Beobachtungsmöglichkeiten im Domain Name System

Angriffe auf die Privatsphäre und Techniken zum Selbstdatenschutz

Disputation

Dominik Herrmann

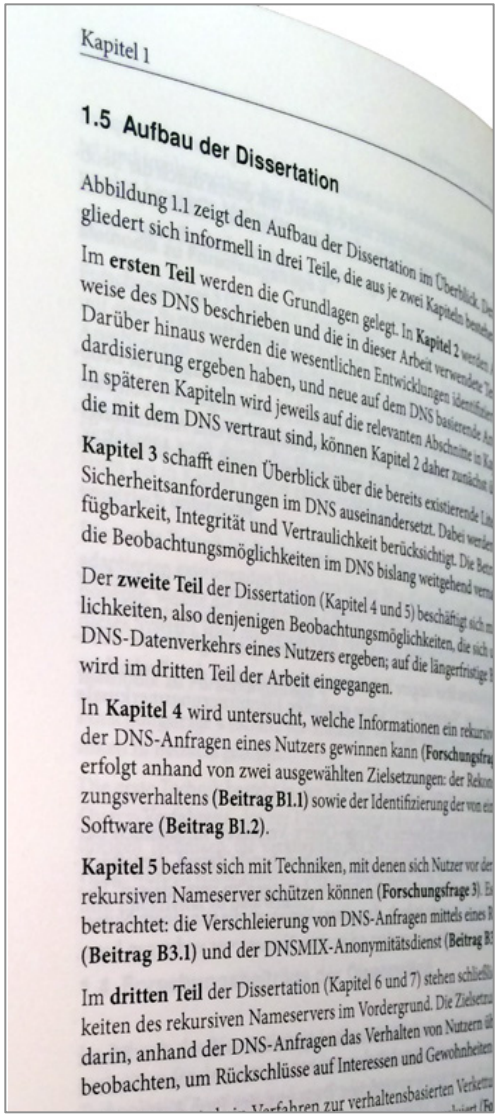
Hamburg, 7. April 2014

AGENDA

Ausgangssituation und Forschungsfragen

Ausgewählte Forschungsbeiträge

Schlussfolgerungen und Ausblick



Kapitel 1

1.5 Aufbau der Dissertation

Abbildung 1.1 zeigt den Aufbau der Dissertation im Überblick. Die
gliedert sich informell in drei Teile, die aus je zwei Kapiteln bestehen

Im **ersten Teil** werden die Grundlagen gelegt. In **Kapitel 2** werden
weise des DNS beschrieben und die in dieser Arbeit verwendete Te
Darüber hinaus werden die wesentlichen Entwicklungen identifizier
dardisierung ergeben haben, und neue auf dem DNS basierende An
In späteren Kapiteln wird jeweils auf die relevanten Abschnitte in Ka
die mit dem DNS vertraut sind, können Kapitel 2 daher zunächst ü

Kapitel 3 schafft einen Überblick über die bereits existierende Lite
Sicherheitsanforderungen im DNS auseinandersetzt. Dabei werden
fügbarkeit, Integrität und Vertraulichkeit berücksichtigt. Die Betr
die Beobachtungsmöglichkeiten im DNS bislang weitgehend vernac

Der **zweite Teil** der Dissertation (Kapitel 4 und 5) beschäftigt sich m
lichkeiten, also denjenigen Beobachtungsmöglichkeiten, die sich
DNS-Datenverkehrs eines Nutzers ergeben; auf die längerfristige B
wird im dritten Teil der Arbeit eingegangen.

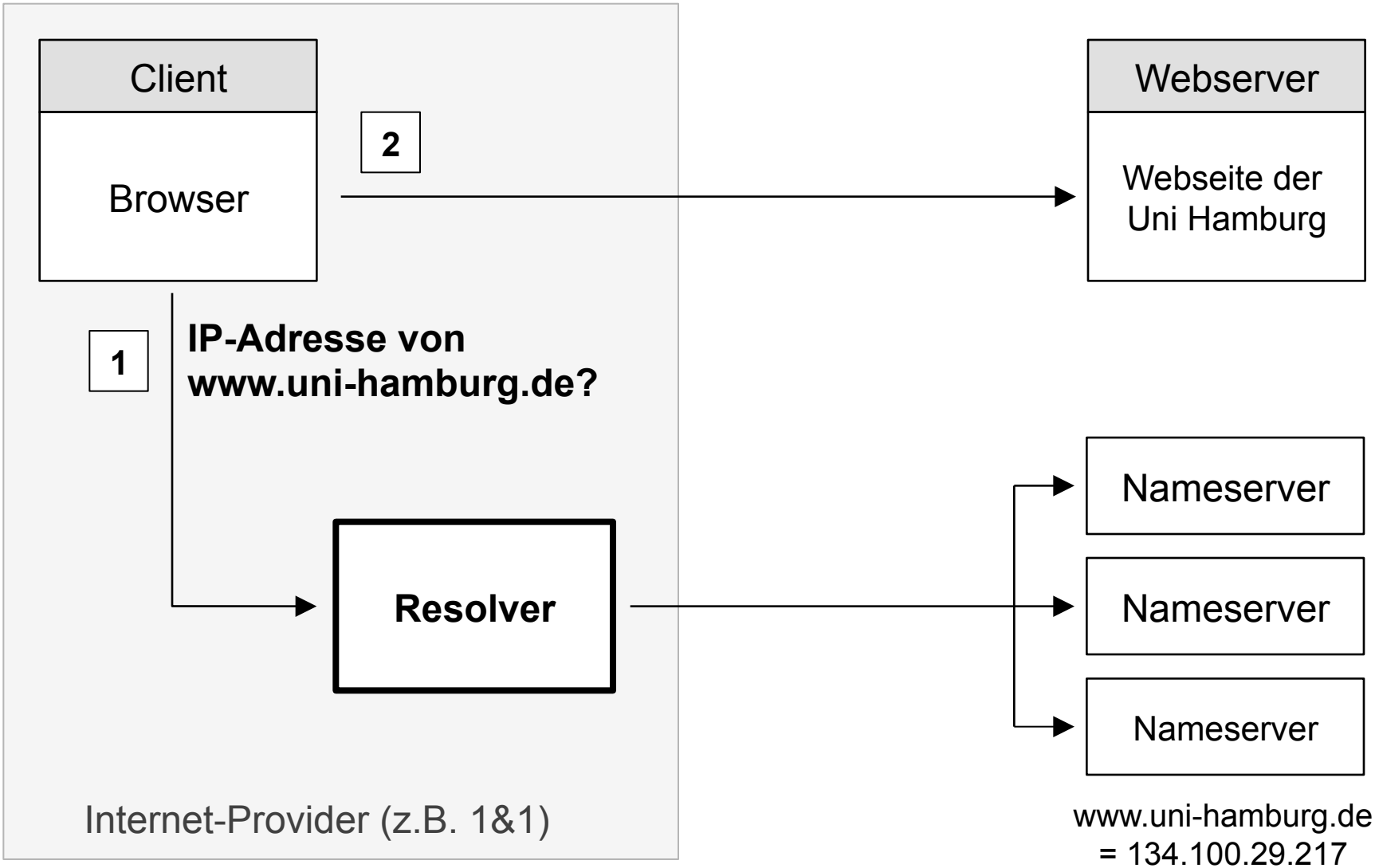
In **Kapitel 4** wird untersucht, welche Informationen ein rekursiv
der DNS-Anfragen eines Nutzers gewinnen kann (**Forschungsfrag**
erfolgt anhand von zwei ausgewählten Zielsetzungen: der Rekon
zungsverhaltens (**Beitrag B1.1**) sowie der Identifizierung der von ein
Software (**Beitrag B1.2**).

