

Beobachtungsmöglichkeiten im Domain Name System

Angriffe auf die Privatsphäre und
Techniken zum Selbstdatenschutz

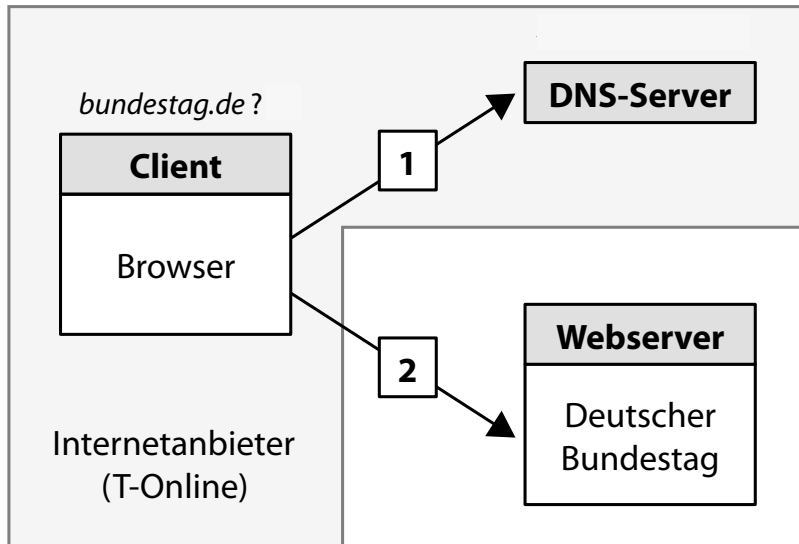
Dr. Dominik Herrmann
Universität Hamburg



Download Handout und Folien
<http://dhgo.to/castfolien> bzw. [/casthandout](http://dhgo.to/casthandout)

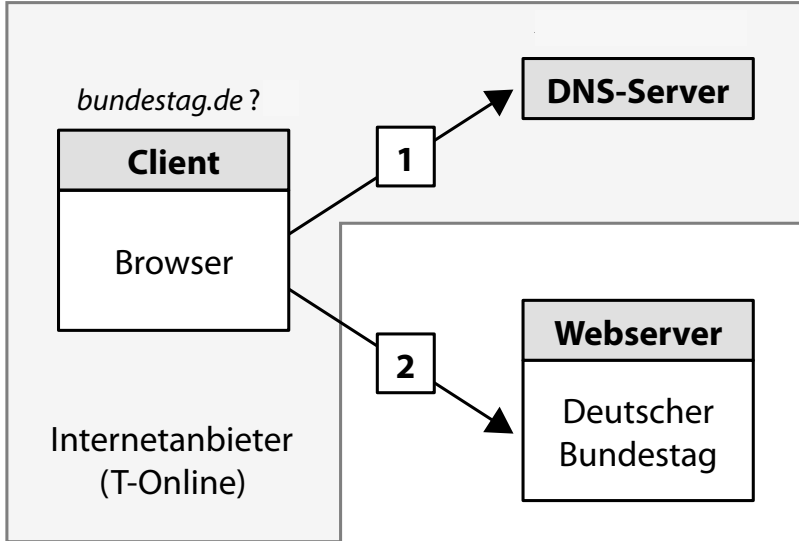
Domain Name System



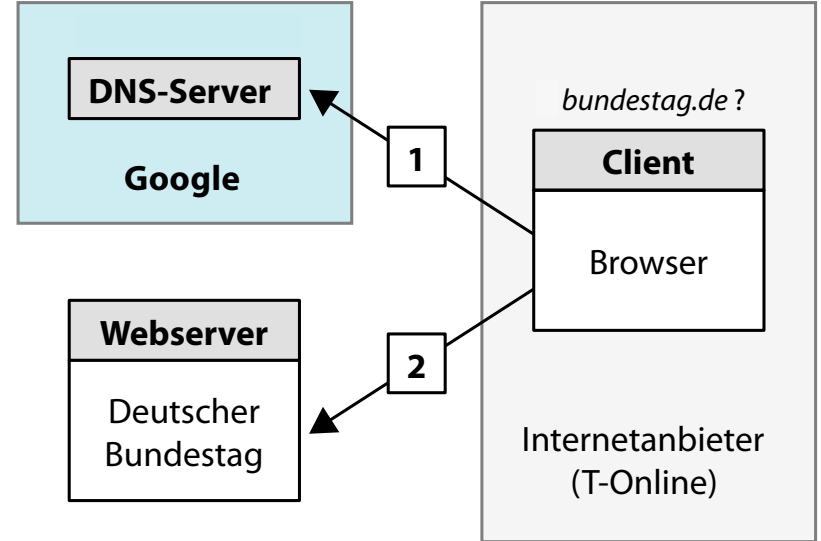


Vertraulichkeit?
Brauchen wir nicht.

```
17-Nov-2014 10:23:49.770 189.11.9.16 #15619: query: www.google.de IN A +
17-Nov-2014 10:23:51.622 42.81.144.1 #12191: query: wikipedia.org IN A +
17-Nov-2014 10:23:52.051 134.9.15.51 #13170: query: www.spiegel.de IN A +
```

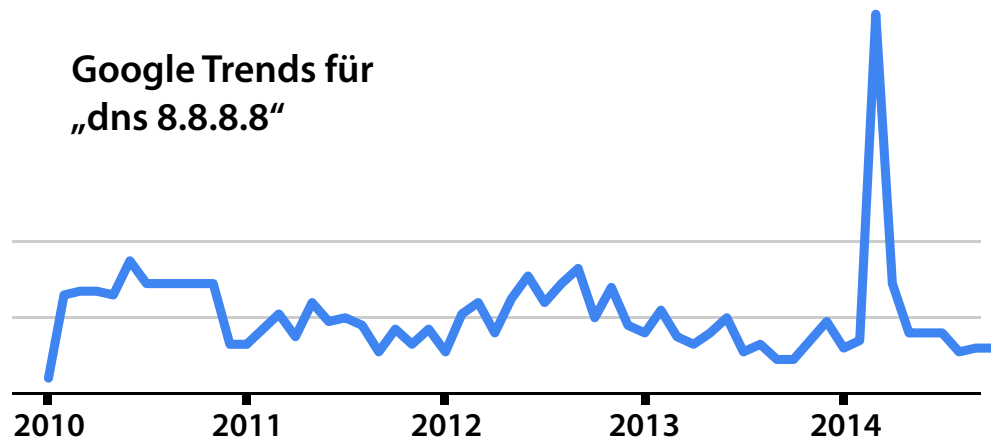


»Angreifer«

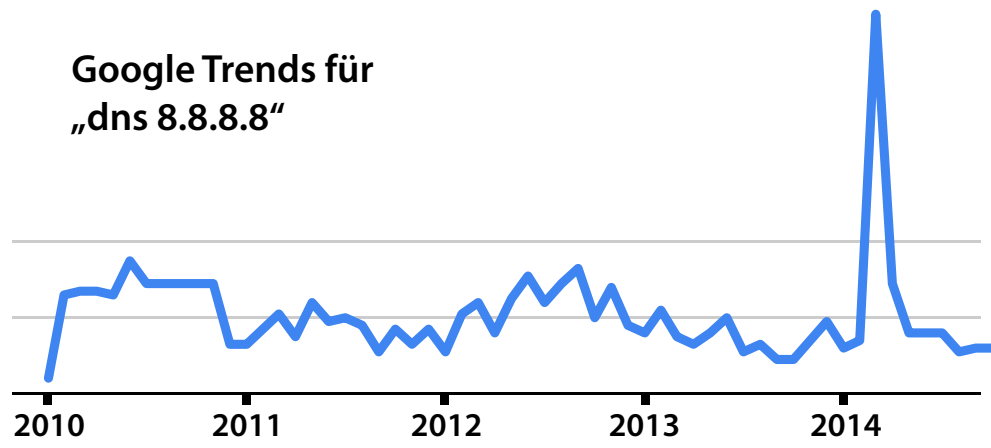


oder doch?

