Zentrum für sichere Informationstechnologie – Austria
Secure Information Technology Center – Austria

A-1030 Wien, Seidlgasse 22 / 9
Tel.: (+43 1) 503 19 63–0
Fax: (+43 1) 503 19 63–66

A-8010 Graz, Inffeldgasse 16a
Tel.: (+43 316) 873-5514
Fax: (+43 316) 873-5520

http://www.a-sit.at
E-Mail: office@a-sit.at

DVR: 1035461 ZVR: 948166612 UID: ATU60778947

SPURENVERWISCHUNG IM INTERNET
STUDIE

Alexander Marsalek - alexander.marsalek@a-sit.at
Bernd Prünster - bernd.pruenster@a-sit.at
Dominik Ziegler - dominik.ziegler@a-sit.at

Zusammenfassung: Bei nahezu allen Internet-Aktivitäten werden Spuren hinterlassen, welche entsprechend zusammengeführt und ausgewertet, Rückschlüsse zulassen, die einzelne Personen mit hinreichender Sicherheit identifizieren können. Einer der Hauptgründe dafür ist, dass der technische Unterbau des Internets aus einer Zeit stammt, in der Privatsphäre und Datenschutz auf dieser Ebene irrelevant waren. Da es sich bei den involvierten Komponenten jedoch vielfach um kritische Infrastrukturen handelt, können diese nicht ersetzt werden. Somit müssen aktiv zusätzliche Maßnahmen ergriffen werden, um die eigene Identität im Internet geheim zu halten und möglichst wenig Spuren zu hinterlassen. Doch vor allem werden auch auf Applikationsebene unzählige Spuren hinterlassen, welche unzweifelhaft einzelnen Personen zugeordnet werden können.
Im Rahmen diese Studie wird zuerst dargelegt, wo welche Spuren bei alltäglichen Internet-Aktivitäten hinterlassen werden und gezeigt, wie diese sich auf einzelne Nutzer und Nutzerinnen zurückführen lassen. Anschließend werden unterschiedliche Methoden und Konzepte zur Vermeidung verschiedenster Arten von Spuren im Internet beschrieben. Aufbauend auf diesen allgemeinen Möglichkeiten zur Spurenverwischung werden nachfolgend konkrete Implementierungen der diskutierten Konzepte vorgestellt. Generell wurden dabei Implementierungen dahingehend ausgewählt, ein möglichst breites Spektrum an Spurenverwischungs-methoden abzudecken. Abschließend werden die gewonnenen Erkenntnisse zueinander in Bezug gesetzt und zusammengefasst.

Dokumenthistorie

Version	Datum	Autor	Kommentar
0.1	01.08.2016	Dominik Ziegler	Struktur
0.2	26.09.2016	Alexander Marsalek	Fingerprinting, Ghostery, NoScript, Tor-Browser, Tails
0.3	10.10.2016	Dominik Ziegler	Surface Web, Deep Web, Darknet, Einleitung, Charakteristik
0.4	20.10.2016	Bernd Prünster	Tor, I2P, SSH-Tunnel/VPN, Methoden
0.5	11.11.2016	Bernd Prünster, Dominik Ziegler	Lokale Eigenschaften, Mix-Netze, Proxies
0.6	14.11.2016	Alexander Marsalek	Bezahlmethoden, Überarbeitung Ghostery, Netzwerkeigenschaften
0.7	12.12.2016	Alexander Marsalek, Bernd Prünster, Dominik Ziegler	Erster Draft
1.0	13.12.2016	Herbert Leitold	Review

Inhaltsverzeichnis

1 Einleitung

Bei der Benutzung des Internets hinterlässt man unweigerlich Spuren. Je nach Art der Anwendung können die dabei hinterlassenen Spuren, wie Konversationsprotokolle oder Kopien von Dateien, variieren. Beispielsweise legen Browser typischerweise eine Historie von besuchten Webseiten, Cookies, Session-Informationen oder Offline-Kopien an, um die Benutzerfreundlichkeit zu steigern oder bestimmte Funktionalität zur Verfügung zu stellen.
Gleichzeitig hinterlässt das Surfen im Internet aber auch Spuren abseits des eigenen Computers. Basierend auf dem Betriebssystem, dem benutzten Browser, zuletzt besuchter Seiten oder der IP-Adresse können Webseiten aber auch Dritte, Profile anlegen. Studien haben gezeigt, dass eine eindeutige Identifizierung eines Nutzers bzw. einer Nutzerin basierend auf diesen Eigenschaften möglich ist [1].
Mit ein Grund dafür sind die zugrundeliegenden Technologien und Protokolle, die nicht auf Anonymität oder den Schutz der Privatsphäre ausgelegt sind. Bei jeder Nachricht sind Sender und Empfänger ersichtlich. Überwachung oder Kontrolle durch Dritte, zum Beispiel durch Beschränkung des Internetverkehrs oder durch Beobachtung des Datenflusses auf einem spezifischen Webserver, ist dadurch technisch möglich.
Je nach Anwendungsfall gibt es somit verschiedene Gründe (z.B. persönlich, politisch, religiös, rechtlich), um das Internet anonym zu nutzen oder Aktivitäten im Internet zu verschleiern. Es können dies legitime Gründe sein, aber auch illegitime oder kriminelle. Die Technologien sind oft dieselben, es ist auch nicht technisch unterscheidbar, ob die Motivation zur Anonymität rechtens ist, oder sie dem Verschleiern unrechtmäßiger Vorgänge dient.
Zur Anonymität gilt, dass es im Prinzip zwei Arten von Anonymität gibt: Soziale Anonymität und technische Anonymität [2]. Während soziale Anonymität, die Anonymität gegenüber Anderen aufgrund fehlender Attribute oder Wissen definiert, man also nicht wirklich anonym ist, beschreibt die technische Anonymität, das Fehlen relevanter Informationen, die die Identifikation des Benutzers bzw. der Benutzerin ermöglichen.
Aus dem Streben nach technischer Anonymität im Internet, und um sicher zu gehen, dass bestimmte Unterhaltungen nicht überwacht oder blockiert werden können, sind verschiedene Technologien und Anwendungen entstanden, um Spuren im Internet zu verschleiern.
Die vorliegende Studie erhebt aus diesem Grund die aktuelle Situation und gibt einen Überblick über den aktuellen Stand der Technik sowie mögliche Anwendungen um Spuren im Internet zu verwischen. Sie gliedert sich in fünf Abschnitte. Abschnitt 2 behandelt die Grundlagen für das Surfen im Internet sowie verwendeter Technologien. Weiters werden, für einen besseren Überblick, verschiedene Bereiche des Internets definiert. Im Anschluss gibt Abschnitt 3 einen Überblick über Methoden zur Spurenverwischung sowie Möglichkeiten Spuren zu vermeiden. In Kapitel 4 werden Beispiele für Implementierungen genannt. Dies beinhaltet sowohl Plugins für bekannte Webbrowser als auch eigenständige Applikationen. Die Studie schließt mit einem Fazit über die diskutierten Methoden und gibt allgemeine Empfehlungen für das Surfen im Internet.

2 Grundlagen

Um ein einheitliches Grundverständnis über die in diesem Dokument diskutierten Technologien sowie deren Implikationen zu gewährleisten, werden im Folgenden die verwendeten Begriffe und Konzepte definiert. Anhand eines Beispiels werden zu Beginn die Elemente, die beim Aufruf einer Webseite involviert sind, identifiziert und genauer erklärt. Anschließend wird ein Überblick über Möglichkeiten zur Identifizierung von Benutzern und Benutzerinnen, über Spuren die lokal aber auch im Internet hinterlassen werden, gegeben. Im Zuge dessen werden die unterschiedlichen Bereiche des Internets, sowie deren Strukturen definiert. Im Anschluss werden (anonyme) Bezahlmethoden im Internet behandelt.

2.1 Was beim Aufruf einer Webseite passiert

Um die typischen Schritte, die beim Aufruf einer Webseite durchlaufen werden, zu veranschaulichen, illustriert Abbildung 1 anhand eines Beispiels diesen Prozess. In diesem Fall wird die Startseite des Webauftrittes der Österreichischen Nationalbank herangezogen, also einer Organisation, wo außer Zweifel steht, dass die Webseite rein der Information und nicht dem Abgreifen von persönlichen Daten dient. Ziel ist zu zeigen, welche Information bei praktisch jedem Web-Aufruf übertragen wird.