Kapitel 5 befasst sich mit Techniken, mit denen sich Nutzer vor der
rekursiven Nameserver schützen können (**Forschungsfrage 3**). Es
betrachtet: die Verschleierung von DNS-Anfragen mittels eines R
(**Beitrag B3.1**) und der DNSMIX-Anonymitätsdienst (**Beitrag B3**

Im **dritten Teil** der Dissertation (Kapitel 6 und 7) stehen schließli
keiten des rekursiven Nameservers im Vordergrund. Die Zielsetzu
darin, anhand der DNS-Anfragen das Verhalten von Nutzern üb
beobachten, um Rückschlüsse auf Interessen und Gewohnheiten

Verfahren zur verhaltensbasierten Verkettu

Im Fokus der Disputation stehen die Beobachtungsmöglichkeiten auf dem DNS-Resolver, an den die Namensauflösung delegiert wird.



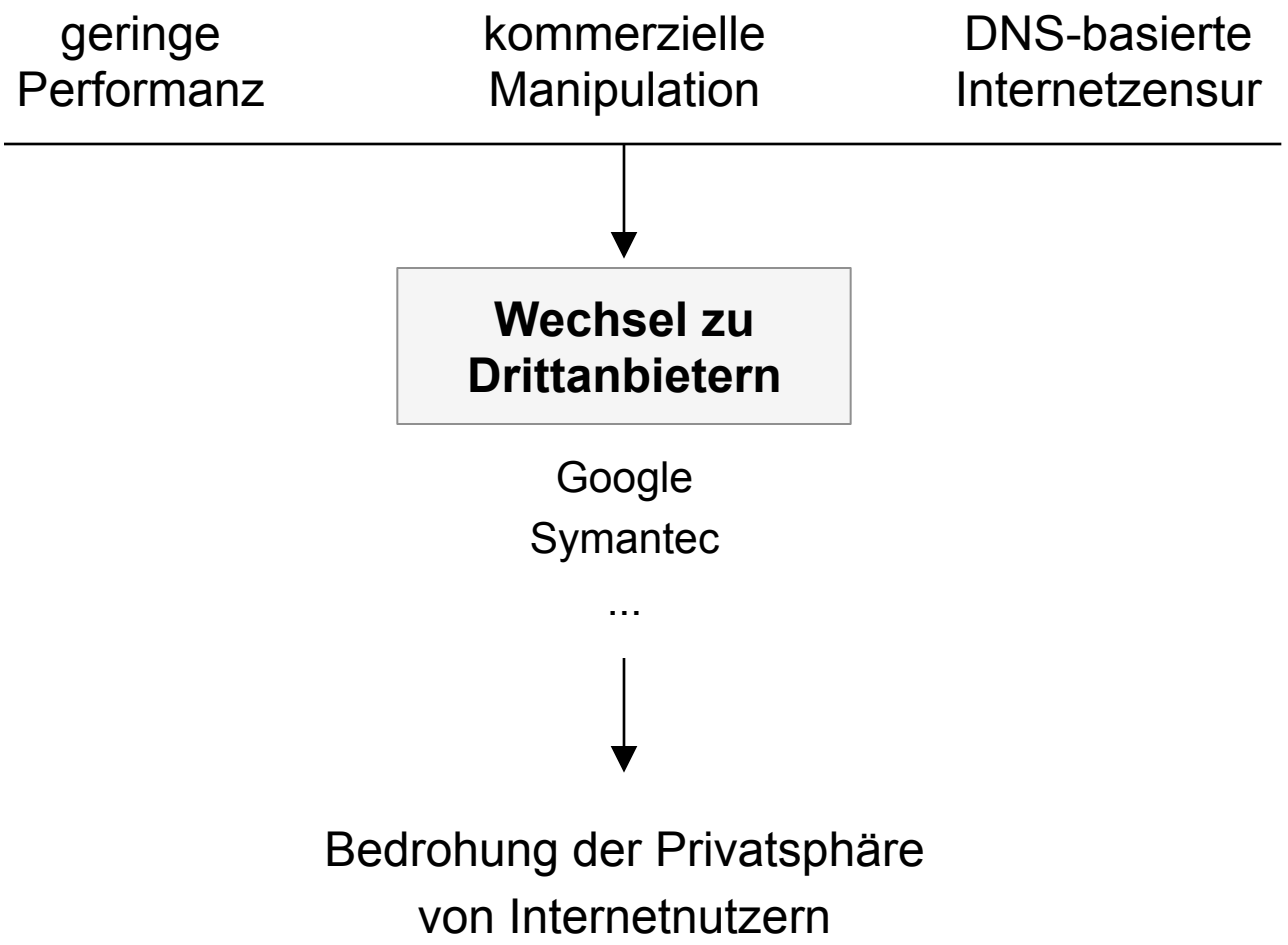
Das Schutzziel Vertraulichkeit wird im DNS bislang vernachlässigt.