Google Trends für
„dns 8.8.8.8“

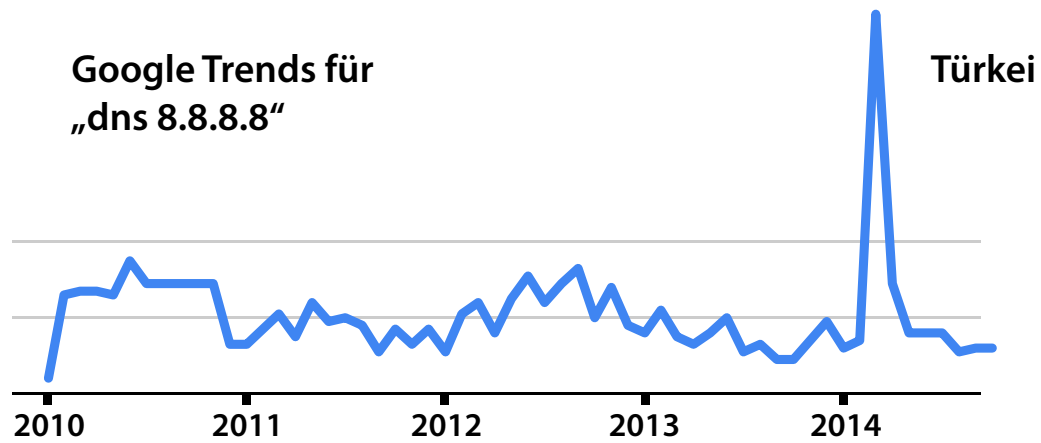
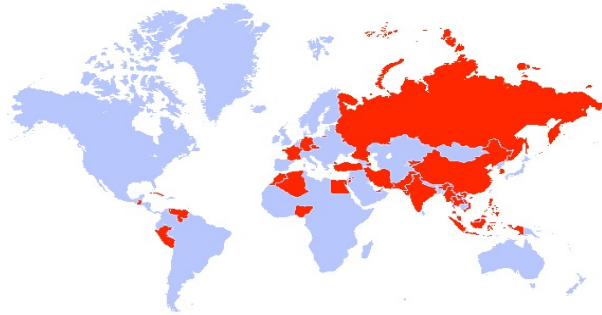


Google Trends für
„dns 8.8.8.8“

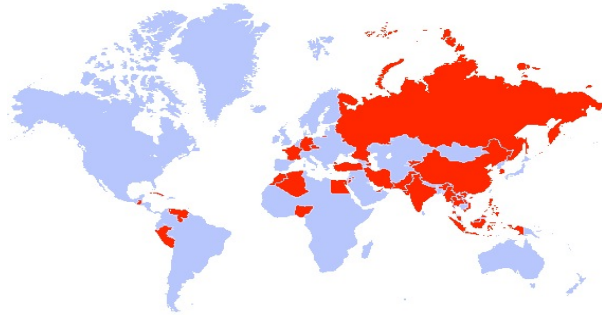


↑
**Disputation
Dr. Herrmann**

Umgehung von DNS-Sperren



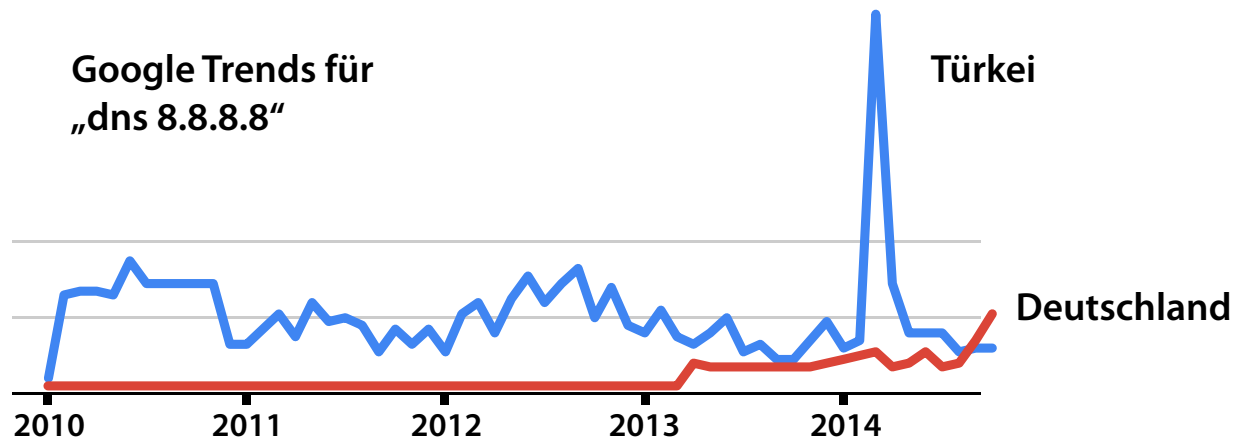
Umgehung von DNS-Sperren



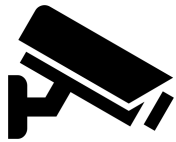
Performanz & Zuverlässigkeit

... nicht ...
... Provider angewiesen und kann einen ...
... bigen Server festlegen. Empfehlenswert sind
... die freien DNS-Server OpenDNS (208.67.
... 222.222 und 208.67.220.220) und Google
... Public DNS (8.8.8.8 und 8.8.4.4). Da die Netz-
... konfiguration Ihres Mac in der Regel m
... DHCP über den Router vorgege

Google Trends für
„dns 8.8.8.8“



Motivation



Bedrohung für die Privatsphäre?
Geeignete Schutzmechanismen?



Nutzen für die IT-Forensik?

A



Ermittlung der
besuchten Webseiten



Motivation: Profiling der Nutzer

Problem 1:

DNS-Server sieht nur Domains, jedoch keine URLs



The screenshot shows a web browser window with the address bar containing `de.wikipedia.org/wiki/Alkoholkrankheit`. The page content includes the Wikipedia logo, navigation tabs for 'Artikel', 'Diskussion', 'Lesen', 'Quelltext anzeigen', and 'Versionsgeschichte'. The main heading is 'Alkoholkrankheit'. Below the heading, there is a paragraph describing the condition and a table titled 'Vergleichende Klassifikation nach' with columns for 'ICD-10' and 'DSM-IV'. The table lists various codes and descriptions for alcohol-related conditions.

de.wikipedia.org/wiki/Alkoholkrankheit

Benutzerkonto erstellen Anmelden

Artikel Diskussion Lesen Quelltext anzeigen Versionsgeschichte Suchen

Alkoholkrankheit

Die **Alkoholkrankheit** (auch *Alkoholabhängigkeit*, *Äthylismus*, *Dipsomanie*, *Potomanie*, *Trunksucht*, *Alkoholsucht* oder *Alkoholismus* genannt) ist die **Abhängigkeit** von der **psychotropen Substanz Ethanol**.

Inhaltsverzeichnis [Anzeigen]

Symptome

Vergleichende Klassifikation nach	
ICD-10	DSM-IV
F10 Psychische und Verhaltensstörungen durch Alkohol	Störungen im Zusammenhang mit Alkohol
F10.0 akute Alkoholintoxikation (akuter Alkoholrausch)	303.00 Alkoholintoxikation
F10.1 schädlicher Gebrauch von Alkohol	305.00 Alkoholmissbrauch
F10.2 Abhängigkeitssyndrom	303.90 Alkoholabhängigkeit

Problem 2:

Domains korrespondieren nicht mit besuchten Seiten

de.wikipedia.org

bits.wikimedia.org

meta.wikimedia.org

counsellingresource.com

upload.wikimedia.org

www.izb.fraunhofer.de

www.spiegel.de

www.stadt-und-gemeinde.de

www.biospektrum.de

w210.ub.uni-tuebingen.de

www.uni-muenster.de

ec.europa.eu

www.alkoholismus-hilfe.de

www.klinik-dr-fontheim.de

www.versorgungsleitlinien.de

de.wikiquote.org

drogenbeauftragte.de

www.sucht-info.ch

www.aafp.org

www.thieme-connect.com

www.kenn-dein-limit.de

www.casusconsult.nl

www.hta.ac.uk

www.stern.de

whqlibdoc.who.int

www.focus.de

Abrufmuster für
<http://de.wikipedia.org/wiki/Alkoholkrankheit>

de.wikipedia.org

bits.wikimedia.org

meta.wikimedia.org

counsellingresource.com

upload.wikimedia.org

www.izb.fraunhofer.de

www.spiegel.de

www.stadt-und-gemeinde.de

www.klinik-dr-fontheim.de

www.versorgungsleitlinien.de

de.wikiquote.org

drogenbeauftragte.de

www.sucht-info.ch

www.aafp.org

www.thieme-connect.com

www.kenn-dein-limit.de



Kann ein DNS-Server (nicht-)besuchte Webseiten anhand charakteristischer Abrufmuster erkennen?