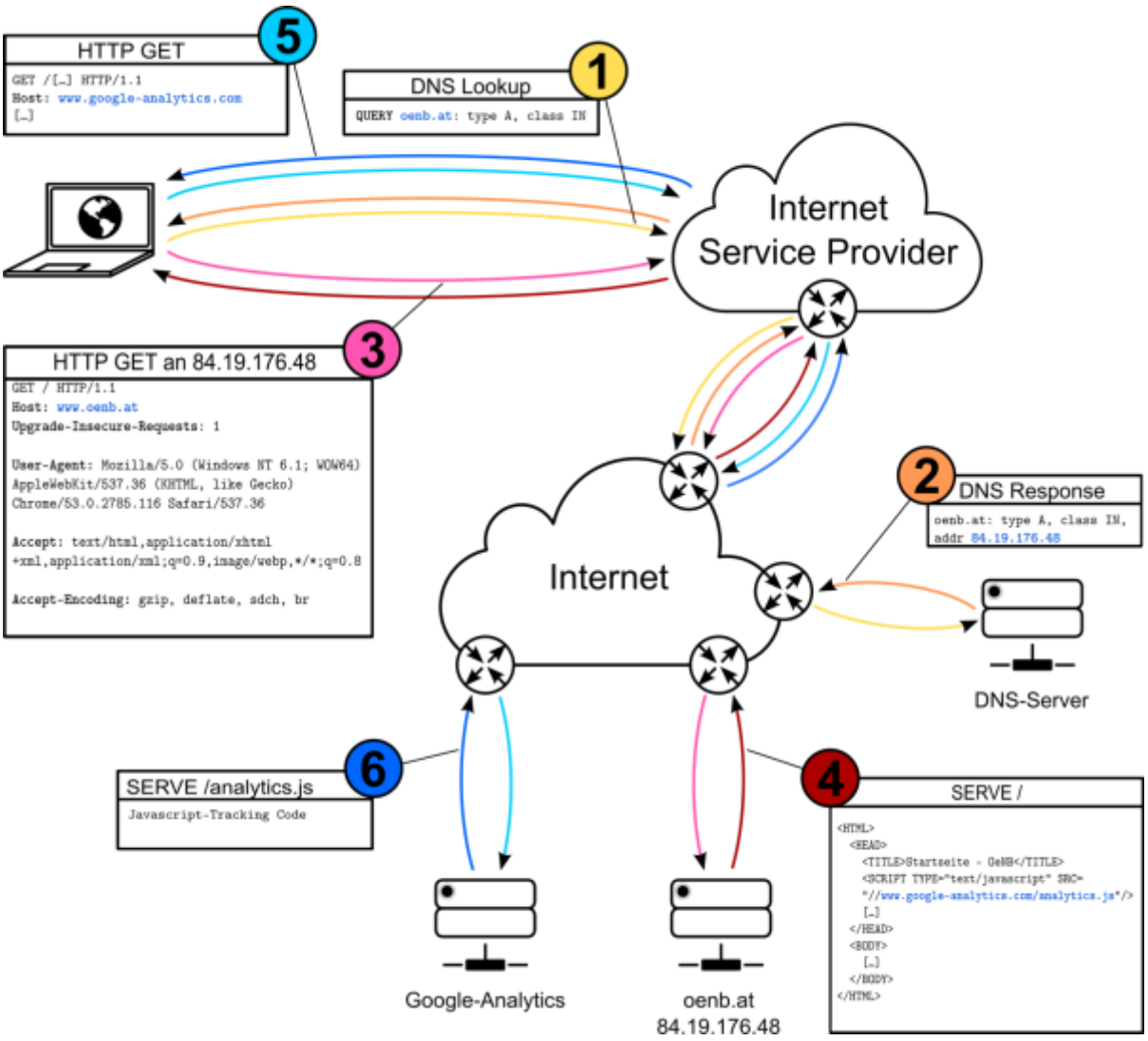

Abbildung 1: Aufruf einer Webseite (www.oenb.at)

Beim Aufruf von https://www.oenb.at/ wird zuerst eine Anfrage an einen Domain Name System (DNS) Server abgesetzt (1). Dieser liefert die zur Domain gehörende IP Adresse 88.19.176.48 zurück (2). Anschließend kontaktiert der Browser die aufgelöste IP-Adresse um die gewünschte Seite zu erhalten (3). In dieser Anfrage sendet der Browser einige Metadaten wie die Art des Gerätes, das verwendete Betriebssystem, den verwendeten Browser, sowie weitere Informationen mit, damit der Webserver eine möglichst gut darstellbare Version der Webseite ausliefern kann. Im Anschluss wertet der Webserver diese Informationen aus und liefert die angefragte Webseite aus (4). Empfängt der Browser die Seite, wird sie interpretiert und dargestellt. Sind externe Dateien wie Bilder oder Skripte eingebunden, werden diese ebenfalls geladen. Beispielsweise bindet die OeNB-Webseite ein Skript von Google-Analytics zur Analyse des Nutzungsverhalten ein[1]. Dieses Skript wird dementsprechend vom Browser abgerufen (5). Analog zum vorherigen Ablauf werden bei dieser Anfrage wieder Metadaten übertragen, sowie im übertragenen Header ein *Referrer-Feld* hinzugefügt, welches die URI der ursprünglichen Seite beinhaltet. Bei jedem dieser Schritte werden zusätzliche Spuren, die Informationen über den Benutzer und die Benutzerin preisgeben können, hinterlassen. Beispielsweise wissen der Internet Service Provider (ISP) und alle beteiligten Knoten zwischen Sender und Empfänger welche Seiten aufgerufen werden. Zusätzlich kann diese Information vom DNS-Server und auch von eventuell eingebundenen Tracking-Seiten wie Google-Analytics genutzt werden. Es werden also sowohl lokal, im Netzwerk und auch bei der Gegenstelle Spuren, beispielsweise in Form von Cookies, der Browser-History oder der IP-Adresse hinterlassen.

[1] Die IP-Adresse von google-analytics.com muss ebenfalls aufgelöst werden. Die hierfür notwendigen Schritte sind aus Gründen der Übersichtlichkeit nicht dargestellt.

Die nächsten Abschnitte beschreiben diese hinterlassenen Spuren bzw. Quellen, die zur Identifizierung eines Browsers oder Geräts benutzt werden können, im Detail.

2.2 Wie Benutzer und Benutzerinnen identifiziert werden können

Die Identifizierung von Benutzern und Benutzerinnen im Internet, ist ein gängiges bzw. notwendiges Mittel um gewisse Funktionalität, welche die Benutzerfreundlichkeit steigert, zur Verfügung zu stellen. Dazu zählt beispielsweise das Speichern von Session-Informationen, um eine erneute Eingabe von Passwörtern zu vermeiden oder um sicher zu stellen, dass gewisse Dienste genutzt werden dürfen. Für diesen Vorgang werden Daten, wie zum Beispiel Cookies, von Webseiten am Ziel-PC abgelegt, oder übermittelte Eigenschaften, wie Browser-Informationen, ausgewertet. Die Verwendung von persistenten Daten oder benutzerspezifischen Attributen, um Benutzer und Benutzerinnen zu identifizieren, beeinträchtigt also nicht grundsätzlich die Privatsphäre und wird von vielen Services für einen reibungslosen Ablauf benötigt.
Allerdings werden diese Informationen häufig auch verwendet, um Aktivitäten von Benutzern und Benutzerinnen für Marketingzwecke oder Ähnliches zu verfolgen. Je nach eingesetzter Technologie oder angestrebter Persistenz, gibt es dafür unterschiedliche Implementierungen. An den richtigen Stellen eingesetzt, ist es somit möglich, die Aktionen von Benutzern und Benutzerinnen eindeutig nachzuvollziehen, sowie Benutzerprofile über deren Angewohnheiten anzulegen.
Aufgrund dieser Datenschutzprobleme implementieren viele Browserhersteller Mechanismen, die es erlauben Technologien, die Benutzer und Benutzerinnen identifizieren, automatisch zu blockieren (sog. *Privater Modus*, *Private Browsing*, oder *Inkognito-Modus*), um eine Nachverfolgung zu verhindern oder zumindest zu erschweren. Gleichzeitig existieren verschiedene Ansätze die, trotz vorhandener Sicherheitsmechanismen, Benutzer und Benutzerinnen eindeutig identifizieren können. Diese Technologien werden im Folgenden behandelt.