Denial-of-Service u.a.	**Verfügbarkeit**	Redundanz, Verteilung
DNS-Spoofing u.a.	**Integrität**	zukünftig DNSSEC
unklare Bedrohung	**Vertraulichkeit**	**keine Mechanismen**

„DNS-Anfragen nur Metadaten“

„Internet-Provider kann sowieso überwachen“

Durch veränderte Rahmenbedingungen könnte die Relevanz der Beobachtungsmöglichkeiten im DNS in Zukunft steigen.



Die Dissertation adressiert 4 Forschungsfragen zur Quantifizierung des Bedrohungspotenzials und zur Gestaltung von Schutzmaßnahmen.

1. Monitoring

[...] Welche Informationen kann ein DNS-Resolver bei der Adressauflösung gewinnen?

2. Tracking

Inwiefern kann ein DNS-Resolver das Verhalten von Nutzern auch dann verfolgen, wenn sie unter verschiedenen IP-Adressen auftreten? [...]

3. Schutz vor Monitoring

Welche Maßnahmen können Nutzer ergreifen? Wie sind praxistaugliche datenschutzfreundliche Techniken zu gestalten; wie effektiv sind sie?

4. Schutz vor Tracking

[...] Welche Schutzmaßnahmen können Nutzer ergreifen und wie effektiv sind diese?

AGENDA

Ausgangssituation und Forschungsfragen

Ausgewählte Forschungsbeiträge

Monitoring

Ermittlung besuchter Webseiten

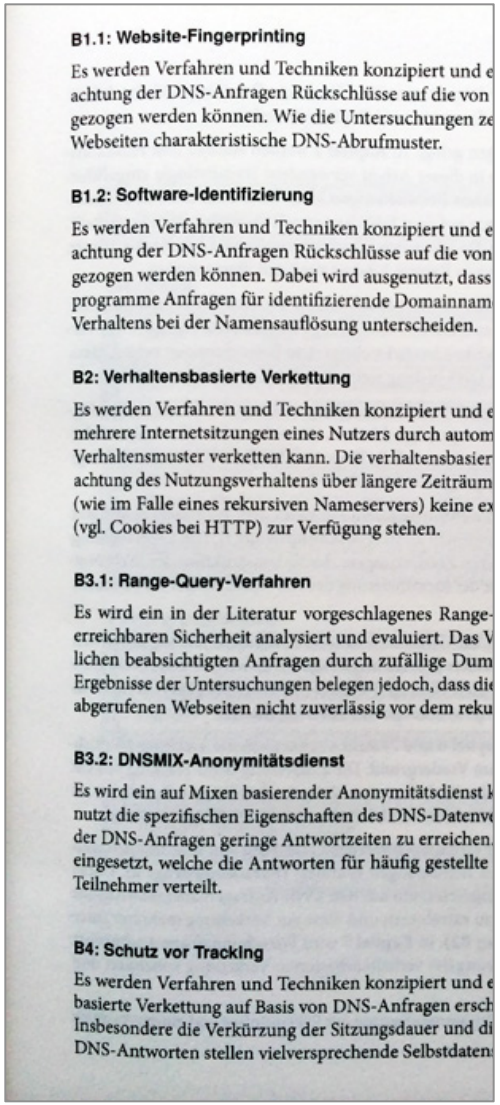
Schutz durch DNSMIX-Dienst

Tracking

Verhaltensbasierte Verkettung

Schutz durch häufigen Adresswechsel

Schlussfolgerungen und Ausblick



B1.1: Website-Fingerprinting

Es werden Verfahren und Techniken konzipiert und e
achtung der DNS-Anfragen Rückschlüsse auf die von
gezogen werden können. Wie die Untersuchungen ze
Webseiten charakteristische DNS-Abrufmuster.

B1.2: Software-Identifizierung

Es werden Verfahren und Techniken konzipiert und e
achtung der DNS-Anfragen Rückschlüsse auf die von
gezogen werden können. Dabei wird ausgenutzt, dass
programme Anfragen für identifizierende Domainnam
Verhaltens bei der Namensauflösung unterscheiden.

B2: Verhaltensbasierte Verkettung

Es werden Verfahren und Techniken konzipiert und e
mehrere Internetsitzungen eines Nutzers durch autom
Verhaltensmuster verketten kann. Die verhaltensbasier
achtung des Nutzungsverhaltens über längere Zeiträum
(wie im Falle eines rekursiven Nameservers) keine ex
(vgl. Cookies bei HTTP) zur Verfügung stehen.

B3.1: Range-Query-Verfahren

Es wird ein in der Literatur vorgeschlagenes Range-
erreichbaren Sicherheit analysiert und evaluiert. Das V
lichen beabsichtigten Anfragen durch zufällige Dum
Ergebnisse der Untersuchungen belegen jedoch, dass die
abgerufenen Webseiten nicht zuverlässig vor dem reku

B3.2: DNSMIX-Anonymitätsdienst

Es wird ein auf Mixen basierender Anonymitätsdienst k
nutzt die spezifischen Eigenschaften des DNS-Datenv
der DNS-Anfragen geringe Antwortzeiten zu erreichen.
eingesetzt, welche die Antworten für häufig gestellte
Teilnehmer verteilt.

B4: Schutz vor Tracking

Es werden Verfahren und Techniken konzipiert und e
basierte Verkettung auf Basis von DNS-Anfragen ersch
Insbesondere die Verkürzung der Sitzungsdauer und di
DNS-Antworten stellen vielversprechende Selbstdaten

Inwiefern kann ein DNS-Resolver nachvollziehen, welche Internetseiten seine Nutzer aufrufen?