Empirische Untersuchung

1. Abruf von Webseiten
2. Aufzeichnen der DNS-Abrufmusters
3. Bestimmung der k -Identifizierbarkeit



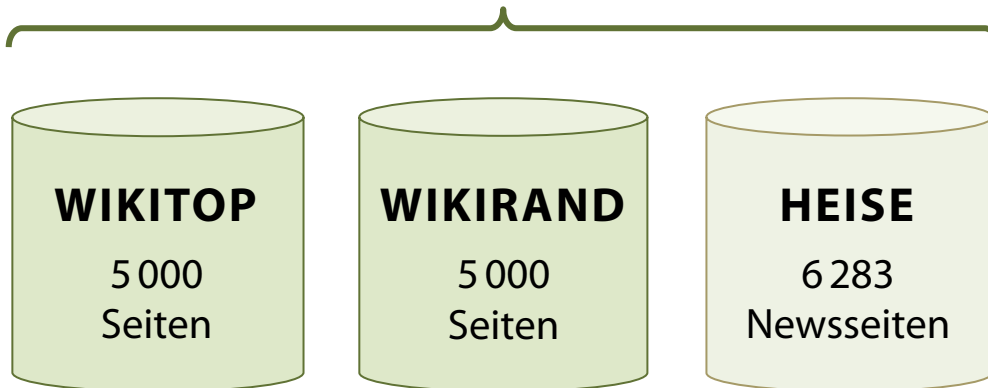
$k = 1$ > 99 %

80 % ← Seiten mit einzigartigem Abrufmuster

$k \leq 5$ > 99 %

94 % ← Seiten mit Abrufmustern, die jeweils höchstens auf 5 Seiten auftreten

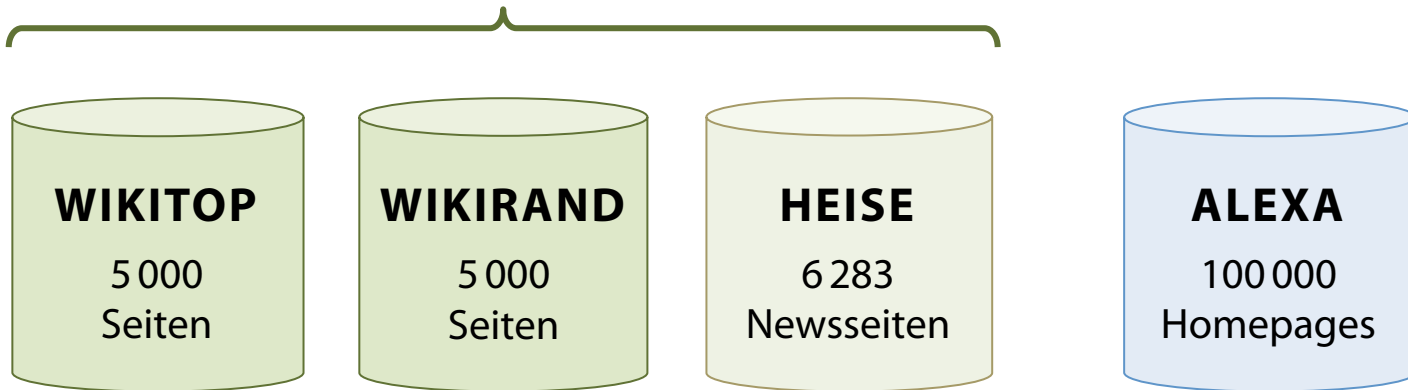
Ermittlung der genauen URL



$k = 1$	> 99 %	80 %	63 %
$k \leq 5$	> 99 %	94 %	76 %

Empirische Untersuchung

Ermittlung der genauen URL



$k = 1$	> 99 %	80 %	63 %	> 99,9 %
$k \leq 5$	> 99 %	94 %	76 %	> 99,9 %

A



Ermittlung der
besuchten Webseiten

genauer als gedacht

B



Identifizierung der
benutzten Software



Motivation: Nachweis der Beteiligung
an kriminellen Handlungen

Ergebnis: gängige Betriebssysteme und
Browser allein anhand ihres Verhaltens
identifizierbar

A



Ermittlung der
besuchten Webseiten

B



Identifizierung der
benutzten Software

C



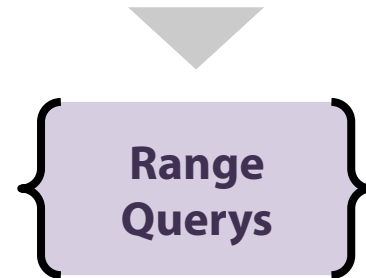
Techniken zum
Selbstdatenschutz



Schutz der Identität
des Nutzers



Verbergen der
wahren Interessen



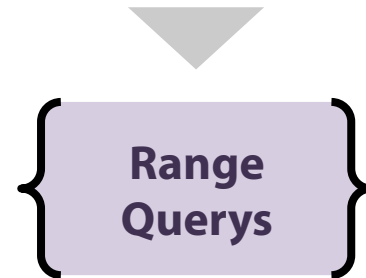
Existierende Ansätze für DNS
ungeeignet oder unsicher