2.2.1 Über lokal hinterlassene Spuren

Das Surfen im Internet kann verschiedene Spuren auf einem Computer hinterlassen. Einerseits speichern viele Internet-Browser automatisch eine Historie besuchter Seiten und eingegebener Formulardaten, andererseits werden von Webseiten unterschiedlichste Informationen, wie beispielsweise Cookies, am Gerät gespeichert um gewünschte Funktionalität zur Verfügung zu stellen. Dies ermöglicht es Seitenbetreibern, Einstellungen über mehrere Sitzungen aufrecht zu erhalten bzw. diese für einen Benutzer und Benutzerin abzulegen. Gleichzeitig ist es dadurch auch möglich User-Aktionen zu rekonstruieren oder je nach eingesetzter Technologie Benutzer und Benutzerinnen und deren Aktivitäten zu verfolgen.
Um Sessioninformationen abzuspeichern, gibt es typischerweise mehrere Ansätze die, abhängig von der gewünschten Persistenz oder verwendeten Technologie, alleine oder in Kombination mit anderen Mechanismen zum Einsatz kommen. Wird keine dauerhafte Persistenz vorausgesetzt, werden herkömmlicherweise HTTP-Cookies verwendet, die relevante Informationen beinhalten. Diese können allerdings einfach gelöscht, oder durch die Verwendung von Browsern im Inkognito-Modus automatisch blockiert bzw. gelöscht werden. Für eine dauerhafte Identifikation von Benutzern und Benutzerinnen werden deswegen, unter Zuhilfenahme von verschiedenen Technologien, meist mehrere Kopien von Cookies an verschiedenen Stellen am Computer abgelegt. Werden HTTP-Cookies oder Teile von Datenbanken gelöscht, ist es somit in den meisten Fällen ausreichend, dass eine dieser Kopien verfügbar ist, um die zuvor gelöschten Daten wiederherzustellen und den Benutzer weiterhin identifizieren und verfolgen zu können. Beispiele hierfür sind die Nutzung von *HTML5 Storage*, die Speicherung von Cookies in einem *HTML5 History Object*, aber auch die Verwendung von *HTML5 Canvas Tags* in Kombination mit dem Browser-Cache.[2] Eine bekannte Bibliothek für persistente Cookies ist beispielswiese *evercookie*[3].

[2] Die hier aufgeführten Technologien werden eingesetzt, um moderne Web-Apps umzusetzen und sind daher breit verfügbar. Deren genaue Funktion ist im Zusammenhang mit Benutzer-Tracking irrelevant, da es sich dabei um eine missbräuchliche Verwendung Webtechnologien handelt.
[3] http://samy.pl/evercookie/

2.2.2 Über lokale Eigenschaften

Auf Grund zunehmender Sicherheitsmaßnahmen welche die Identifizierung von Benutzern und Benutzerinnen über lokal hinterlassene Spuren unterbinden oder erschweren, wurden verschiedene Methoden entwickelt, um die Aktivitäten von Benutzern und Benutzerinnen anhand übermittelter Eigenschaften weiterhin zu verfolgen zu können. Diese werden im Folgenden diskutiert.

2.2.2.1 Browser-Fingerprinting

Unter *Browser-Fingerprinting* versteht man die Erkennung bzw. Identifizierung eines Browsers bzw. einer Browser-Instanz anhand der Eigenschaften. Im Gegensatz zur Erkennung über gespeicherte Daten müssen bei dieser Technik keine Daten am lokalen PC abgelegt werden. Dadurch ist es auch schwieriger, sich dieser Nutzerverfolgungs-Technik zu entziehen. Gängige Bibliotheken verwenden unter anderem vom Browser übermittelte Informationen, wie Header-Informationen, Bildschirmauflösung, Sprache oder Zeitzone, sowie Informationen über installierte Plugins. Abbildung 2 zeigt beispielsweise wieder jene Informationen, die beim Aufruf des Webauftrittes der Österreichischen Nationalbank gesendet werden. Auch hier soll das Beispiel zeigen, dass in der Kommunikation mit einer vertrauenswürdigen Organisation möglicherweise identifizierende Informationen übermittelt werden, selbst wenn dies nicht der Zweck ist, sondern etwa die Information über das zugreifende System ein besseres Benutzer- und Benutzerinnenerlebnis über angepasste Darstellung ermöglichen soll. Das Beispiel soll damit veranschaulichen, dass andere Organisationen solche Information nutzen können, um Besucherinnen und Besucher des Webauftritts ohne deren Wissen zu identifizieren. Es ist für Benutzerinnen oder Benutzer kaum unterscheidbar, ob Information des Systems abgefragt wird, um den Webauftritt besser zu gestalten, oder eine Organisation dies dazu verwendet, das Benutzerinnen- und Benutzerverhalten besser auszuwerten.
Browser-Fingerprinting ist jedoch nicht auf oben genannte Methoden limitiert. Vielmehr werden laufend neue Fingerprinting-Techniken entwickelt. Trotzdem bietet jede Informationsquelle für sich nur recht wenige aussagekräftige Informationen. Außerdem sind auch nicht immer alle Quellen benutzerspezifischer Informationen verfügbar. Daher kombinieren gängige Bibliotheken alle verfügbaren Informationen zu einem einzigen Fingerabdruck. Dieser Fingerabdruck ist meist recht eindeutig. Beispielsweise gelang es Eckersley [3] 94,2% der Browser mit aktiviertem Flash und Java

```
GET / HTTP/1.1
Host: www.oenb.at
Connection: keep-alive
Pragma: no-cache
Cache-Control: no-cache
Upgrade-Insecure-Requests: 1
User-Agent: Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/53.0.2785.116 Safari/537.36
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,*/*;q=0.8
Accept-Encoding: gzip, deflate, sdch
Accept-Language: de-DE,de;q=0.8,en-US;q=0.6,en;q=0.4
```

Abbildung 2: Request Header beim Aufruf von www.oenb.at

eindeutig zu identifizieren. Die meisten dieser Informationen verändern sich kaum oder nur sehr selten, wodurch der gewonnene Fingerabdruck über längere Zeit die Identifizierung des Browsers ermöglicht. Auch der Inkognito-Modus schützt nur begrenzt, da einige Werte trotzdem abrufbar sind. Einige der Informationen sind zudem unabhängig vom verwendeten Browser, wodurch teilweise auch eine browserübergreifende Identifizierung möglich wird. Eine bekannte Open-Source-Fingerprinting-Bibliothek ist beispielsweise *fingerprintjs2*[4].

2.2.2.2 Geräteübergreifende Identifizierung

Ein weiteres Mittel um Benutzer und Benutzerinnen über lokale Eigenschaften zu identifizieren, ist die Verwendung von *cross-device targeting* oder *cross-device tracking*. Dabei wird versucht, Internetnutzer und Internetnutzerinnen über verschiedene Geräte wie Notebooks, Smartphones, Tablets aber auch Fernseher oder Wearables zu identifizieren bzw. einem Nutzer oder einer

[4] https://github.com/Valve/fingerprintjs2/

Nutzerin mehrere Geräte zuzuordnen. Dies erlaubt es beispielsweise gezielte Werbung unabhängig vom verwendeten Endgerät oder dem aktuellen Standort anzuzeigen.
Um Benutzer und Benutzerinnen geräteunabhängig zu identifizieren, gibt es mehre Möglichkeiten. Ein einfacher Ansatz ist ein verpflichtender Authorisierungsprozess vor der ersten Verwendung von Geräten. Durch diesen Loginvorgang werden mehrere Geräte einem eindeutigen Benutzerprofil zugeordnet. Ebenso existieren komplexere Methoden die eine Identifikation von Gerätebesitzern und Gerätebesitzerinnen zulassen. Dazu zählt die Verwendung von statistischen Modellen über die Gerätenutzung und übermittelter Daten sowie die Verwendung von nicht hörbaren Tönen im Ultraschallbereich, welche in Fernsehwerbung oder Werbung im Webbrowser abgespielt werden. Geräte in der Umgebung (wie z.B. Smartphones) können diese Töne aufzeichnen, und ermöglichen somit eine Zuordnung von mehreren Endgeräten zu einem Nutzer bzw. einer Nutzerin. Bis dato existieren jedoch keine Möglichkeiten diese Audio-Cookies zu blockieren oder zu umgehen. Aus diesem Grund soll im Projekt SoniControl [4] eine Applikation entwickelt werden, die automatisiert akustische Cookies identifiziert, und diese gegebenenfalls blockiert.

