, Seiten 375ff., Deutscher Fachverlag GmbH, Frankfurt a.M., 2007.

[9] S. Meyer, “Aktuelle Rechtsentwicklung bei Suchmaschinen im Jahre 2012.” in Kommunikation und Recht K&R, Volume 15, Issue 4, Seiten 221ff., Deutscher Fachverlag GmbH, Frankfurt a.M., 2013.

[10] A. Dähn and C. H. Cap, "Application transparency: How and why are providers manipulating our information?." in IEEE Computer, Volume 47, Issue 2, Seiten 56ff., IEEE Computer Society, Los Alamitos, 2014.

[11] T. Wu, "Network neutrality, broadband discrimination." in Journal of Telecommunications and High Technology Law, Volume 2, Issue 1, Seiten 141ff., Boulder, 2003.

[12] G. M. Bullinger, "Netzneutralität: Pro und Contra einer gesetzlichen Festschreibung." in Wissenschaftliche Dienste des Deutschen Bundestages, Dokumentennummer WD 10 - 3000/065-10, Berlin, 2010.

[13] M. Dischinger, M. Marcon, S. Guha, K. P. Gummadi, R. Mahajan, and S. Saroiu, "Glasnost: Enabling End Users to Detect Traffic Differentiation." in Proceedings des 7. Symposium on Networked Systems Design and Implementation NSDI, Usenix, San Jose, 2010.

[14] M. B. Tariq, M. Motiwala, N. Feamster, and M. Ammar, "Detecting Network Neutrality Violations with Causal Inference." in Proceedings der 5. International Conference on Emerging Networking Experiments and Technologies CoNEXT, ACM, Seiten 289ff., New York, 2009.

[15] M. Dhawan, J. Samuel, R. Teixeira, C. Kreibich, M. Allman, N. Weaver, and V. Paxson, "Fathom: a browser-based network measurement platform." in Proceedings der Internet Measurement Conference IMC 2012, ACM, Seiten 73ff., Boston, 2012.

[16] "Herdict." Webressource `http://herdict.org`, betrieben vom Berkman Center for Internet and Society der Harvard University, zuletzt besucht am 2014-04-15.

[17] D. Kaminsky, "Black Ops Of TCP/IP 2011." Vortrag auf der 19. DefCon, Foliensatz auf der Seite des Autors unter `http://dankaminsky.com/2011/08/05/bo2k11` (zuletzt abgerufen am 2014-04-15), Las Vegas, 2011.

[18] R. Dewenter, "Effiziente Regeln für Telekommunikationsmärkte der Zukunft." erschienen als 7. Ausgabe der Reihe „Wettbewerb und Regulierung von Märkten und Unternehmen", Hrsg. von Haucap, Kühling, erschienen bei Nomos (ISBN 978-3-8329-3736-2), 2009.

[19] C. Schwarz-Schilling, "Netzneutralität aus der Perspektive der Bundesnetzagentur." in „Netzneutralität in der Informationsgesellschaft", Hrsg. von Kloepfer als 27. Teil der Reihe „Beiträge zum Informationsrecht", erschienen bei Duncker & Humblot (ISBN 978-3-428-13677-3), Seiten 133ff., Berlin, 2011.

[20] H. Gersdorf, "Netzneutralität: Landesrechtliche Plattformregulierung als Referenzmodell?." in „Digitalisierungsbericht 2010", Hrsg. Kommission für die Zulassung und Aufsicht der Landesmedienanstalten ZAK, Berlin, 2010.

[21] A. S. Tannenbaum, "Computernetzwerke." erschienen bei Pearson Studium (ISBN 3-8273-7046-9), 4. Auflage, München, 2003.

[22] R. Lindhorst, "Sicherheit von drahtlosen Netzwerken." Diplomarbeit, Universität Rostock, Rostock, 2007.

[23] A. Dähn, "Entwicklung eines empirischen Verfahrens zum Nachweis der An- oder Abwesenheit von Netzneutralität." Diplomarbeit, Universität Rostock, Rostock, 2011.

[24] C. Stoll, "Stalking the wily hacker." in Communications of the ACM, Volume 31, Issue 5, ACM, Seiten 484ff., New York, 1988.

[25] B. Quoitin, C. Pelsser, L. Swinnen, O. Bonaventure, and S. Uhlig, "Interdomain Traffic Engineering with BGP." in Communications Magazine, Volume 41, Issue 5, IEEE, Seiten 122ff., 2003.