„Nutzen“ für den DNS-Resolver

- Zielgerichtetere Werbung
- Möglichkeit der Diskriminierung

Zwei Herausforderungen

- Abruf von Webseiten korrespondiert nicht mit DNS-Anfragen
- nur Domains beobachtbar, jedoch keine URLs

http://de.wikipedia.org/wiki/Alkoholkrankheit

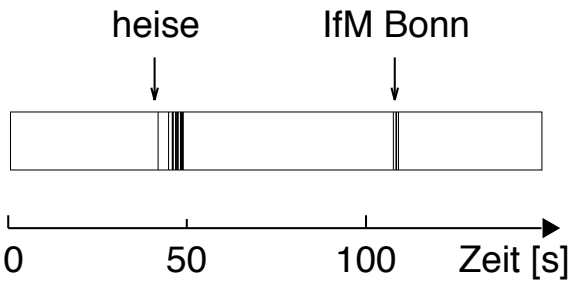
Wegen der vielen DNS-Anfragen ist eine zuverlässige Ermittlung der abgerufenen Webseiten nicht ohne Weiteres möglich.

Präemptive Namensauflösung

Inhalte fremder Webserver

www.heise.de/-1973600

www.heise.de
script.ioam.de
heise.ivwbox.de
ad-emea.doubleclick.net
www.destatis.de
www.ifm-bonn.org
www.mywai.com
www.dci.de... [48 Domains]



www.ifm-bonn.org

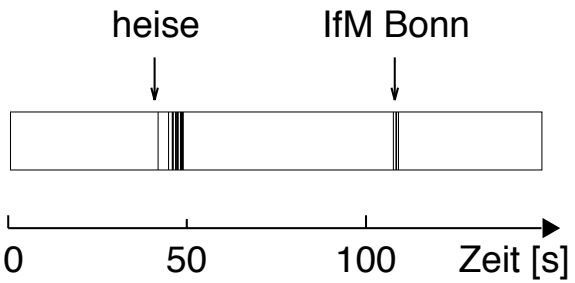
s7.addthis.com
ct1.addthis.com
www.google-analytics.com
en.ifm-bonn.org
m.addthis.com

Wegen der vielen DNS-Anfragen ist eine zuverlässige Ermittlung der abgerufenen Webseiten nicht ohne Weiteres möglich.

HEURISTIK	RESULTAT
1. www.[???].[???]	**www.heise.de** ✚ **www.ifm-bonn.org** ✚ www.destatis.de ✖ www.mymai.com ✖ www.dci.de ✖ www.google-analytics.com ✖
2. Mindestabstand	**www.heise.de** ✚ s7.addthis.com ✖

www.heise.de/-1973600

www.heise.de
script.ioam.de
heise.ivwbox.de
ad-emea.doubleclick.net
www.destatis.de
www.ifm-bonn.org
www.mywai.com
www.dci.de... [48 Domains]

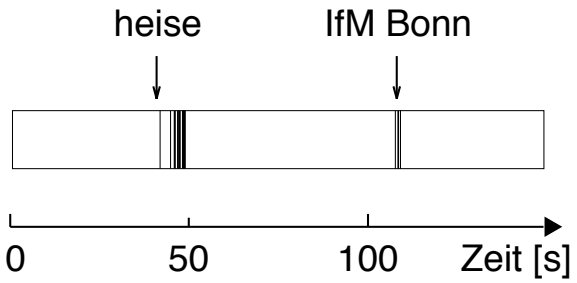


www.ifm-bonn.org

s7.addthis.com
ct1.addthis.com
www.google-analytics.com
en.ifm-bonn.org
m.addthis.com

Durch das vorgeschlagene **Website-Fingerprinting-Verfahren** könnte der DNS-Resolver die Nutzeraktivitäten dennoch nachvollziehen.

www.heise.de
script.ioam.de
heise.ivwbox.de
ad-emea.doubleclick.net
www.destatis.de
www.ifm-bonn.org
www.mywai.com
www.dci.de... [48 Domains]



s7.addthis.com
ct1.addthis.com
www.google-analytics.com
en.ifm-bonn.org
m.addthis.com

WEBSEITE	ABRUFMUSTER	ÜBEREINSTIMMUNG	
heise.de/-1973600	**www.heise.de, www.destatis.de, www.ifm-bonn.org,** ...	**100%**	✚
IfM	**www.ifm-bonn.org, en.ifm-bonn.org, m.addthis.com,**...	**100%**	✚
DCI AG	**www.dci.de**, codes.wai.it, mailer.dci.de, data.dci-se.de, ...	**5%**	✚
Stat. Bundesamt	**www.destatis.de**	**100%**	✖

Ergebnis der Experimente: viele Webseiten haben einzigartige Abfrufmuster.

Neben den besuchten Webseiten können DNS-Anfragen auch andere sensible Informationen preisgeben, wie weitere Untersuchungen zeigen.

RÜCKSCHLUSS AUF	BEISPIEL
Hauptseiten	www.magersucht.de
Unterseiten	de.wikipedia.org/wiki/AIDS
Verwendetes Betriebssystem	Apple MacOS X
Verwendeter Web-Browser	Mozilla Firefox
Verwendete Anwendungen	Virenscanner vorhanden?
Eigener Endgeräte-Name	dominik-pc.arbeitsgruppe
Weitere Endgeräte-Namen	hp-laserjet2100.fritz.box