Schutz der Identität
des Nutzers



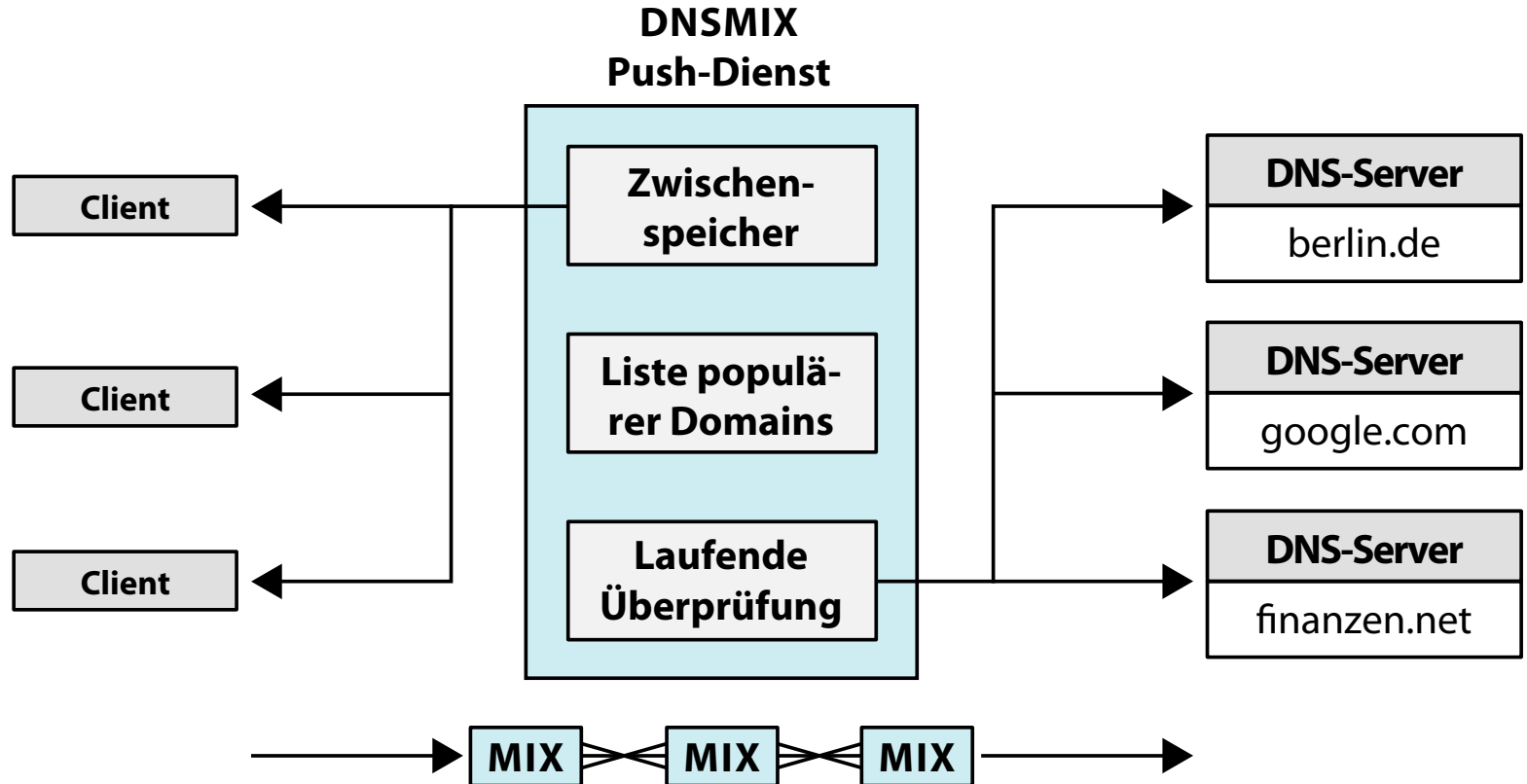
Verbergen der
wahren Interessen



Mit **DNSMIX** lässt sich
Unbeobachtbarkeit erreichen!