[26] B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira, "Avoiding traceroute anomalies with paris traceroute." in Proceedings der 6. ACM SIGCOMM Conference on Internet Measurement IMC 06, Seiten 153ff., ACM, New York, 2006.

[27] A. Chaudhary and A. Sardana, "Software Based Implementation Methodologies for Deep Packet Inspection." in Proceedings der International Conference on Information Science and Applications ICISA 2011, 2011.

[28] R. K. Lenka and P. Ranjan, "A Comparative Study on DFA-Based Pattern Matching for Deep Packet Inspection." in Proceedings der 3. International Conference on Computer and Communication Technology ICCCT 2012, IEEE, Seiten 255ff., 2012.

[29] X. Kefu, Q. Deyu, Q. Zhengping, and Z. Weiping, "Fast dynamic pattern matching for deep packet inspection." in Proceedings der International Conference on Networking, Sensing and Control 2008, IEEE, Seiten 802ff., 2008.

[30] A. Ali and R. Tervo, "Traffic identification using Bayes' classifier." in Proceedings der 2. Canadian Conference on Electrical and Computer Engineering 2000, IEEE, Seiten 687ff., Halifax, 2000.

[31] C. Liu, G. Sun, and Y. Xue, "DRPSD: An novel method of identifying SSL/TLS traffic." in Proceedings des World Automation Congress WAC 2012, IEEE, Seiten 415ff., Puerto Vallarta, 2012.

[32] R. Archibald, Y. Liu, C. Corbett, and D. Ghosal, "Disambiguating HTTP: Classifying web Applications." in Proceedings der 7. International Wireless Communications and Mobile Computing Conference IWCMC 211, IEEE, Seiten 1808ff., Istanbul, 2011.

[33] H. Gersdorf, "§ 41a Netzneutralität." in Kommentar zum TKG aus der Reihe „Informations- und Medienrecht", BeckOK, 3. Edition. C.H. Beck, München, 2014.

[34] "Glasnost: Test if your ISP is shaping your traffic." Webresource `http://broadband.mpi-sws.org/transparency/bttest.php`, Betrieben vom Max Planck Institute for Software Systems, Saarbrücken, zuletzt besucht am 2014-04-15.

[35] Nick Feamster and Mostafa Ammar and Muhammad Mukarram bin Tariq and Murtaza Motiwala, "GTNOISE Network Access Neutrality Project." Webresource `http://gtnoise.net/nano/`, betrieben vom Georgia Institute of Technology, Georgia, zuletzt besucht am 2014-04-15.

[36] P. Kanuparthy, "Shaperprobe." Webresource `http://www.cc.gatech.edu/~partha/diffprobe/shaperprobe.html`, betrieben vom Georgia Tech College of Computing, Georgia, zuletzt besucht am 2012-08-29.

[37] S. Alich and P. Voigt, "Mitteilsame Browser – Datenschutzrechtliche Bewertung des Trackings mittels Browser-Fingerprints." in Computer & Recht CR, Jahrgang 2012, Issue 5, Seiten 344ff., Verlag Dr. Otto Schmidt, Köln, 2012.

[38] K. Tan, G. Yan, J. Yeo, and D. Kotz, "A Correlation Attack Against User Mobility Privacy in a Large-scale WLAN Network." in Proceedings des ACM Workshop on Wireless of the Students, by the Students, for the Students, Seiten 33ff., Chicago, 2010.

[39] K. Tan, G. Yan, J. Yeo, and D. Kotz, "Privacy analysis of user association logs in a large-scale wireless LAN." in Proceedings der 30. IEEE International Conference on Computer Communications IEEE INFOCOM 2011, Seiten 31ff., IEEE, Shanghai, 2011.

[40] J. Ghosh, M. J. Beal, H. Q. Ngo, and C. Qiao, "On profiling mobility and predicting locations of wireless users." in Proceedings des 2. International Workshop on Multi-Hop Ad Hoc Networks, Seiten 55ff., ACM, New York, 2006.

[41] S. Krüger, S.-A. Maucher, "Ist die IP-Adresse wirklich ein personenbezogenes Datum?." in Multimedia und Recht MMR, Jahrgang 2011, Issue 7, Seiten 433ff., C.H. Beck, München, 2011.