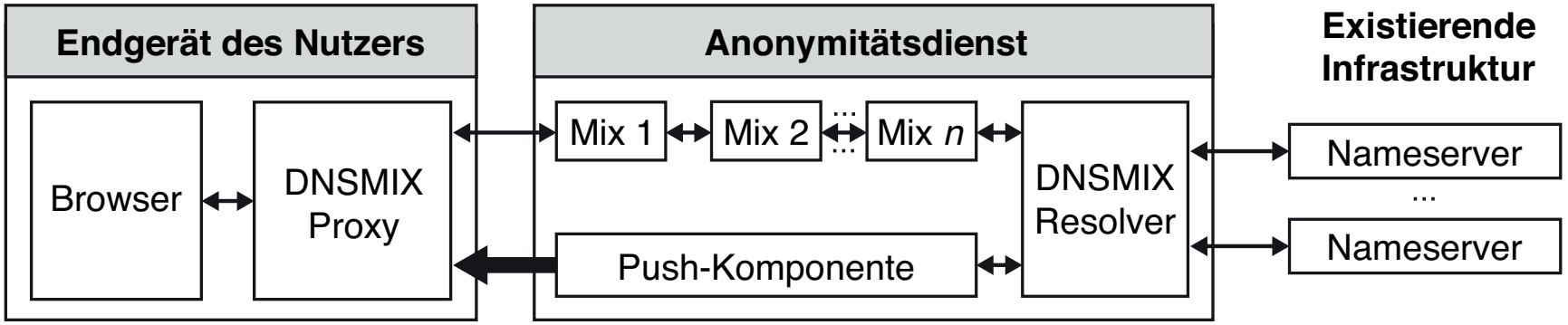
In der Dissertation werden verschiedene Techniken zum Schutz vor Monitoring analysiert.

Verschlüsselung der Anfragen	Verzicht auf Resolver	Verschleierung mit Dummy-Anfragen
↓	↓	↓
schwierig	nicht ausreichend	aufwändig

Mix-Netze: OK, aber langsam
DNS über Tor: 1,4s [Fab+10]

→

DNSMIX-Dienst mit Push-Komponente



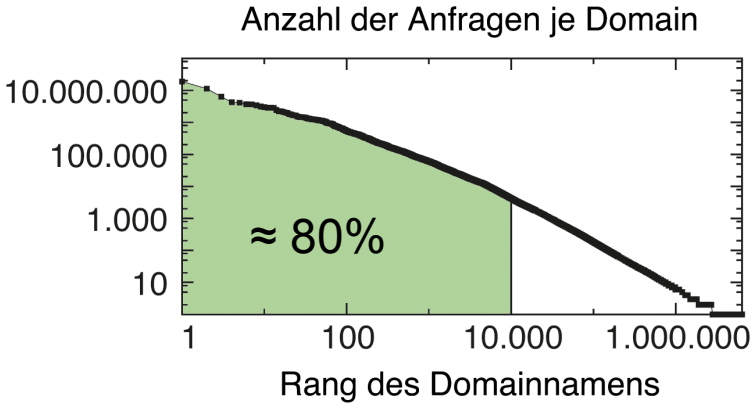
Die Push-Komponente erlaubt völlig verzögerungsfreie Auflösung und Unbeobachtbarkeit für Großteil der Anfragen bei vertretbarem Aufwand.

Domains sind relativ stabil

Studie von [HG05] mit 37.000 Domains: 87% unverändert innerhalb von 57 Tagen

Benutzerverhalten folgt Potenzgesetz

Studie mit 4000 Nutzern über 5 Monate



AGENDA

Ausgangssituation und Forschungsfragen

Ausgewählte Forschungsbeiträge

Monitoring

Ermittlung besuchter Webseiten

Schutz durch DNSMIX-Dienst

Tracking

Verhaltensbasierte Verkettung

Schutz durch häufigen Adresswechsel

Schlussfolgerungen und Ausblick

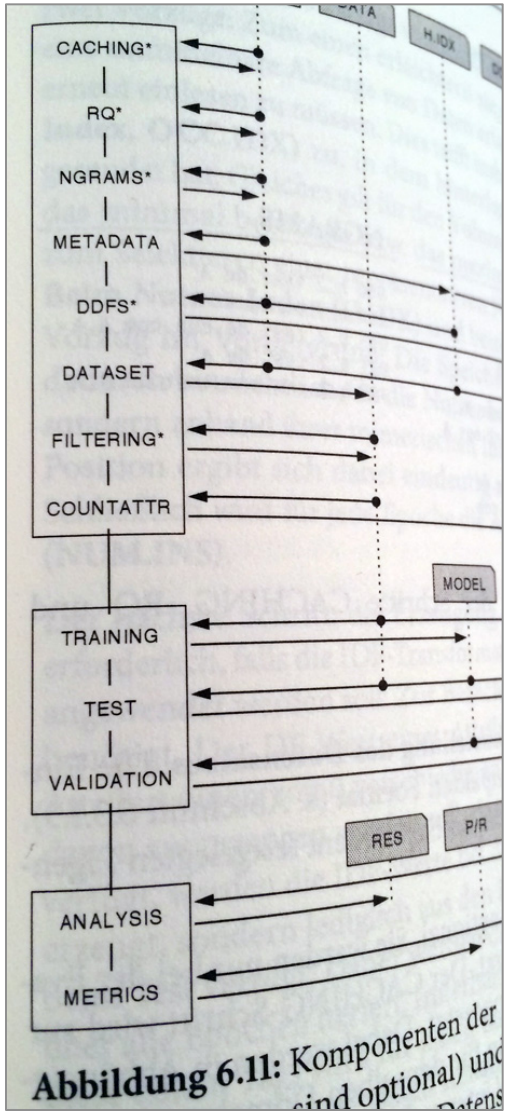
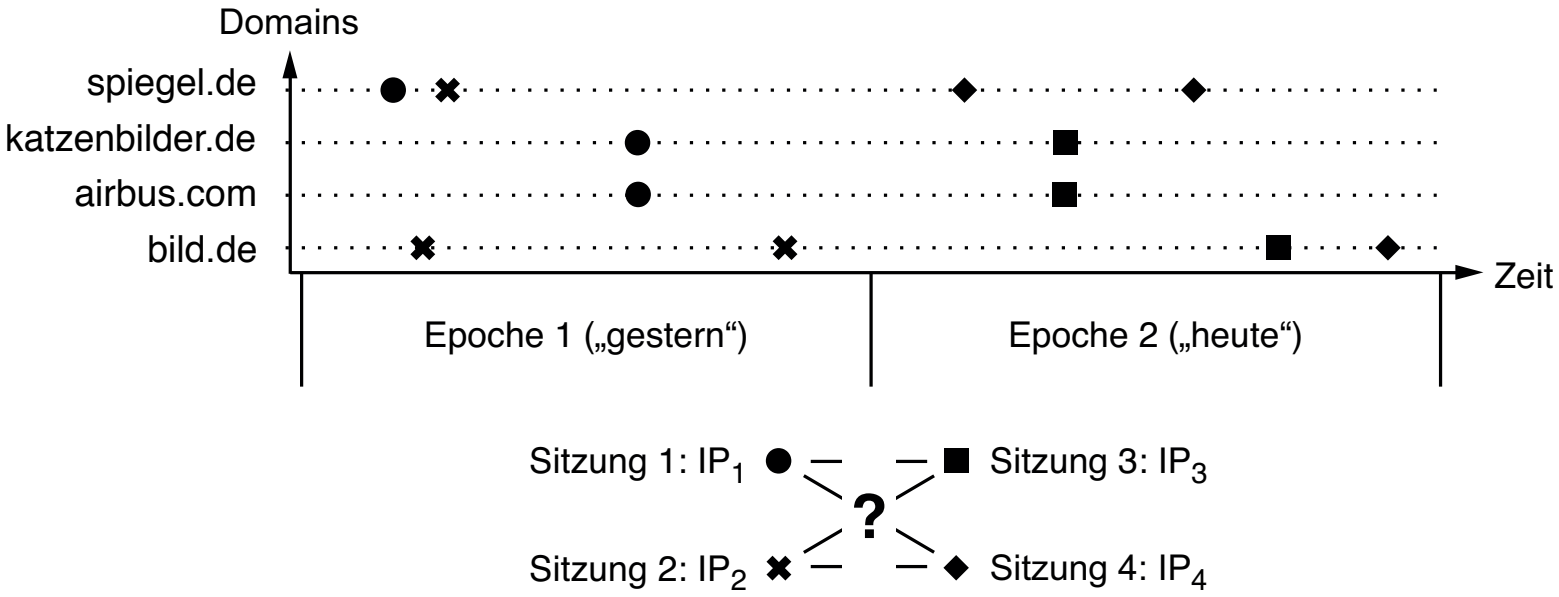