Idee von DNSMIX

populäre DNS-Einträge automatisch an alle Nutzer senden



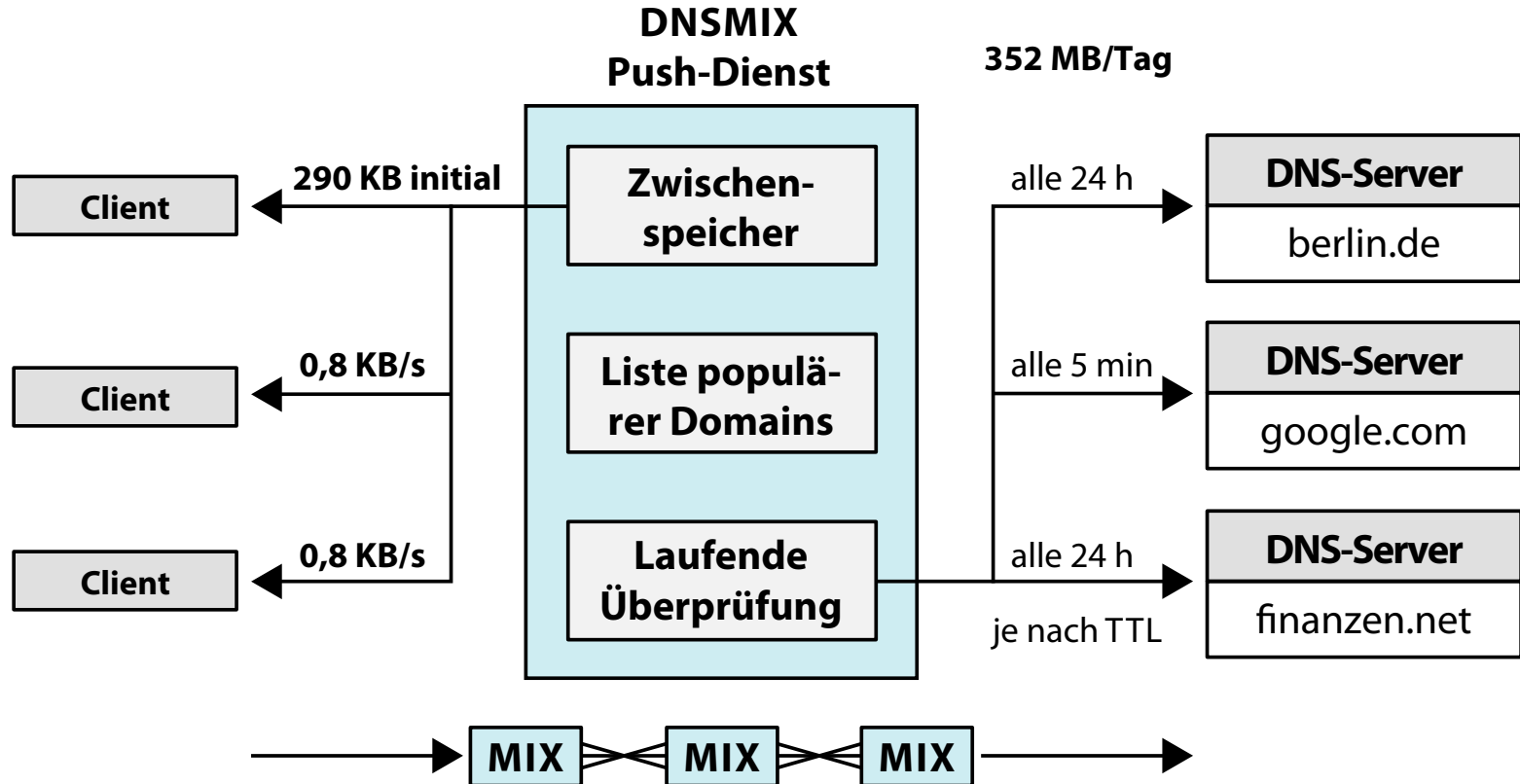
Empirische Untersuchung

Anfragen
von 2082
Nutzern

Pushen von 10.000
populären Domains



Auflösung von 84 % der Anfragen
unbeobachtbar und **unmittelbar**



A



Ermittlung der
besuchten Webseiten

B



Identifizierung der
benutzten Software

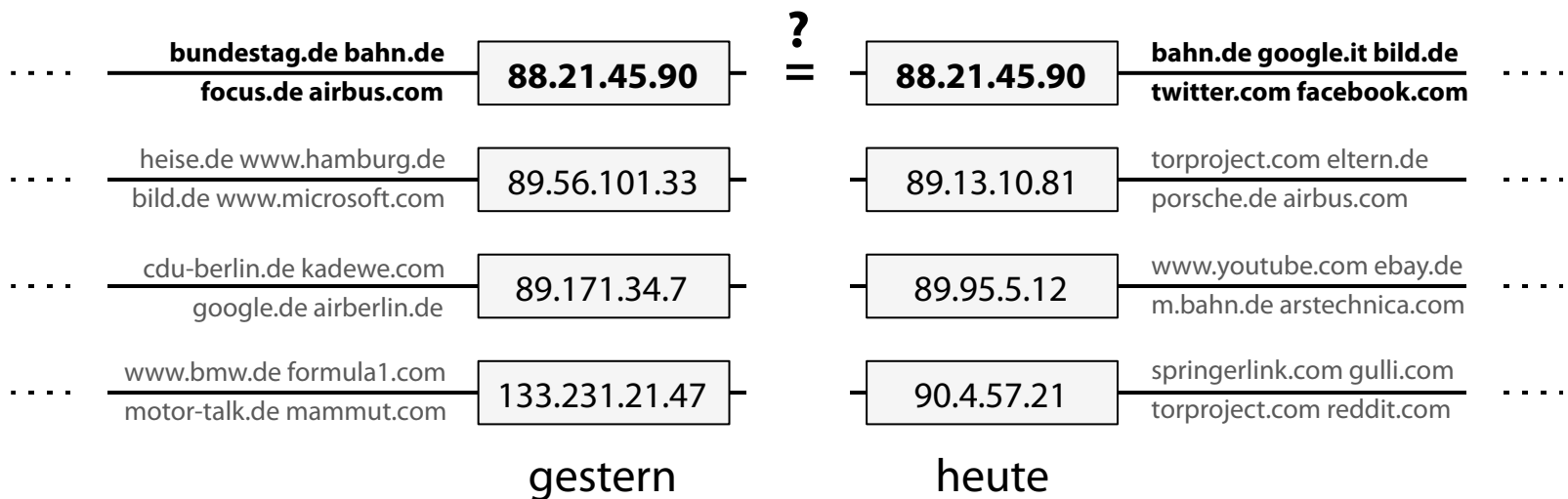
C



Techniken zum
Selbstdatenschutz

Bislang ignoriertes Problem

Beobachtung von Nutzern mit dynamischen IP-Adressen

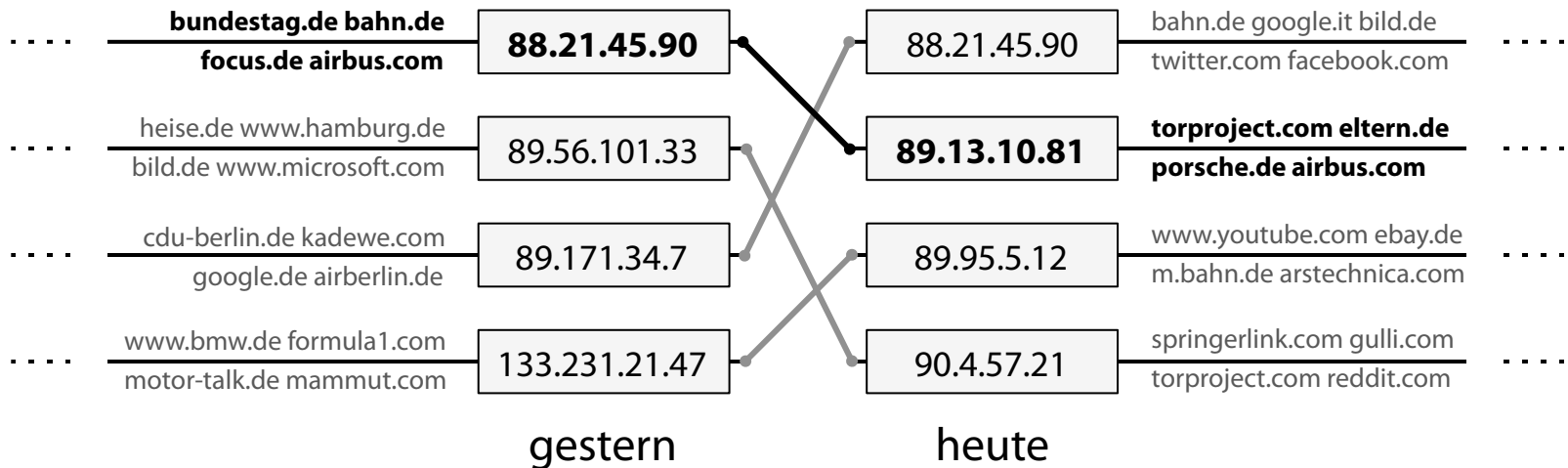


D



Verhaltensbasierte Verkettung von Sitzungen

erlaubt Tracking von Nutzern mit dynamischen IP-Adressen



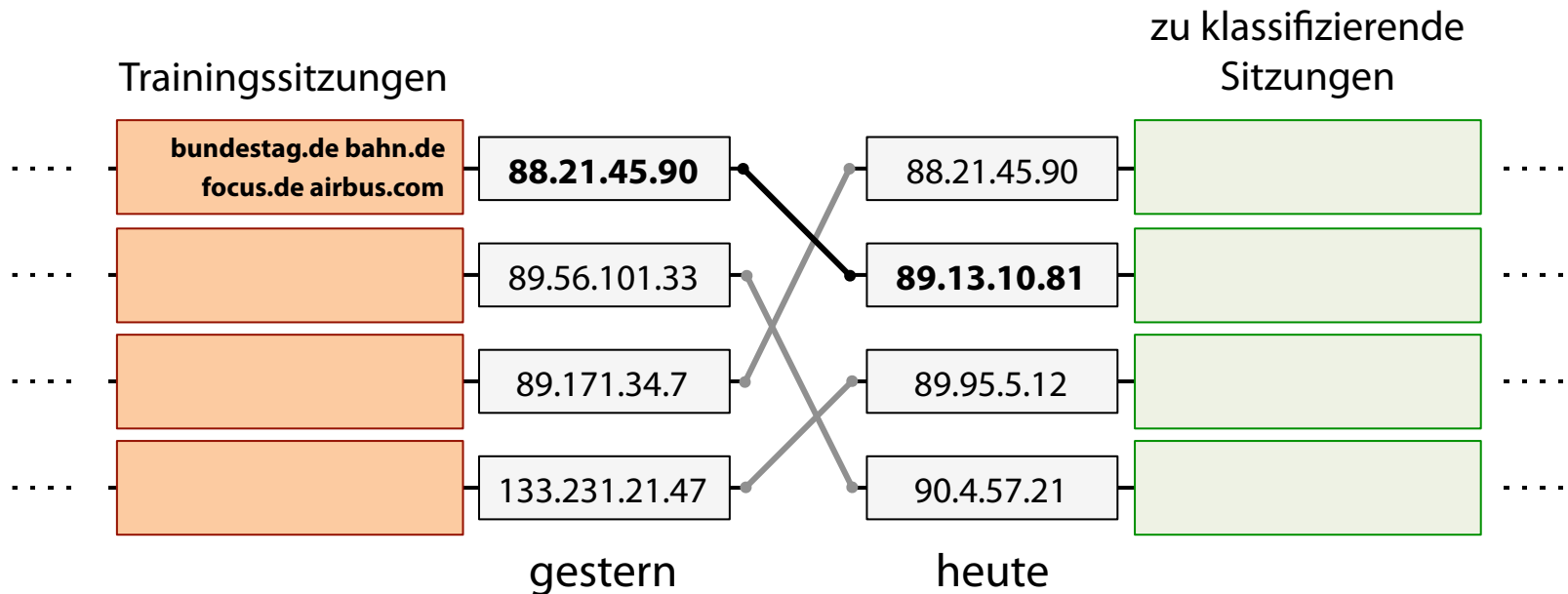
Ausnutzung menschlicher Eigenschaften

individuelle Vorlieben

tägliche Routine



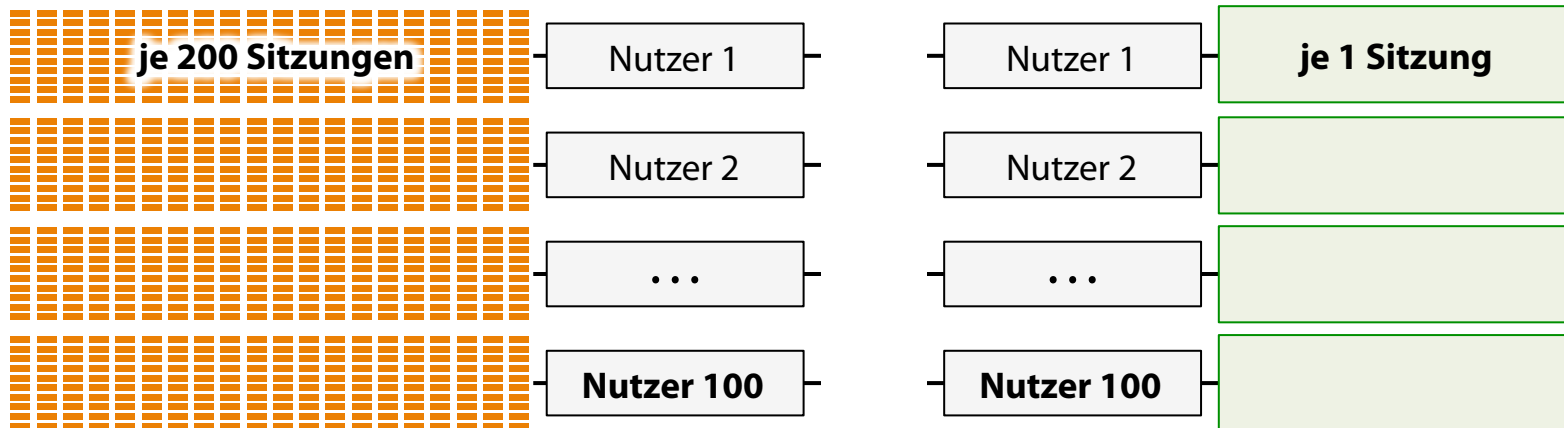
Trainieren eines Klassifikators zur Wiedererkennung