[42] C. H. Cap and A. Dähn and T. Mundt, "Network Neutrality – a Survey." in Proceedings der 5. International Conference on Evolving Internet INTERNET 2013, (Autoren in alphabetischer Reihenfolge), Seiten 56ff., Nizza, 2013.

[43] J. Postel, "Internet Protocol." Standard RFC 791, veröffentlicht durch IETF, 1981. Aktualisiert durch RFC 1349.

[44] J. MacKie-Mason and H. Varian, "Pricing the Internet." in Proceedings der Public Access to the Internet at JFK School of Government, EconWPA, 1993.

[45] J. MacKie-Mason and H. Varian, "Pricing Congestible Network Resources." in IEEE Journal on Selected Areas in Communications, Volume 13, Issue 7, Seiten 1141ff., 1995.

[46] B. van Schewick, "Internet Architecture and Innovation." Promotionsschrift, erschienen bei MIT Press (ISBN 026-201-397-5), Cambridge, 2010.

[47] C. Wright, L. Ballard, S. Coull, F. Monrose, and G. Masson, "Spot Me if You Can: Uncovering Spoken Phrases in Encrypted VoIP Conversations." in Proceedings des IEEE Symposium on Security and Privacy, IEEE, Seiten 35ff., 2008.

[48] J. Han, D. Watson, and F. Jahanian, "Topology aware overlay networks." in Proceedings der 24. Annual Joint Conference of the IEEE Computer and Communications Societies INFOCOM 2005, IEEE, Volume 4, Seiten 2554ff., 2005.

[49] V. Ciancaglini, L. Liquori, and G. N. Hoang, "Towards a Common Architecture to Interconnect Heterogeneous Overlay Networks." in Proceedings der 17. International Conference on Parralell and Distributed Systems ICPADS 2011, IEEE, Seiten 817ff., 2011.

[50] J. Ding, I. Balasingham, and P. Bouvry, "Management of Overlay Networks: A Survey," 2009. in Proceedings der 3. International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies UBICOMM '09, IARIA, Seiten 249ff., 2009.

[51] A. Weaver, J. Luo, and X. Zhang, "Monitoring and control using the Internet and Java." in Proceedings der 25. Annual Conference of the IEEE Industrial Electronics Society IECON '99, IEEE, Volume 3, Seiten 1152ff., 1999.

[52] L. Rizzo, "Dummynet: a simple approach to the evaluation of network protocols." in SIGCOMM Computer Communication Review, Volume 27, Issue 1, ACM, Seiten 31ff., New York, 1997.

[53] S. Horvath, "Definitionsansätze für den Begriff „Diskriminierungsfreiheit"im Zusammenhang mit der aktuellen Diskussion über Netzneutralität." in Wissenschaftliche Dienste des Deutschen Bundestages, Dokumentennummer WD 10 - 3000/014-11, Berlin, 2011.

[54] S. Even and R. Tarjan, "Network Flow and Testing Graph Connectivity." in SIAM Journal of Computing, Issue 4(4), Seiten 507ff., Society for Industrial and Applied Mathematics, Philadelphia, 1975.

[55] R. Srikant, "The Mathematics of Internet Congestion Control." Monographie, erschienen in der Reihe „Systems & Control: Foundations & Applications" bei Birkhäuser (ISBN 0-8176-3227-1), Boston/Basel/Berlin, 2004.

[56] C. Adams, P. Cain, D. Pinkas, and R. Zuccherato, "Internet X.509 Public Key Infrastructure Time-Stamp Protocol (TSP)." Standard RFC 3161, veröffentlicht durch IETF, 2001, Aktualisiert durch RFC 5816.

[57] S. Santesson and N. Pope, "ESSCertIDv2 Update for RFC 3161." Vorgeschlagener Standard RFC 5816, veröffentlicht durch IETF, 2010.

[58] B. Pan, H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka, "In Google We Trust: Users' Decisions on Rank, Position, and Relevance." in Journal of Computer-Mediated Communication, Volume 12, Issue 3, Blackwell Publishing Inc, Seiten 801ff., 2007.

[59] Berger, P. and Cap, C.H. and Brumme, R., "Überwachung des digitalen Raums. Verhaltensänderung von Internetnutzern." Zur Veröffentlichung angenommen in „Soziale Welt", 2014.