Abbildung 6.11: Komponenten der

Wegen dynamisch zugewiesener IP-Adressen kann der DNS-Resolver die Aktivitäten seiner Nutzer nicht über längere Zeiträume beobachten.



Reichen die DNS-Anfragen aus <u>einer</u> Sitzung, um Nutzer wiederzuerkennen?

Vermutungen

- Menschen sind Individuen
- Menschen haben Gewohnheiten

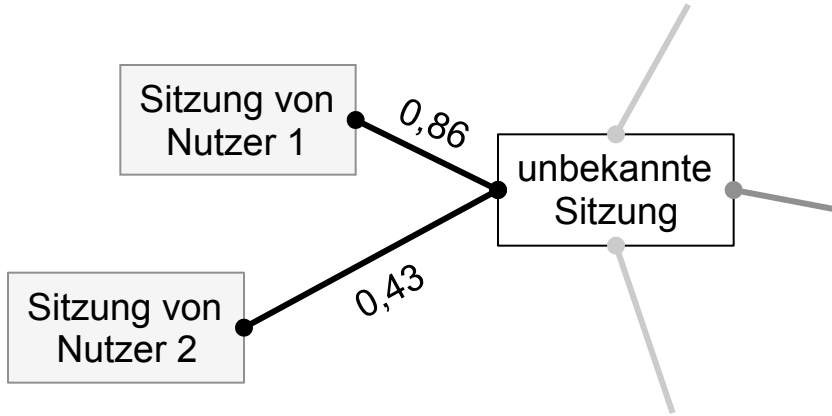
Bislang beste Arbeit

100 Nutzer, je 200 bekannte Sitzungen: 62% korrekt [Yan10]

Es wird ein **Verkettungsverfahren** vorgeschlagen u. implementiert, ein **Datensatz** erhoben und eine realitätsnahe **Evaluation** durchgeführt.

Verfahren des überwachten Lernens

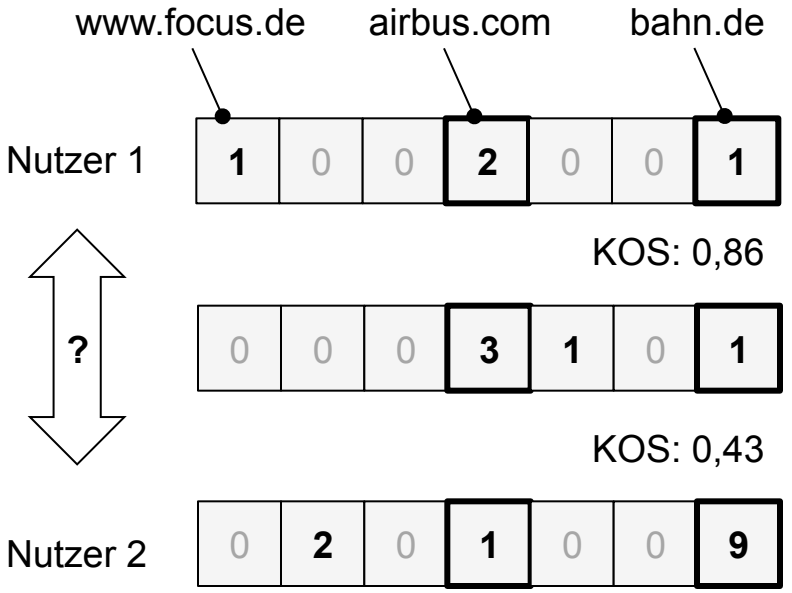
- u.a. 1-Nächste-Nachbarn (1NN) mit Jaccard- und Kosinus-Ähnlichkeit
- Normalisierung, IDF, n-Gramme
- Implementierung mit MapReduce



Datenverkehr für Evaluation

- 61 Tage, 3862 Nutzer
- 431 Mio. Anfragen, 5 Mio. Domains
- statische Pseudonyme

```
1271195950.882 fe9 google.de A
```



Ergebnisse deuten darauf hin, dass die Sitzungsverkettung mit DNS-Anfragen gelingt, wenn IP-Adressen der Nutzer täglich wechseln.