2.2.3 Über Netzwerkeigenschaften

Trotz vorhandener Sicherheitsmaßnahmen in Browsern bzw. Browserimplementierung wie dem Inkognito-Modus ist die Identifikation von Benutzern und Benutzerinnen grundsätzlich über die im Netzwerk hinterlassenen Spuren wie zum Beispiel der IP-Adresse oder DNS-Anfragen möglich. Der Grund dafür ist, dass die zugrundeliegende Infrastruktur, basierend auf IP-Adressen, nicht auf Anonymität ausgerichtet ist. Dadurch besitzt jedes Gerät im Internet eine, nach außen hin öffentliche IP-Adresse, die für korrektes Routing im Netzwerk auch erforderlich ist. Besucht man Seiten im Web, so propagiert diese IP-Adresse durch das Netzwerk. Angeforderte Webseiten sowie beteiligte Infrastrukturen wie der ISP oder der DNS-Server können somit Benutzer und Benutzerinnen eindeutig identifizieren und deren Aktivitäten nachverfolgen.
Außerdem erlaubt die Verwendung von MAC-Adressen im lokalen Netzwerk zusätzlich Rückschlüsse auf Benutzer und Benutzerinnen abseits des Internets zu ziehen.

2.3 Bereiche des Internets

Durch die zugrundeliegenden Protokolle bietet das Internet, ohne zusätzliche Maßnahmen, relativ wenig Anonymität. Verschiedene Netzwerke mit unterschiedlichen Routingmechanismen und Technologien versuchen diese Probleme zu beheben. Für eine bessere Differenzierbarkeit haben sich die Begriffe *Clearnet* für den unverschlüsselten bzw. öffentlichen Teil des Internets, und *Darknet* als Gegenstück entwickelt. Die Bedeutung dieser Begriffe wird im Folgenden behandelt.

2.3.1 Clearnet

Der Begriff *Clearnet* beschreibt das Gegenstück zum Darknet. Das Clearnet umfasst all jene Netzwerke über die standardmäßig unverschlüsselt oder nicht anonym kommuniziert wird. Im Clearnet ist es im Allgemeinen möglich, Benutzer und Benutzerinnen via deren IP-Adresse zu identifizieren. Als wohl bekannteste Clearnet-Anwendung kann das World Wide Web genannt werden. Dieses besteht wiederum aus zwei Bereichen: dem *Surface Web* und dem *Deep Web*. Die nächsten Abschnitte beschreiben diese Begriffe ausführlicher.

2.3.1.1 Surface Web – Visible Web

Bright Planet [5] definierte 2001 den Begriff *Surface Web* (oder auch *Visible Web*) als jenen Teil des World Wide Web, der öffentlich erreichbar und durch Suchmaschinen zugänglich und indiziert ist. Das Surface Web besteht hauptsächlich aus statischen Webseiten die mittels Webcrawlern gefunden werden können. Dabei werden, ausgehend von einer Webseite, alle weiterführenden Hyperlinks untersucht und die Seiten in einen Index aufgenommen. Seiten, die keine eingehenden Links haben, sind laut dieser Definition folglich nicht Teil des Surface Web.
In der Fachliteratur finden sich allerdings teils unterschiedliche Definitionen und Interpretationen des Begriffs *Surface Web*. Eine einheitliche Definition scheint nicht zu existieren. Aus Gründen der Einfachheit wird der Begriff *Surface Web* deshalb wie folgt definiert: Das Surface Web wird als der Teil des World Wide Web bezeichnet, der öffentlich, ohne zusätzliche Software oder Zusatzmechanismen, erreichbar ist.

2.3.1.2 Deep Web – Hidden Web – Invisible Net

Laut Bergman [5] sind, im Gegensatz zu Seiten aus dem Surface Web, Seiten des *Deep Web* (auch *Hidden Web* oder *Invisible Net*) über herkömmliche Suchmaschinen nicht auffindbar. Sie beinhalten dynamischen Inhalt, der über Suchanfragen an Datenbanken generiert wird oder der nicht öffentlich zugänglich ist. Herkömmliche Link-Crawler können folglich über Hyperlinks nicht direkt zu Deep Web Seiten gelangen.
In der Fachliteratur finden sich auch zum Begriff *Deep Web* unterschiedliche Definitionen. Der Begriff wird im Rahmen diese Studie folgendermaßen definiert: Das Deep Web ist jener Teil des World Wide Web, der nicht öffentlich, also nur mit zusätzlichem Wissen (z.B. Passwort) oder Mechanismen (z.B. Zuordnung einer bestimmten IP Adresse), erreicht werden kann. Bekannte Beispiele für Seiten im Deep Web sind private Facebookprofile, oder nur hinter Paywalls verfügbare Dokumente und Fachliteratur, wie sie z.B. von *Springer Link* angeboten werden.

2.3.2 Darknet

Für den Begriff *Darknet* existiert keine einheitliche Definition. Aked et al. [6] beispielsweise definieren ihn als ein verschlüsseltes Kommunikationsnetzwerk, das es erlaubt, über ein *Peer-to-Peer-Netz* (P2P-Netz), (anonym) zu kommunizieren bzw. Daten auszutauschen. Um das Darknet vom regulären Internet zu unterscheiden, wurde der Begriff *Clearnet* definiert. Zwar agieren Netzwerke im Darknet ebenfalls innerhalb des Internets, allerdings unterscheiden sie sich in deren Charakteristiken und verfolgten Prinzipien vom Clearnet. Während Benutzer im Clearnet im Allgemeinen einfach identifiziert und deren Aktivitäten verfolgt werden können, erschweren Maßnahmen, wie Verschlüsselung und spezielle Routingverfahren, in Darknets die Überwachung durch Dritte. Aufgrund dieser Eigenschaft, wird meist spezielle Software vorausgesetzt, um Darknets zu erreichen. Abbildung 3 gibt einen Überblick über diese Topologie.
Im Zusammenhang mit dem Begriff *Darknet* wurden verschiedene Terminologien identifiziert, die fälschlicherweise als Synonym für diesen Begriff verwendet werden. Um eine klare Unterscheidung zwischen dem Deep Web oder genauer gesagt zwischen dem Clearnet und Darknet zu erreichen wird der Begriff *Darknet* als ein verschlüsseltes Peer-to-Peer-Netzwerk, über das anonym kommuniziert werden kann, definiert.

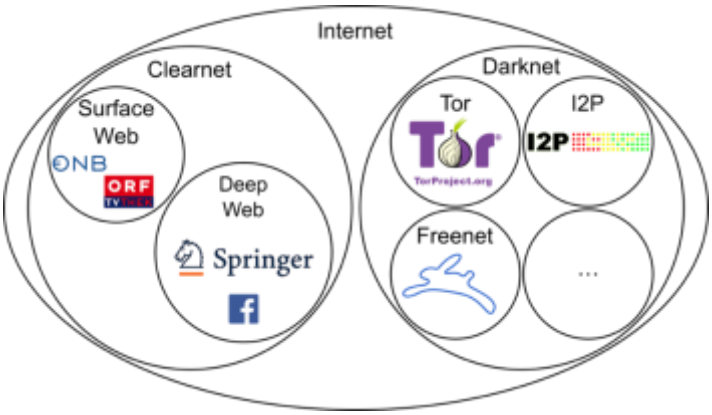

Abbildung 3: Aufbau von Clearnet und Darknet

2.3.2.1 Grundlegende Eigenschaften von Darknets

Ein definiertes Ziel von Darknets ist es, dass Datenverkehr von Dritten nicht abgefangen, gelesen oder modifiziert werden kann. Der Einsatz von Verschlüsselung ist deshalb ein wesentlicher Bestandteil von Darknets. Ein weiteres Ziel ist die Wahrung der Anonymität von Benutzern und Benutzerinnen. Durch spezielle Routingmechanismen, in Kombination mit Verschlüsselung, soll so der Zusammenhang zwischen Ursprung und Endpunkt einer Anfrage verborgen werden. Im Gegensatz zu Routing im Clearnet, das unter anderem den Standort, die Bandbreite und Auslastung zur Berechnung heranzieht, haben Darknets das Ziel, die Identität eines Users zu verschleiern. Aus diesem Grund wird die Route nicht nur nach Latenz gewählt, sondern wechselt regelmäßig, um eine Nachverfolgung zu erschweren. Aked et al. [6] definieren fünf Eigenschaften anhand derer Darknets identifiziert werden können. Dazu zählen der Einsatz von Verschlüsselung, die Wahrung von

Anonymität, die verfügbaren Applikationen, die Art des Routings sowie der Sichtbarkeit nach außen. Populäre Darknets, die diese Eigenschaften aufweisen, sind zum Beispiel *Tor*[5], *I2P*[6] oder *Freenet*[7].