[60] N. Härting, "Rotlichtgerüchte: Haftet Google?," 2012. in Kommunikation und Recht K&R, Volume 15, Issue 10, Seiten 633ff., Deutscher Fachverlag GmbH, Frankfurt am Main, 2012.

[61] A. Dähn and U. Grohnwaldt, "Trafficshaping in Wohnheimen." Studienarbeit, Universität Rostock, Rostock, 2009.

# Weitere Quellen

In diesem Abschnitt finden sich Quellen, die nicht der klassischen „Wissenschaftlichen Literatur“ i.S. von Artikeln, Monographien oder Hochschularbeiten zugeordnet werden können.


[62] N.N., “Skype im Netz von T-Mobile: Abgehackte Gespräche ohne die Option „Internet Telefonie“.” Webresource, `http://layer9.wordpress.com/2013/03/13/skype-im-netz-von-t-mobile-abgehackte-gesprache-ohne-die-option-internet-telefonie/`, 2013, zuletzt besucht am 2014-03-08.

[63] A. Dähn, “There is something ugly going on in the network.” Webresource, `http://ad001.de/there-s-something-in-the-net.html`, 2013, zuletzt besucht am 2014-04-15.

[64] B. Schwan, “Netzneutralitätsabschaltung per API.” Webresource, `http://heise.de/-1378089`, zuletzt besucht am 2014-04-15, erschienen bei „Heise Online“, 2011.

[65] O. Reissmann, “Bevorzugte Dienste: Telekom bremst Spotify-Konkurrenz aus.” Webresource, `http://www.spiegel.de/netzwelt/netzpolitik/netzneutralitaet-was-der-telekom-spotify-deal-bedeutet-a-853246.html` zuletzt besucht am 2014-04-15, erschienen bei „Spiegel Online Netzwelt“, 2012.

[66] S. Hage, “Obermann will Google zur Kasse bitten.” Webresource, `http://www.manager-magazin.de/unternehmen/it/a-684172.html`, zuletzt besucht am 2014-04-15, erschienen bei „Manager Magazin Online“, 2010.

[67] N. Boeing, “Droht ein Zwei-Klassen-Internet?.” in „Die Zeit“, Ausgabe 6 des Jahres 2010.

[68] A. Wragge, “Zugangsprovider dürfen keine Hilfssheriffs werden.” Webresource, `http://www.golem.de/news/copyright/copyright-war-zugangsprovider-duerfen-keine-hilfssherrifs-werden-1207-93420.html`, zuletzt besucht am 2014-04-15, erschienen bei Golem.de, 2012.

[69] N.N., “Netzneutralität ist der Schlüssel zur Wahrung des feien Internets.” Webresource, `http://pro-netzneutralitaet.de/`, zuletzt besucht am 2013-06-14, mittlerweile (2014-04-15) verschollen, „Initiative pro Netzneutralität!“, 2010.

[70] N.N., "Configuring Load-Balance Per-Packet Action." Webresource, http://www.juniper.net/techpubs/software/junos/junos70/swconfig70-policy/html/policy-actions-config11.html, Juniper, zuletzt besucht am 2013-07-11.

[71] T. B. Lee, "The Journal Misunderstands Content-Delivery Networks." Webresource, http://www.freedom-to-tinker.com/blog/tblee/journal-misunderstands-content-delivery-networks/, in „Freedom to Tinker", 2008, zuletzt besucht am 2014-04-15.

[72] E. Felten, "Three Flavors of Net Neutrality." Webresource, http://www.freedom-to-tinker.com/blog/felten/three-flavors-net-neutrality, in „Freedom to Tinker", 2008, zuletzt besucht am 2014-04-15.

[73] S. Krempl, "Filesharing-bremsen der provider in aller welt beleuchtet." Webresource, http://heise.de/-1665358, in „Heise Online", 2012, zuletzt besucht am 2014-04-15.

[74] N.N., "Initiative Netzqualität: Netzneutralitäts-Test," 2013. Webresource, http://www.initiative-netzqualitaet.de/netzneutralitaetstest/, zafaco GmbH im Auftrag der Bundesnetzagentur, 2013, zuletzt besucht am 2014-04-15.

[75] N. Walter, "Neutralität kontra Kontrolle – Das Ringen um die Freiheit des Internets." Webresource, http://www.fkyter.de/de/120/thema/11365/, erschienen in „Fluter.", 2013, zuletzt besucht am 2014-04-30.