Evaluation mit Sitzungen von 24h

Tag-zu-Tag-Zuordnung über 61 Tage

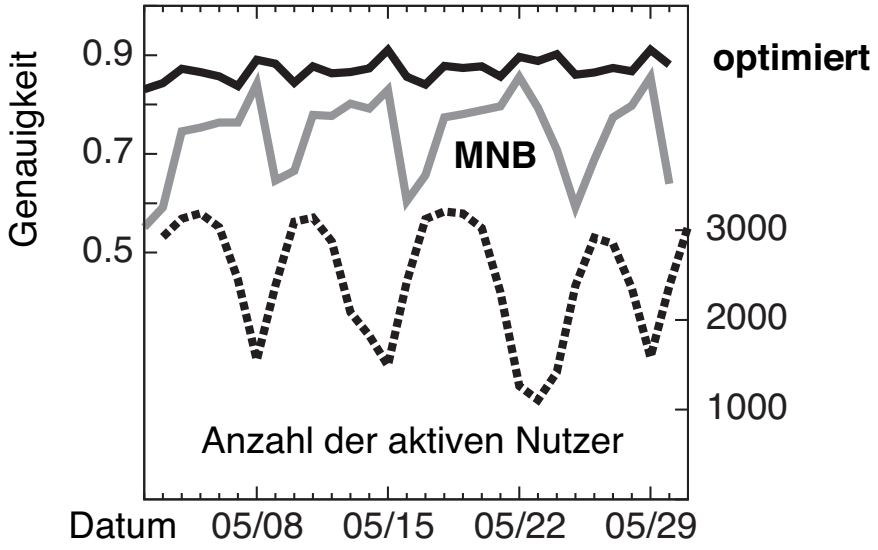
1100 – 3200 Nutzer pro Tag

VERFAHREN	GENAUIGKEIT
1NN JAC	66%
1NN KOS	75%

Optimierung

Berücksichtigung der Nutzerfluktuation

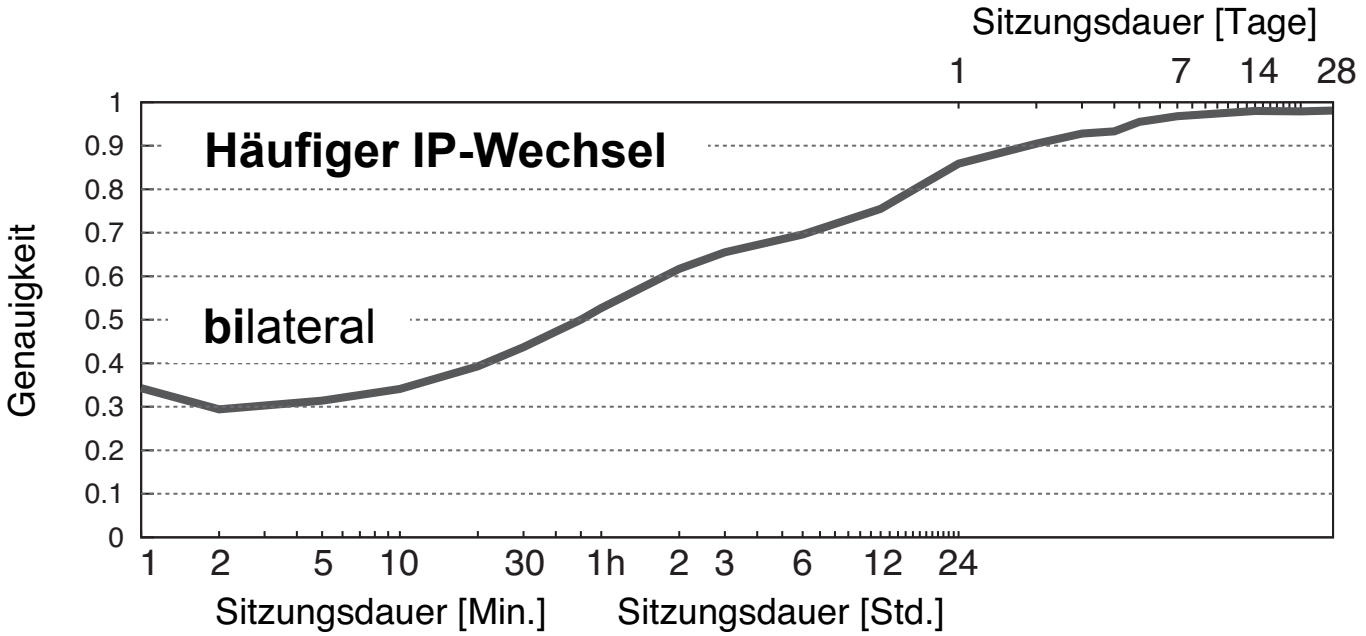
Genauigkeit >86%



Ergebnisse weiterer Analysen deuten darauf hin, dass das Verfahren zur verhaltensbasierten Verkettung von Sitzungen robust ist.

FRAGESTELLUNG	ERGEBNIS
Verkettung bei allen Nutzern möglich?	nein: f. 90% d. Nutzer >60%
Alle DNS-Anfragen erforderlich?	nein: bei 500 pop. Domains 80%
Bei größerer Nutzergruppe möglich?	ja: bei 12.000 Nutzern ca. 80%
Mit alten Trainingsdaten möglich?	ja: nach 28 Tagen ca. 76%
Rufen Nutzer tägl. dieselben Seiten ab?	nein (nicht notwendig), aber: **benutzerspezifische Domains**

Von den analysierten Maßnahmen ist v.a. häufiges Wechseln der IP-Adresse zum Schutz vor Tracking vielversprechend.