2.3.2.2 Dark Web

Als *Dark Web* werden Webseiten im Darknet bezeichnet, welche nur über spezielle Software erreichbar sind. Owen et al. [7] haben gezeigt, dass eine Vielzahl der Seiten im Dark Web illegale Inhalte, wie z.B. Online-Schwarzmärkte oder Zugang zu Botnetzen, bereitstellen. Gleichzeitig bietet das Dark Web aber auch Whistleblowerplattformen oder Seiten mit politisch motivierten Inhalten an. Um eine eindeutige Abgrenzung zwischen Deep Web und Dark Web zu gewährleisten, wird das Dark Web im Rahmen dieser Studie als Darknet-Gegenstück zum World Wide Web definiert, das nur durch zusätzliche Software erreicht werden kann. Jedoch gibt es auch zu diesem Begriff unterschiedliche Definitionen in der Literatur.

2.4 Bezahlung im Internet

Neben gängigen Zahlungsmitteln wie zum Beispiel der Zahlung per Rechnung, Überweisung oder Kreditkarte, gibt es im Internet weitere Zahlungsmöglichkeiten, wie beispielsweise PayPal[8], um Einkäufe im Internet zu tätigen. Die Nutzung dieser Bezahlformen zieht allerdings einen Nachteil mit sich. Anonyme Bezahlvorgänge sind dadurch kaum oder nur schwer realisierbar. Dabei kann es verschiedene Gründe für das Verlangen nach anonymen Zahlungsmitteln geben. Neben der Bezahlung von illegalen Produkten gibt es auch eine Reihe (legaler) Gründe und Argumente für anonyme online Bezahlmöglichkeiten. Während man beispielsweise in der realen Welt Produkte nahezu jederzeit ohne Preisgabe der eigenen Identität kaufen kann, wird bei jedem Kauf im Internet die Identität von Benutzern und Benutzerinnen festgestellt – ein Eingriff in die Privatsphäre, der nicht zwingend notwendig ist. Aus diesem Grund existieren einige Anbieter, die auf anonyme Online-Zahlungen spezialisiert sind. Beispiele hierfür sind unter anderem *paysafecard*[9], Kryptowährungen wie *Bitcoin* oder auch Prepaid-Kreditkarten. Meist kann bei diesen Anbietern vor dem eigentlichen Bezahlvorgang Guthaben in bar eingekauft werden und später anonym im Internet ausgeben werden.

2.5 Spurenverwischung

Betrachtet man die in Abschnitt 2 diskutierten Aspekte, wird offensichtlich, dass bereits beim Aufruf einer statischen Webseite eine Vielzahl an Spuren hinterlassen wird. Dem steht das Recht auf bzw. das Verlangen nach Privatsphäre gegenüber. Auf Grund der schier unüberblickbaren Menge von Aktivitäten, welche Spuren am benutzten Gerät, im Netzwerk, auf Webseiten und bei Serviceanbietern hinterlassen, ist es jedoch für den Normalverbraucher bzw. die Normalverbraucherin nahezu unmöglich, sich wirklich anonym im Internet zu bewegen. Meist ist vollständige technische Anonymität auch ein unerreichbares Ideal. Kombiniert man jedoch die verschiedenen Möglichkeiten, unterschiedliche Spuren zu unterdrücken oder zu verwischen, ist es möglich, sich im Internet nahezu unsichtbar für Außenstehende zu bewegen. Tabelle 1 gibt einen Überblick darüber, welche Art von Spuren beim Aufruf einer unverschlüsselten Webseite hinterlassen werden.

---

[5] https://torproject.org/
[6] https://geti2p.net/
[7] https://freenetproject.org/
[8] https://www.paypal.com
[9] https://www.paysafecard.com

	Am Gerät	Im Netzwerk	Bei Dritten	Am Webserver
MAC-Adresse	/	○		
IP-Adresse	/	●	●	●
Position	/		●	●
DNS-Anfrage	●	●		
Verwendeter Browser	●	●	●	●
Betriebssysteminformation	/	●	●	●
Zeitzone	/	●	●	●
Fingerprint	●	●	●	●
Benutzter Dienst	●	●	●	●
Zugangsdaten	●	●		●
Payload	●	●		●
Cookies	●	●	●	●
Verlauf	●	●	○	○
(Logs, DB-Einträge, …)	●		●	●

Tabelle 1: Hinterlassene Spuren beim Aufruf einer Webseite

(/... nicht Anwendbar, ●... Information verfügbar, ○... Information teilweise verfügbar)

Je nach durchgeführter Aktivität im Internet, beispielsweise durch online Bezahlvorgänge, lassen sich diese Spuren allerdings noch erweitern. Abgesehen vom digitalen Fingerabdruck ist es in einigen Fällen notwendig, seine persönlichen Daten, wie z.B. Adresse, oder auch die eigenen Zahlungsinformationen, bei Servicebetreibern zu hinterlegen. Sobald bei der Benutzung von Online-Diensten, ein Zusammenhang zwischen realer Welt und Online-Welt hergestellt werden muss, ist Anonymität somit nicht mehr möglich. Derartige Aspekte werden aus diesem Grund im Rahmen dieser Studie nicht miteinbezogen. Dennoch ist es unter bestimmten Umständen möglich, vollständig anonyme Geschäfte im Internet abzuwickeln. Dazu zählt beispielsweise der Handel von ausschließlich digitalen Gütern unter der Verwendung von anonymen Bezahlmethoden.
Im Folgenden werden zuerst unterschiedliche Methoden zur Verwischung, bzw. Vermeidung von Spuren im Internet diskutiert und anschließend konkrete Implementierungen dieser Konzepte analysiert und (sofern dies sinnvoll möglich ist) verglichen. Hierbei gilt zu beachten, dass nur wenige Implementierungen drauf ausgelegt sind, falsche Spuren zu legen, bzw. einmal bestehende Spuren zu verwischen. Ein Großteil der existierenden Methoden beschränkt sich lediglich auf die Vermeidung von Spuren.

3 Methoden

Die Vermeidung von Spuren im Internet kann auf Grund der vielen Faktoren, die dabei zusammenspielen, nur durch den Einsatz unterschiedlicher Methoden, in ausreichendem Maß sichergestellt werden. In diesem Abschnitt werden deshalb existierende Methoden diskutiert, welche verschiedenste Arten von Spuren vermeiden können. In diesem Zusammenhang spielen auch völlig untechnische und technologieunabhängige Aspekte eine Rolle. Dazu zählen beispielsweise die im anschließenden Abschnitt 3.1 diskutierten anonymen Netzzugänge.

3.1 Anonyme Netzzugänge (Öffentliches WLAN, Internetcafés, Wertkarten, ...)

Eine Möglichkeit die eigene Identität zu verschleiern ist die Benutzung eines fremden Netzzuganges. Ist keine Authentifizierung bei der Benutzung des Netzzuganges erforderlich bzw. ist dem Betreiber die eigene Identität nicht bekannt, hat man einen anonymen Zugang zum Internet. Dazu eignen sich

beispielsweise öffentliche oder schlecht gesicherte Drahtlosnetzwerke die von vielen Flughäfen, Cafés, Restaurants und Privatpersonen zur Verfügung gestellt bzw. betrieben werden. Auch öffentliche Internet-Terminals, welche oft von Universitäten, Bibliotheken oder teilweise sogar von Städten angeboten werden, eignen sich für einen anonymen Zugang. Ebenso eignen sich Daten-Wertkarten, die anonym gekauft werden, für das anonyme Surfen im Internet. Dabei sollte beachtet werden, dass das benutzte Gerät nicht identifizierbar ist, also die IMEI nicht einem Benutzer oder einer Benutzerin zugeordnet werden kann. Trotzdem gilt zu beachten, dass durch Techniken wie Triangulierung dennoch der Standort des Gerätes herausgefunden werden kann. Darüber hinaus werden bereits beim Aufruf einer einfachen Webseite eine Vielzahl anderer Spuren hinterlassen, die eine Wiedererkennung des benutzten Geräts bzw. des Benutzers oder der Benutzerin ermöglichen. Daher bieten anonyme Netzzugänge keine vollständige Anonymität. Vielmehr sind öffentlich zugängliche Netze ein nützliches Mittel, welches ergänzend zu anderen Maßnahmen eingesetzt werden kann.