Verhaltensbasierte Verkettung
einzelner Sitzungen im Open-World-Szenario
noch nicht erforscht

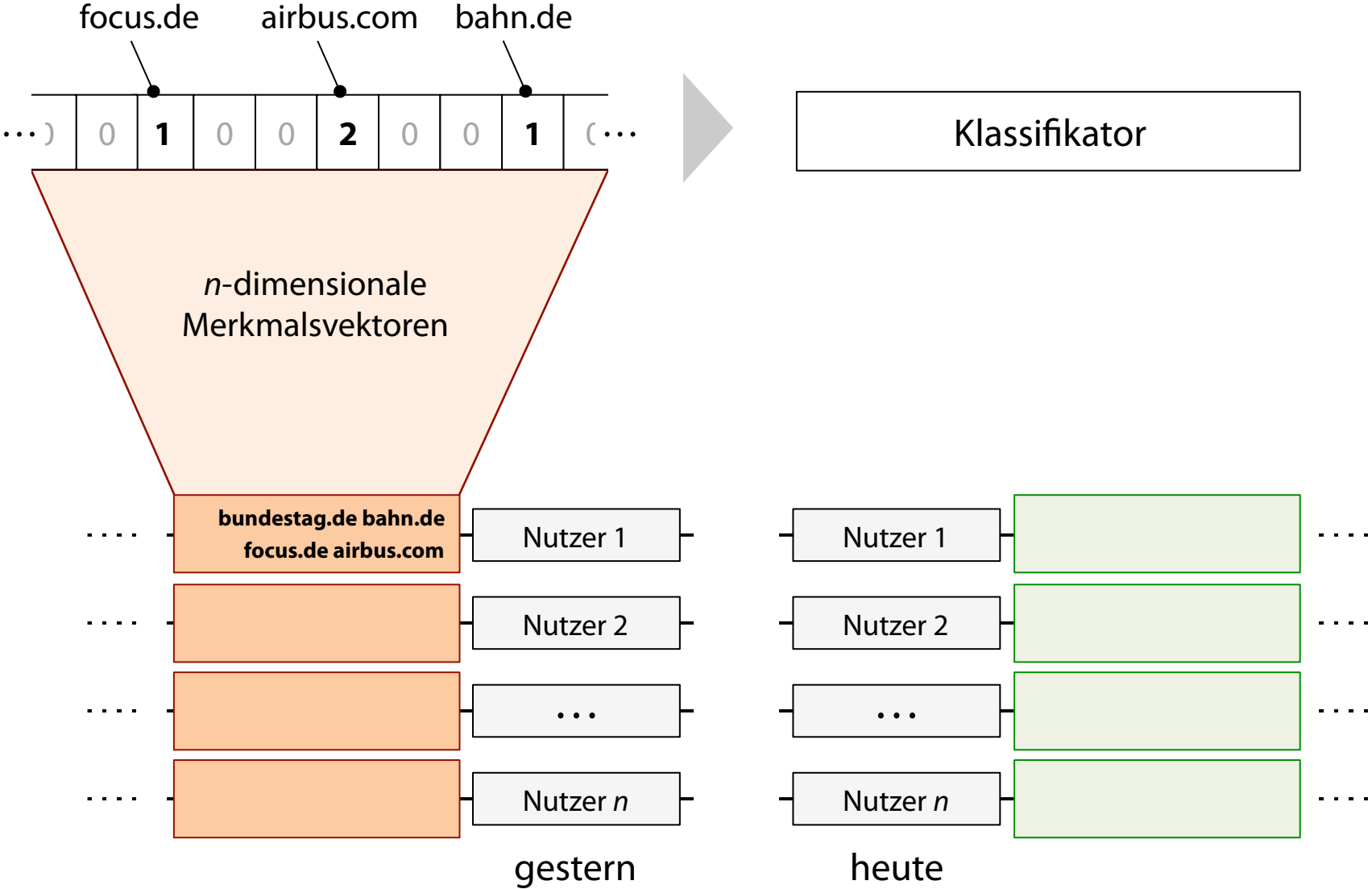
Reicht **eine** Trainings Sitzung?
Genauigkeit bei großen **fluktuierenden** Nutzergruppen?

Closed-World-Untersuchung von Yang (2010)

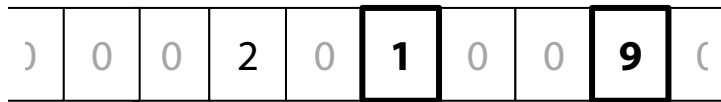
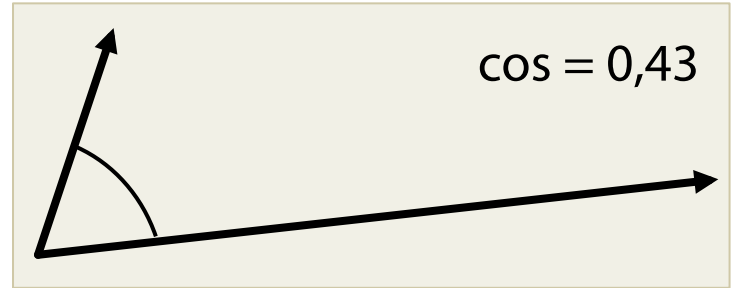
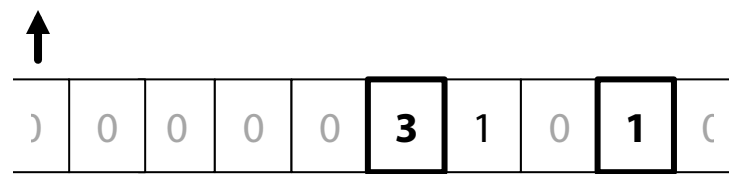
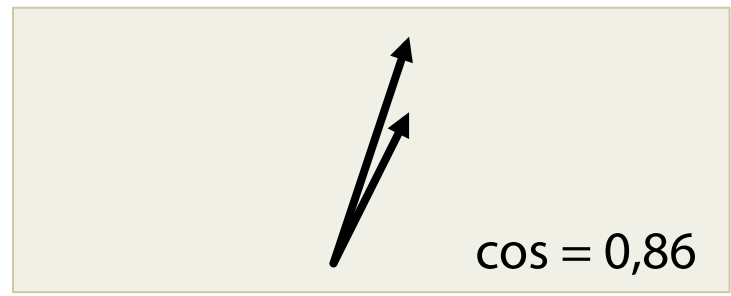
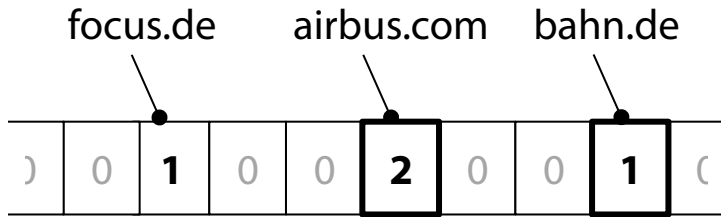