[76] N.N., "Tor Project: Anonymity Online," 2013. Webresource, https://www.torproject.org/, zuletzt besucht 2014-04-15.

[77] C. Duhigg, "How Companies Learn Your Secrets." Webresource, http://nytimes.com/2012/02/19/magazine/shopping-habits.html, in „The New York Times Magazine", 2012, zuletzt besucht 2014-04-15.

[78] T. Spring, "Search engines gang up on Microsoft." Webresource, http://www.cnn.com/TECH/computing/9911/15/search.engine.ms.idg/, in „CNN.com", 1999, zuletzt besucht 2014-04-15.

[79] S. Hansell, "AltaVista invites advertisers to pay for top ranking." in „The New York Times", Ausgabe vom 1999-04-15.

[80] J. Ihlenfeld, "Google setzt BMW vor die Tür." Webresource, http://www.golem.de/0602/43155.html, zuletzt besucht am 2014-04-15, erschienen bei Golem.de, 2006.

[81] K. Biermann, "Vor dem Modem sind längst nicht alle gleich." Webresource, http://www.zeit.de/digital/internet/2009-12/netzneutralitaet-filesharing, erschienen in „Zeit Online", 2009, zuletzt besucht am 2014-04-15.

[82] C. Fiedler, "Pressefreiheit nur noch auf dem Papier." erschienen in Promedia, Hrsg. vom Verband Deutscher Zeitschriftenverleger VDZ, Ausgabe 1 2014, Seiten 6f., 2014.

[83] J. Almunia, "Statement of VP Almunia on the Google antitrust investigation." Webresource, `http://europa.eu/rapid/press-release_SPEECH-12-372_en.htm?locale=en`, 2012, zuletzt besucht am 2014-04-15.

[84] N.N., "Verwenden der Gruppenrichtlinie zum Verteilen von Zertifikaten." Webresource, `http://technet.microsoft.com/de-de/library/cc772491(v=ws.10).aspx`, erschienen in „Microsoft Technet", zuletzt besucht am 2014-04-15.

[85] Y. Tse, "Is Yahoo Mail blocking emails that mention Occupy Wall Street?." Webresource, `https://100gf.wordpress.com/2011/09/19/is-yahoo-mail-blocking-emails-that-mention-occupy-wall-street-occupywallstreet/`, 2011, zuletzt besucht am 2014-04-15.

[86] Yahoo, "We apologize 4 blocking 'occupywallst.org' It was not intentional & caught by our spam filters. It is resolved, but may be a residual delay." Webresource, `http://twitter.com/#!/YahooCare/status/116220596987232256`, 2011, zuletzt besucht am 2014-04-15.

[87] N.N., "Automatische Vervollständigung." Webresource, `http://support.google.com/websearch/bin/answer.py?hl=de&answer=106230`, zuletzt besucht am 2014-05-15.

[88] H. Schmale, "Die Mühen der Wulffs." Webresource, `http://www.fr-online.de/bettina-wulff/bettina-wulff-gegen-google-die-muehen-der-wulffs,17221048,20790384.html`, erschienen in „Frankfurter Rundschau ", 2012, zuletzt besucht am 2014-04-15.

[89] N.N., "Bettina Wulff begrüßt Urteil gegen Google." Webresource, `http://www.fr-online.de/bettina-wulff/google-autocomplete-bettina-wulff-begruesst-urteil-gegen-google,17221048,22765802.html`, erschienen in „Frankfurter Rundschau", 2013, zuletzt besucht am 2014-04-15.

[90] Mike Kortsch, "Klare Feindbilder – Wie sich Konzerne gegen Recherche wehren." Webresource, `http://www.ndr.de/fernsehen/sendungen/zapp/medien_politik_wirtschaft/zapp666.html`, erschienen in „NDR.de", 2008, zuletzt besucht am 2014-04-15.

[91] S. Levy, "Inside Google's China misfortune." Webresource, `http://tech.fortune.cnn.com/2011/04/15/googles-ordeal-in-china/`, erschienen bei „CNN Money", 2011, zuletzt besucht am 2014-04-15.

# Selbständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Ich erkläre, dass in der Arbeit verwendete fremde Quellen als solche kenntlich gemacht wurden, sowohl im Fall wörtlicher Übernahmen wie auch bei Paraphrasen.

Rostock, den _____