3.2 Mix-Netze

Mixe (auch: *Mix-Netze*) beschreiben ein von David Chaum [5] vorgeschlagenes Konzept für ein Routingverfahren zur anonymen Kommunikation innerhalb von Netzwerken. Ziel ist es, die Kommunikation zweier Personen im Netzwerk anderen Netzwerkteilnehmern gegenüber zu verschleiern, ohne dass dabei eine zusätzliche, vertrauenswürdige Instanz involviert ist. Dafür werden Nachrichten nicht direkt übertragen, sondern über mehrere Teilnehmer (genannt *Mixe*), im Netzwerk geroutet. Der Routingpfad wird im Gegensatz zu traditionellen Routingverfahren wie *Open Shortest Path First (OSPF)* [8] oder *Routing Information Protocol (RIP)* [9] vor dem Versenden der Nachricht definiert. Eine weitere Besonderheit des Verfahrens besteht darin, dass jeder Mix über ein *public/private key pair* verfügt, dessen öffentlicher Schlüssel allen anderen Teilnehmern und Teilnehmerinnen bekannt ist, bzw. bekannt gemacht wird.
Sobald der Pfad und die öffentlichen Schlüssel aller Knoten dieses Pfades bekannt sind wird die Nachricht verschlüsselt. Hierbei handelt es sich um eine mehrschichtige Verschlüsselung, welche in der dem Pfad entgegengesetzten Reihenfolge der Knoten erfolgt. Soll beispielsweise eine Nachricht an einen Mix *B* entlang des Pfades $C \rightarrow A \rightarrow B$ übertragen werden, so wird die Nachricht zuerst für *B* verschlüsselt. Das resultierende Kryptogramm wird anschließend für *A* verschlüsselt und schließlich unter dem öffentlichen Schlüssel von *C* (siehe Abbildung 4) ein letztes Mal verschlüsselt. Insgesamt kommen in diesem Beispiel somit drei Schichten von Verschlüsselung zum Einsatz.
Empfängt ein Mix eine Nachricht, so wird die äußerste Schicht entfernt, indem die Nachricht entschlüsselt wird. Dabei spielt es zu diesem Zeitpunkt keine Rolle, ob es sich dabei um einen Mix entlang des Pfades oder den endgültigen Empfänger handelt. Tatsächlich kann ein Mix erst nach der Entschlüsselung feststellen ob er der finale Empfänger einer Nachricht ist, oder ob diese an einen weiteren Knoten weitergeleitet werden muss. Durch den Einsatz asymmetrischer Kryptografie kann jeder Teilnehmer nur für ihn verschlüsselte Nachrichten – und folglich immer nur eine Schicht einer Nachricht – entschlüsseln. Die Konsequenz dieser Eigenschaft ist, dass garantiert werden kann, dass Knoten im Netzwerk nichts über den Inhalt der Nachricht sowie den ursprünglichen Sender oder endgültigen Empfänger in Erfahrung bringen können. Jeder Mix kann folglich nur Informationen über dessen unmittelbare Nachbarknoten innerhalb eines Pfades extrahieren, nicht jedoch über die gesamte Route. Abbildung 5 illustriert die Übertragung einer Nachricht im Netzwerk. Diese Technik zur Verschlüsselung bzw. Übertragung von Nachrichten wird auch (angelehnt an die mehrschichtige Verschlüsselung) *Onion Routing* genannt. Ein Einsatzgebiet dieses Verfahrens ist neben der Anonymisierung von Traffic innerhalb eines Mix-Netzes auch die Verschleierung der eigenen Identität gegenüber externen Kommunikationspartnern. Dabei fungiert der letzte Knoten innerhalb eines Pfades als Gateway zur Außenwelt.

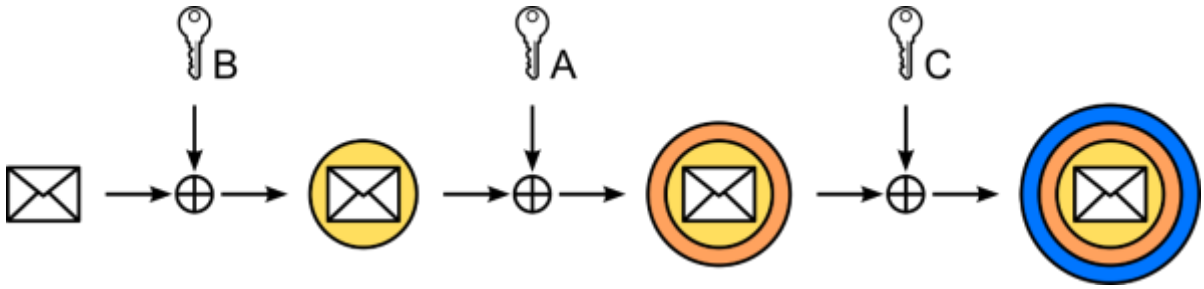

Abbildung 4: Verschlüsselung einer Nachricht vor dem Senden

Ein weiteres wichtiges Ziel des Mix-Konzepts ist, dass selbst der endgültige Empfänger einer Nachricht keine Rückschlüsse auf die Identität des Senders ziehen kann, gleichzeitig aber in der Lage sein muss, diesem zu antworten. Realisiert wird dieses Vorhaben durch anonyme Rückadressen. Diese Rückadressen müssen für jede gesendete Nachricht neu generiert werden, um Deanonymisierungsversuchen auf Basis von Korrelationen des Nachrichtenverkehrs oder sich wiederholenden Mustern entgegenzuwirken. Auf technischer Ebene wird dieses Konzept durch eine Kombination von asymmetrischer und symmetrischer Kryptografie umgesetzt. Das notwenige Schlüsselmaterial wird vom Sender generiert und jeder Nachricht beigelegt. Dieses besteht aus einem Einweg-public/private-keypair, sowie einem symmetrischen Schlüssel für jeden Knoten im (Rücksende-)Pfad. Tatsächlich werden die symmetrischen Schlüssel zusätzlich unter dem öffentlichen Schlüssel des zugehörigen Mixes verschlüsselt, damit jeder Mix nur auf den für ihn bestimmten symmetrischen Schlüssel Zugriff hat.
Ausgehend vom Empfänger einer Nachricht werden folgende Schritte durchgeführt: Der Empfänger verwendet den öffentlichen Teil des vom ursprünglichen Sender generierten Einweg-Schüsselpaars, um seine Antwort zu verschlüsseln und leitet diese inklusive aller (verschlüsselten) symmetrischen Schlüssel an den ersten Knoten des Rücksendepfades weiter. Dieser ist in der Lage den für ihn bestimmten symmetrischen Schlüssel zu dechiffrieren und die empfange Nachricht damit zu verschlüsseln. Die Nachricht wird somit um eine zusätzliche Verschlüsselungsschicht ergänzt, bevor diese an den nächsten Knoten im Pfad weitergeleitet wird. Folglich wird die Nachricht am Rückweg an jedem Knoten um eine weitere Verschlüsselungsschicht ergänzt. Dieser Vorgang ist notwendig, um zu garantieren, dass sich Nachrichten auch am Rückweg verändern, um somit nicht nachverfolgt werden zu können. Am Empfänger (bzw. dem ursprünglichen Sender, der auf eine Antwort wartet) angekommen, wird die Verschlüsselung wieder Schicht für Schicht abgetragen, bis schließlich der private Teil des Einwegschlüsselpaares zum Einsatz kommt, um auch die letzte Verschlüsselungsschicht zu entfernen. Da der Empfänger selbst dieses Schlüsselmaterial generiert hat, sind ihm auch alle Schlüssel bekannt, welche zur vollständigen Entschlüsselung notwendig sind.