62% Genauigkeit

Konstruktion des Verkettungsverfahrens



Funktionsweise des verwendeten Klassifikators

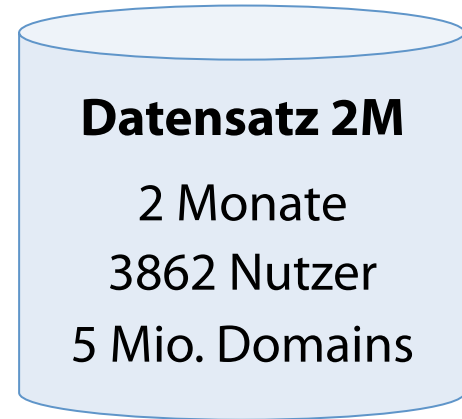


gestern

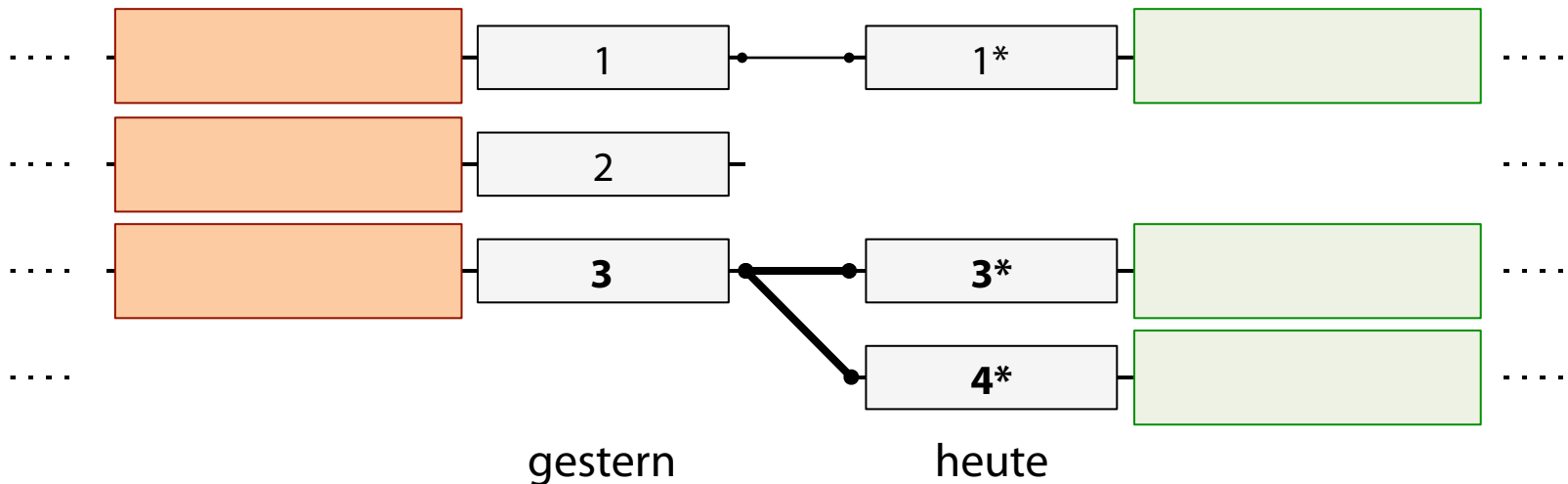
heute

Empirische Untersuchung

1. Implementierung mit MapReduce
2. Erhebung eines DNS-Datensatzes
3. Experimente im Open-World-Szenario



Beobachtung: Fehlentscheidungen wegen Nutzerfluktuation



Empirische Untersuchung

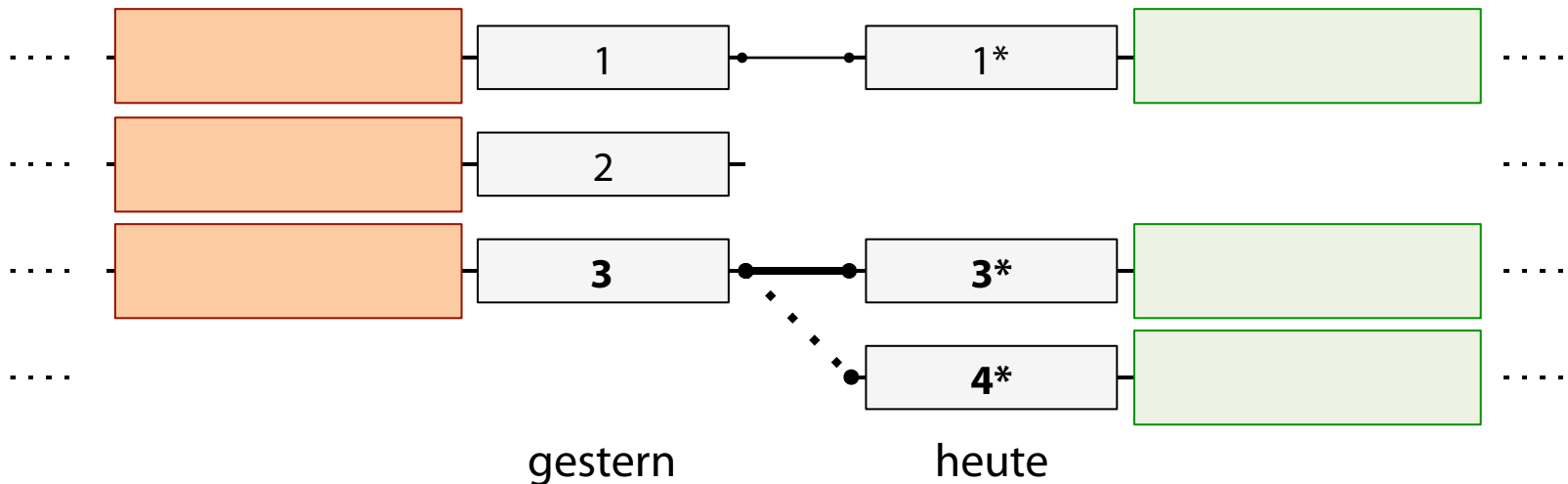
Standardverfahren

75 % Genauigkeit

nach Optimierung

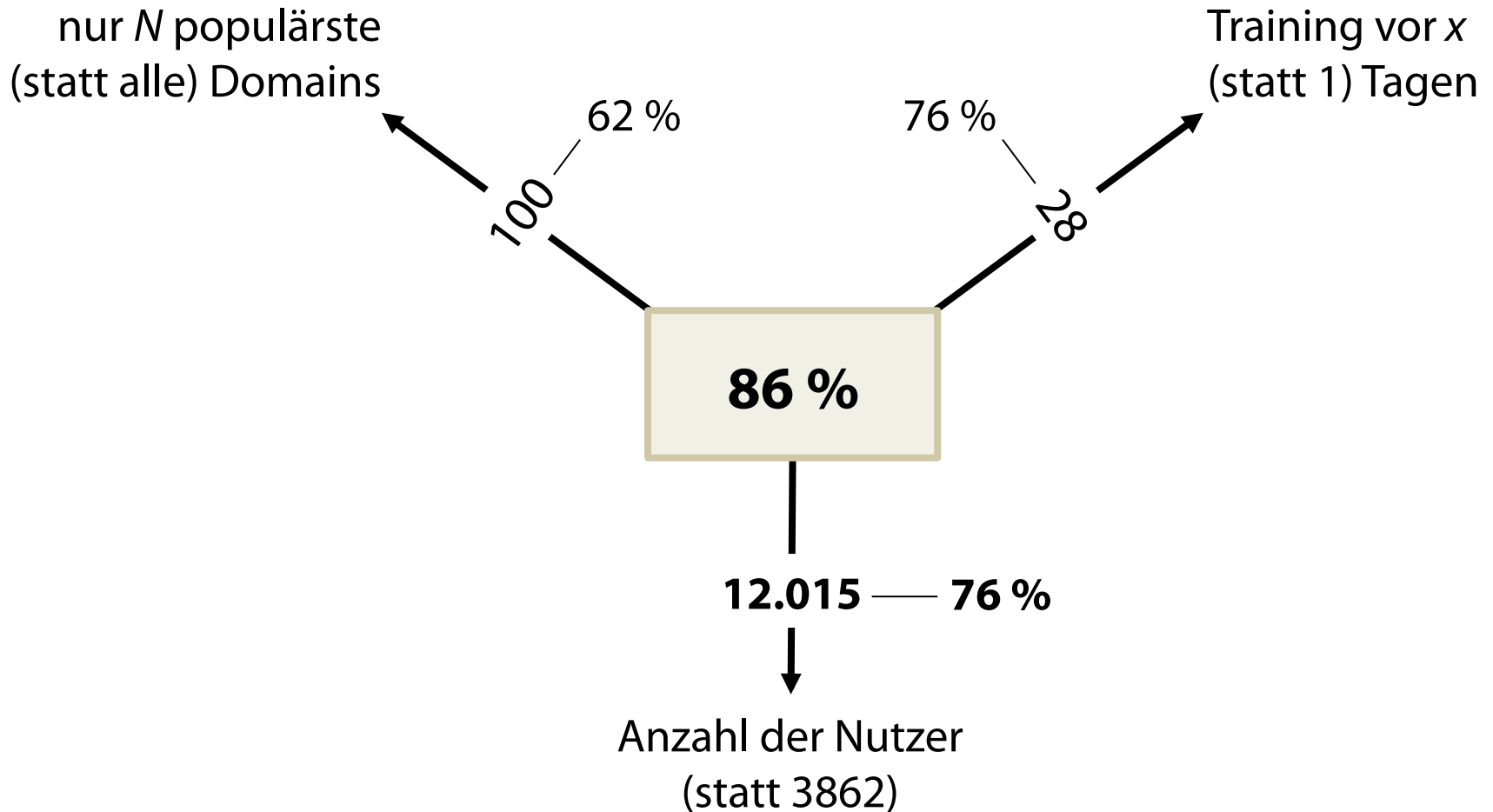
86 % Genauigkeit

Optimierung: Ermittlung der ähnlichsten Sitzung



Verkettung erweist sich als robust

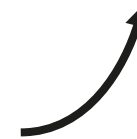
gelingt auch unter widrigen Umständen



Einsatz auch unabhängig vom DNS möglich

HTTP- und Flash-Cookies

Browser-Fingerprinting



Tracking- und Profiling-Dienste