DNSMIX-Dienst → **multi**lateral

Verlängertes Caching → **uni**lateral

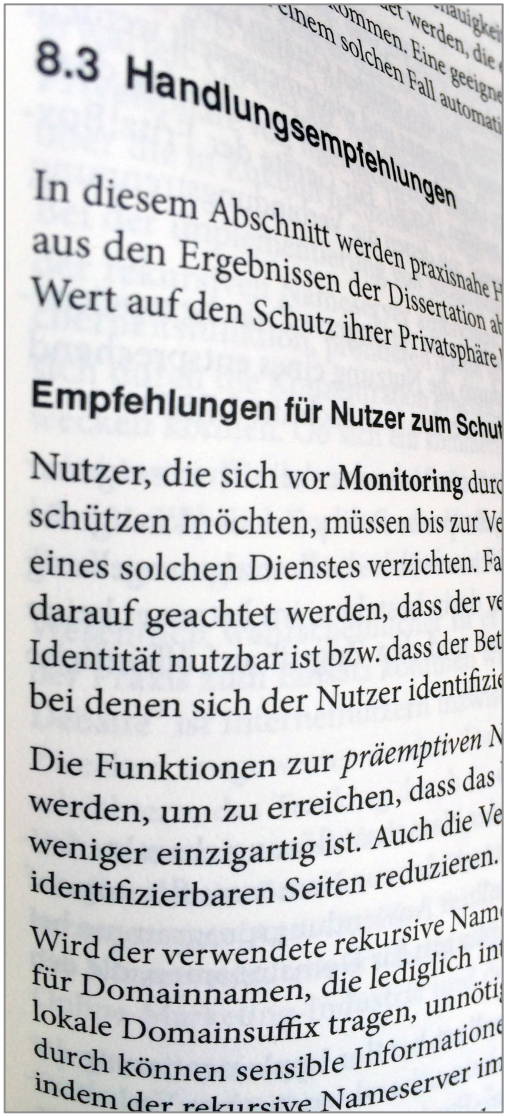
Verschleierung mit Dummy-Anfragen → wenig effektiv

AGENDA

Ausgangssituation und Forschungsfragen

Ausgewählte Forschungsbeiträge

Schlussfolgerungen und Ausblick



8.3 Handlungsempfehlungen

In diesem Abschnitt werden praxisnahe
aus den Ergebnissen der Dissertation
Wert auf den Schutz ihrer Privatsphäre

Empfehlungen für Nutzer zum Schut

Nutzer, die sich vor **Monitoring** durc
schützen möchten, müssen bis zur Ve
eines solchen Dienstes verzichten. Fa
darauf geachtet werden, dass der ve
Identität nutzbar ist bzw. dass der Bet
bei denen sich der Nutzer identifizie

Die Funktionen zur *präemptiven* N
werden, um zu erreichen, dass das
weniger einzigartig ist. Auch die Ve
identifizierbaren Seiten reduzieren.

Wird der verwendete rekursive Nam
für Domainnamen, die lediglich in
lokale Domainsuffix tragen, unnöti
durch können sensible Informatione
indem der rekursive Nameserver im

Die Ergebnisse der Dissertation sind – nicht nur – für DNS-Nutzer, die sich um ihre Privatsphäre sorgen, von Bedeutung.

Verhaltensbasierte Verkettung praktikabel und generalisierbar	Gefahr →	Nächstes Supercookie?
Dynamische IP-Adressen bieten kaum Schutz	Chance →	Möglichkeit zum häufigen Wechsel mit IPv6
DNS-Anfragen verraten besuchte Webseiten und verwendete Software	Chance →	Neue Analysetechniken bei forensischen Untersuchungen
Unbeobachtbare, schnelle Auflösung durch Push	Future Work →	Praxistaugliche Implementierung

Beobachtungsmöglichkeiten im DNS
Angriffe auf die Privatsphäre und Techniken zum Selbstdatenschutz

FORSCHUNGSFRAGEN

FF1 Monitoring-Moglichkeiten?
Website-Fingerprinting

FF2 Tracking-Möglichkeiten?
Verhaltensbasierte Verkettung

FF3 Schutz vor Monitoring?
DNSMIX mit Push-Komponente

FF4 Schutz vor Tracking?
Häufiger Adresswechsel

THESEN AUS VORTRAG

- Vertraulichkeit im DNS vernachlässigt
- Webseiten haben DNS-Abrufmuster
- DNSMIX bietet geringe Verzögerung
- Unbeobachtbarkeit für Top 10.000
- Individuelle Gewohnheiten der Nutzer erlauben Langzeitverfolgung
- Erkenntnisse übertragbar und von übergeordneter Bedeutung

Die Forschungsbeiträge wurden bereits teilweise publiziert und auf internationalen Konferenzen und Workshops präsentiert.

JOURNALS

(mit H. Federrath & C. Gerber) Verhaltensbasierte Verkettung von Internetsitzungen, **Datenschutz und Datensicherheit** 35.11, 2001.

(mit C. Banse & H. Federrath) Behavior-based Tracking: Exploiting Characteristic Patterns in DNS Traffic, **Computers & Security** 39A, 2013.

CONFERENCES & WORKSHOPS (PEER-REVIEWED)

(mit H. Federrath & R. Wendolsky) Website Fingerprinting: Attacking Popular Privacy Enhancing Technologies with the Multinomial Naïve-Bayes Classifier, **CCSW 2009, ACM.**

(mit C. Banse, H. Federrath & C. Gerber) Analyzing Characteristic Host Access Patterns for Re-identification of Web User Sessions, **NordSec 2010**, **Springer**.

CONFERENCES & WORKSHOPS (PEER-REVIEWED)

(mit H. Federrath, K.-P. Fuchs & C. Piosecny) Privacy-Preserving DNS: Analysis of Broadcast, Range Queries and Mix-Based Protection Methods, **ESORICS 2011, Springer**.

(mit C. Banse & H. Federrath) Tracking Users on the Internet with Behavioral Patterns: Evaluation of Its Practical Feasibility, **IFIP SEC 2012, Springer**.

(mit C. Arndt & H. Federrath) IPv6 Prefix Alteration: An Opportunity to Improve Online Privacy, **PDPT Workshop, Amsterdam, 2012**.

(mit H. Federrath & K.-P. Fuchs) Fingerprinting Techniques for Target-oriented Investigations in Network Forensics, **SICHERHEIT 2014, GI**.

(mit H. Federrath & M. Maaß) Evaluating the Security of a DNS Query Obfuscation Scheme for Private Web Surfing, **IFIP SEC 2014, Springer**.