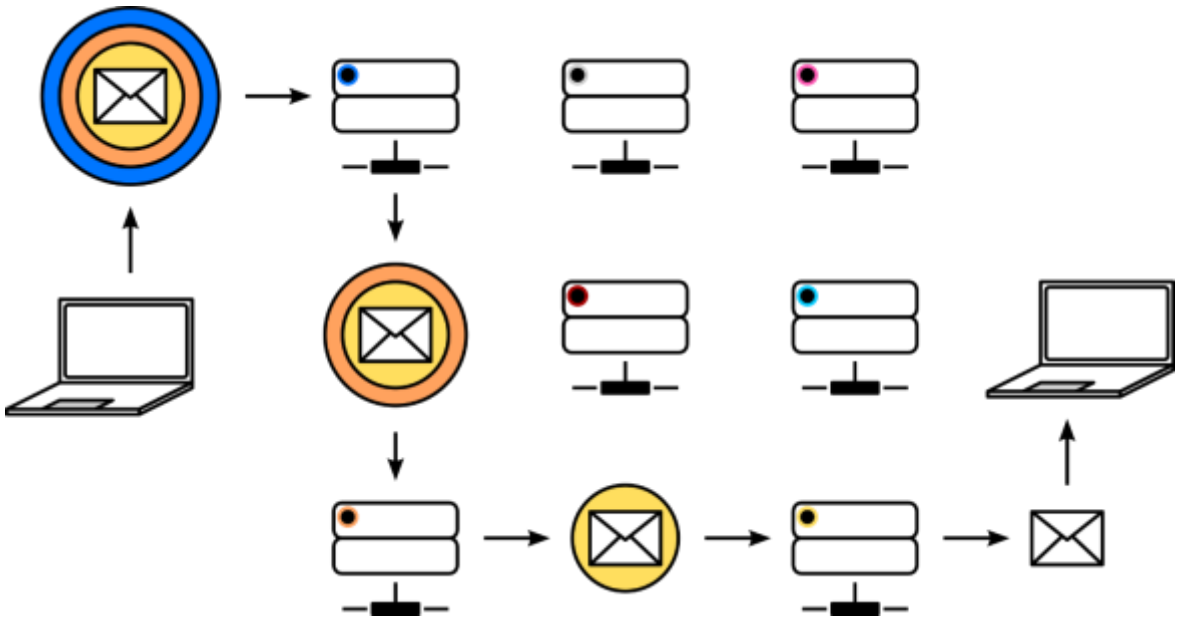

Abbildung 5: Übertragung einer Nachricht im Mix-Netzwerk

Trotz dieser Mechanismen ist es unter bestimmten Umständen immer noch möglich Informationen über Sender sowie Empfänger einer Nachricht zu erhalten. Beobachtet man den Zeitpunkt des Eintreffens einer Nachricht, sowie den Zeitpunkt deren Weiterleitung, lassen sich Zusammenhänge ableiten und die Route einer Nachricht kann nachverfolgt werden. Jeder Mix hat daher zu garantieren, dass es keine von außen feststellbare Korrelation zwischen angenommenen Nachrichten und weitergeleiteten Nachrichten gibt. Um dies sicherzustellen, werden die Nachrichten von mehreren Sendern gesammelt, und diese in anderer Reihenfolge an den nächsten Knoten gesendet.
Für größtmögliche theoretische Sicherheit wird die Verwendung von *Mix-Kaskaden*, bei denen es nur einen einzigen vertrauenswürdigen Pfad (oder eine Menge von vorgegebenen Pfaden) gibt, empfohlen. Dieser Vorgang minimiert die Gefahr von bösartigen Knoten. Dadurch ergibt sich allerdings auch ein *Single Point of Failure.* Weiteres wird hierbei ein uneingeschränktes Vertrauen gegenüber dem Betreiber der Mix-Kaskade vorausgesetzt, was einen Widerspruch zur ursprünglichen Definition von Chaum darstellt. Aufgrund der Struktur von Mix-Kaskaden benötigen diese außerdem hohe Netzwerkbandbreiten. Aus diesen Gründen werden häufig Mix-Netze, bei denen vom Sender ein zufälliger Pfad aus einer Liste von Mixen ausgewählt wird, verwendet.

3.3 Proxies

Proxies werden im Allgemeinen als Schnittstelle in einem Netzwerk eingesetzt. Sie agieren dabei als Vermittler um Anfragen von einem Kommunikationspartner an den Nächsten weiterzureichen. Nach außen hin bleibt somit typischerweise die ursprüngliche Adresse von Teilnehmern im Netzwerk verborgen. Im Gegensatz zu *Network Address Translation (NAT*) [10], das typischerweise in Routern eingesetzt wird, können Proxies aktiv den Netzwerkverkehr beeinflussen, Anfragen protokollieren oder je nach verwendeter Technologie Datenströme analysieren sowie gegebenenfalls sogar modifizieren. Proxies können aber auch genutzt werden um länderspezifische Sperren oder Firewalls zu umgehen. Ebenso können sie verwendet werden, um Zugriff auf beispielsweise durch Zensur gesperrte Dienste zu erhalten. Im Web finden sich deswegen eine große Anzahl an offenen anonymen Proxies, die die Herkunft, respektive die IP-Adresse, von Anfragen verstecken sollen. Hierbei gilt es allerdings zu beachten, dass die eigene Adresse nur nach außen hin verschleiert wird. Anfragen zum Proxy sowie der zugehörige Datenverkehr können sehr wohl einer IP-Adresse zugeordnet werden.
Als sogenannte *Reverse Proxies* eingesetzt, ermöglichen es Proxies außerdem, Zugriffe auf einen Server oder ein privates Netzwerk zu limitieren bzw. dieses zu schützen sowie Load-Balancing oder Caching durchzuführen.
Im Grunde kann man zwischen zwei Arten von Proxies unterscheiden: Generische Proxies und protokollspezifische Proxies. Diese werden im folgenden Absatz im Detail diskutiert.

3.3.1 Generische Proxies

Generische Proxies dienen gewöhnlich dazu, Datenverkehr protokollunabhängig zwischen einem Server und einem Client weiterzuleiten. Als Beispiel sei hierbei *SOCKS (Socket Secure) genannt*, das in der aktuellen Version 5 nebst Benutzer-Authentifizierung, Unterstützung für UDP Verkehr zur Verfügung stellt. SOCKS agiert völlig unabhängig vom zugrundeliegenden Verkehr, mit dem Ziel, Anfragen bzw. Datenverkehr weiterzuleiten, um beispielsweise Firewalls zu umgehen.
Vor der eigentlichen Anfrage ist eine Verbindungsanfrage an den Proxy notwendig. Nach erfolgreichem Handshake, der unter anderem Authentifizierungsmaßnahmen voraussetzen kann, agiert SOCKS im Anschluss völlig transparent. Im Gegensatz zu herkömmlichen HTTP-Proxies, modifizieren SOCKS-Proxies also den Datenverkehr nicht. Dadurch können im Grunde beliebige Anwendungen über SOCKS genutzt werden.

3.3.2 Protokollspezifische Proxies

Protokollspezifische Proxies agieren, im Gegensatz zu SOCKS, auf dem gesendeten Datenverkehr auf der Applikationsebene. Ihr Ziel ist es, Funktionalität für die Weiterleitung von Web-Verkehr bereitzustellen oder um sicherzustellen, dass der Datenverkehr festgelegten Richtlinien entspricht. Den Großteil im Netz verfügbarer protokollspezifischer Proxies machen so genannte Web-Proxies aus.

Da protokollspezifische Proxies auf den Daten arbeiten, können sie im Grunde die Anfragen und auch die Antworten beliebig modifizieren, und somit beispielsweise automatisiert Header-Informationen hinzufügen oder entfernen sowie die Nutzdaten anpassen. Nach außen hin ist es somit möglich Attribute, die es erlauben Benutzer und Benutzerinnen zu identifizieren, automatisch zu filtern. Der Vorteil von protokollspezifischen Proxies liegt also darin, dass sie genutzt werden können, um Spuren im Datenverkehr automatisch zu entfernen. Gleichzeitig bedeuten diese Proxies auch einen massiven Einschnitt in die Privatsphäre der Benutzer und Benutzerinnen, da der Datenverkehr analysiert, sowie Anfragen modifiziert und protokolliert werden können. Ein Vertrauen hinsichtlich des Betreibers ist also in diesem Fall notwendig. Wird ein Proxy selbst gehostet, können derartige Bedenken klarerweise von vornherein ausgeräumt werden.
Eine Ausnahme stellen HTTPS-Proxies dar, bei denen der Client zuerst nur eine Verbindungsanfrage an den Web-Proxy stellt. Ähnlich wie bei SOCKS findet dazu erst ein Handshake statt. Nach erfolgreicher Verbindung wird der Datenverkehr anschließend analog zu SOCKS getunnelt. Durch die Verwendung von Zertifikaten, ist es dadurch dem Client möglich eine sichere Verbindung zu einem Server herzustellen, ohne dass der Datenverkehr mitgelesen oder modifiziert werden kann. Voraussetzung dafür ist natürlich eine intakte Zertifikatskette.