Verhaltensbasierte Verkettung
erfolgt rein passiv, ist also nicht erkennbar

D



Verhaltensbasierte
Verkettung von Sitzungen

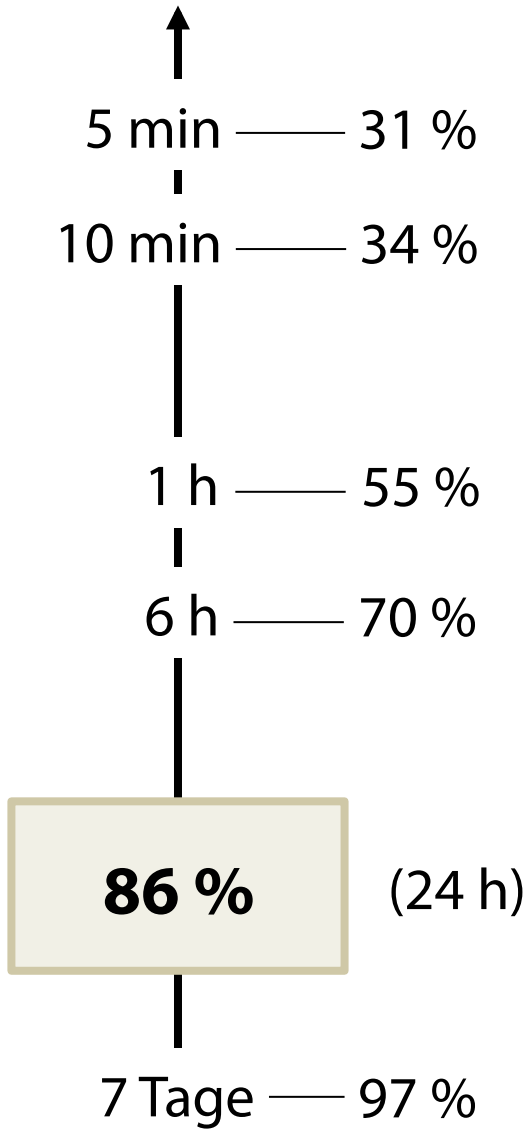
E



Schutz vor Tracking durch
verhaltensbasierte Verkettung

ist nicht nur für Nutzer öffentlicher
DNS-Server von Interesse

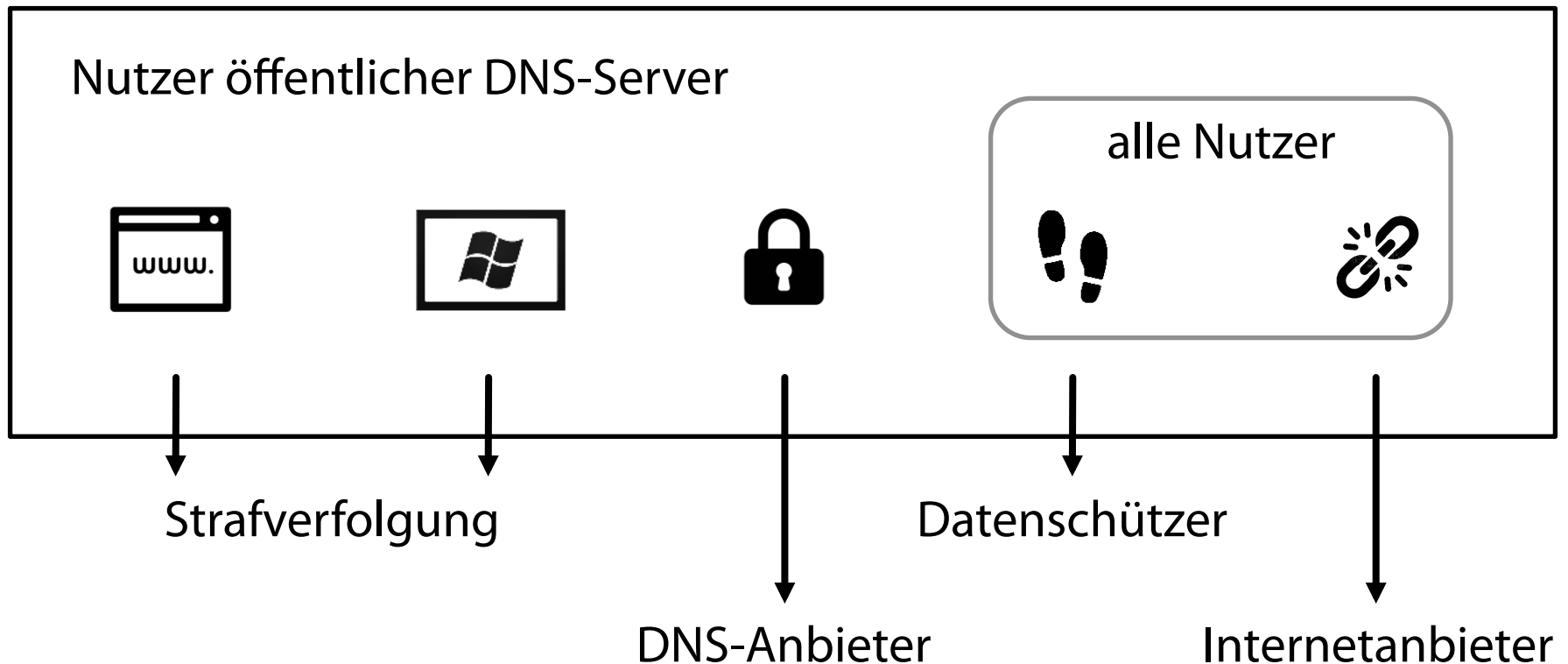
Sitzungsdauer



IP-Adresse häufig wechseln
heute: Komforteinbußen

Chance
»Privacy by Default«
mit IPv6

Zusammenfassung und Relevanz der Ergebnisse



Beobachtungsmöglichkeiten im Domain Name System

... **umfangreicher und genauer** als zu erwarten wäre;

DNSMIX: **dienstspezifische Konzepte** lohnenswert;

unbemerkt Tracking – auch abseits von DNS – möglich;

komfortabler Schutz durch **häufigen Adresswechsel** mit IPv6.



Generalisierbarkeit der Ergebnisse?

Konkrete Handlungsempfehlungen?

Weitere Experimente und Future Work?