3.4 SSH-Tunnel/VPN

Eine Möglichkeit Spuren sowohl auf Netzwerkebene, als auch teilweise auf Applikationsebene zu verschleiern, bieten *Virtual Private Networks* (VPNs). VPNs bieten die Möglichkeit ein logisches (virtuelles) Netzwerk über bestehende Infrastruktur sowie unterschiedliche Netzwerke aufzubauen. Teilnehmer eines VPNs können (je nach Konfiguration) direkt miteinander kommunizieren, da diese sich innerhalb eines einzigen (wenn auch virtuellen) Netzwerks befinden. Im Regelfall setzen VPNs auch Verschlüsselung ein, um die übertragenen Daten vor Dritten (wie z.B. anderen Teilnehmern, die sich im selben physischen Netzwerk befinden) zu schützen. Des Weiteren können VPNs so konfiguriert werden, dass Verbindungen nach außen über ein Gateway erfolgen, wodurch die eigene IP-Adresse vor externen Kommunikationspartnern verborgen wird. In jedem Fall werden jedoch Spuren innerhalb des VPN hinterlassen.
Eine weitere Möglichkeit, um einen ähnlichen Verschleierungseffekt zu erzielen, bietet *Secure Shell* (SSH). Ursprünglich wurde SSH entwickelt, um entfernte Logins auf Workstations oder Mainframes zu ermöglichen. SSH setzt zwingend Verschlüsselung voraus, um jegliche Kommunikation abzusichern. Mittlerweile bieten SSH-Implementierung auch die Möglichkeit, beliebigen Netzwerkverkehr über einen so genannten *SSH-Tunnel* umzuleiten. Hierzu wird eine SSH-Verbindung zu einem entfernten Rechner aufgebaut, welcher dieselbe Funktion erfüllt, wie auch ein VPN-Gateway. Ist ein SSH-Tunnel aufgebaut, tritt man auf Netzwerkebene gegenüber anderen Kommunikationsteilnehmern nicht mehr mit der eigenen IP-Adresse, sondern mit der IP-Adresse der SSH-Gegenstelle, auf.
Hier gilt es jedoch ebenfalls zu beachten, dass allein die Verschleierung der eigenen Identität auf Netzwerkebene nicht verhindert, dass auf Applikationsebene andere Identifizierungsverfahren zum Einsatz kommen.

3.5 Browser-Plugins

Die meisten modernen Browser unterstützen Erweiterungen in Form von Add-ons oder Plugins. Dadurch kann die Funktionalität des Browsers von Drittanbietern über definierte Schnittstellen erweitert werden. Im Rahmen dieser Studie sind besonders Erweiterungen, die die Sicherheit oder Anonymität des Nutzers bzw. der Nutzerin verbessern, interessant. Neben Werbeblockern sind außerdem Erweiterungen, welche die automatische Ausführung von Skripten stoppen, relevant. Beispiele hierfür sind *Ghostery*[10] und *NoScript*[11].

3.6 Werbeblocker

Bei Werbeblockern oder auch Werbefiltern handelt es sich um Computerprogramme oder Module eines Programms (Plugins), welche unerwünschte (elektronische) Werbung filtern bzw. ausblenden. Werbung kommt auf Webseiten in vielen verschiedenen Ausprägungen vor, beispielsweise als Pop-

[10] https://www.ghostery.com/
[11] https://noscript.net/

ups, Texte, Bilder, Videos oder in Form von Flash-Anwendungen. Typischerweise verwenden Werbeblocker eine Blacklist die laufend aktualisiert wird. Bindet eine Website Daten von einer Domain ein, welche auf der Blacklist steht, greift der Werbeblocker ein und verhindert je nach Ansatz das Laden oder zumindest das Anzeigen der Werbung. Zusätzlich blockieren gängige Werbeblocker meist Pop-ups. *AdBlock*[12] ist ein Beispiel für einen weitverbreiteten Werbefilter. Werbeblocker können nicht nur die Privatsphäre schützen, sie können auch die Sicherheit erhöhen, indem verhindert wird, dass Daten an Werbenetzwerke übertragen werden. Weiters kann auch der Stromverbrauch reduziert werden, da weniger (oftmals animierte) Inhalte angezeigt werden, was insgesamt weniger Rechenleistung erfordert.

4 Implementierungen

Die in diesem Abschnitt beschriebenen Implementierungen zeigen den konkreten Einsatz der in Abschnitt 3 beschriebenen Methoden zur Spurenverwischung und anonymen Nutzung des Internets. Dabei werden sowohl eigenständige Applikationen, als auch Erweiterungen für bestehende Applikationen diskutiert, die alleine oder in Kombination mit anderen Techniken dazu führen können, dass Spuren im Internet vermieden werden. Je nach angestrebter Anonymität oder gewünschter Funktionalität können dabei unterschiedlichste Implementierungen relevant sein.

4.1 Tor

Bei *Tor* (ursprünglich: *TOR – The Onion Routing* [11]) handelt es sich um ein Anonymitätsnetzwerk, das auf dem Prinzip eines Mix-Netzes basiert (siehe Abschnitt 3.2). Besonderen Wert wurde bei der Entwicklung von Tor einerseits auf geringe Latenzen, andererseits auf hohe Sicherheit und einen hohen Grad von Anonymität gelegt. Aus diesem Grund verfügt Tor auch über integrierte Abwehrmechanismen gegen verschiedene Deanonymisierungsmethoden. Um auch gegenüber Zensurmaßnahmen bestehen zu können, bietet Tor die Möglichkeit, den Zugang zum Tor-Netz entweder über geheime Zugangspunkte (sogenannte *Bridges*) herzustellen [12], oder den Netzwerkverkehr mit Hilfe von *Pluggable Transports* derart zu verschleiern, dass dieser nicht mehr als Tor-spezifischer Traffic identifizierbar ist [13].
Prinzipiell gibt es zwei Einsatzgebiete für Tor: Einerseits kann es verwendet werden, um bestehende Dienste und Webseiten anonym zu benutzen, andererseits ist es auch möglich, Server innerhalb des Tor-Netzes zu betreiben, sogenannte *Hidden Services*, welche nur innerhalb des Tor-Netzwerks erreichbar sind. Da sich diese Anwendungsfälle in einigen wesentlichen Punkten unterscheiden, werden sie in den folgenden Abschnitten separat diskutiert.

4.1.1 Tor als Anonymisierungsproxy für bestehende Dienste

Die Tor-Software umfasst unter anderem einen SOCKS5-Proxy, welcher dazu benutzt wird, um von beliebigen Anwendungen (und Web-Browsern) ins Tor-Netz einzusteigen. Soll nun ein anonymer Zugriff auf eine Webseite erfolgen, wird (ausgehend vom Proxy) ein Pfad durch das Tor-Netz bestehend aus drei Knoten aufgebaut: *Entry/Guard Node*, *Relay* und *Exit Node*. Wie der Name vermuten lässt, verbinden sich Nutzer zu einem Entry Node, welcher den Netzwerkverkehr über einen Relay Node an einen Exit Node weiterleitet. Dieser wiederum stellt die Verbindung zur Außenwelt bzw. zum gewünschten Dienst her. Obwohl Tor das Prinzip eines Mix-Netzwerks umsetzt und verschachtelte Verschlüsselung wie in Abschnitt 3.2 einsetzt, bleibt ein gewisses Restrisiko bestehen, deanonymisiert zu werden. Wenn ein einzelner Angreifer den ersten und letzten Knoten im Pfad kontrolliert, können Nutzer vollständig identifiziert werden [14].
Prinzipiell kann jede Anwendung und jede Art von Netzwerkverkehr über das Tor-Netz umgeleitet werden, so lange die Übertragung auf TCP basiert [15]. Hierbei muss jedoch unbedingt bedacht werden, dass alleine die Nutzung von Tor nicht bedeutet, dass man sich vollständig anonym im Internet bewegen kann [16]. Alle in Abschnitt 2 diskutierten Aspekte müssen weiterhin berücksichtigt werden. Tor garantiert lediglich Anonymität aus Sicht des Netzwerks. Wenn eine Webseite beispielsweise Cookies, oder spezielle Header setzt und/oder andere Trackingmechanismen auf der Anwendungsschicht zum Einsatz kommen, ist man auch als Tor-Nutzer deanonymisierbar. Es gibt aber auch Anwendungsfälle, die ohnehin voraussetzen, dass sich Nutzer gegenüber einem Dienst identifizieren. In solchen Situationen kann das Ziel sein, sicherzustellen, dass man im lokalen

[12] https://getadblock.com/

Netzwerk (oder innerhalb eines Landes, welches starke Zensurmaßnahmen verfolgt) aus Sicht des Netzwerks anonym bleibt.
Der in Abschnitt 4.7 beschriebene *Tor-Browser* bietet einen einfachen Einstieg um möglichst anonym im Internet zu surfen, da dieser bereits so vorkonfiguriert ist, dass zumindest einige Trackingmöglichkeiten umgangen werden. Trotzdem muss man als Nutzer Webseiten sehr bewusst nutzen und sein Surfverhalten entsprechend ändern, wenn man das höchstmögliche Maß an Anonymität verfolgt. Wirklich vollständig anonym können mittels Tor die im nachfolgenden Abschnitt beschriebenen Hidden Services benutzt werden.

4.1.2 Hidden Services

Zusätzlich zur anonymen Nutzung bestehender Dienste, bietet Tor die Möglichkeit, Server innerhalb des Tor-Netzwerks zu betreiben. Diese so genannten *Hidden Services* sind von außen nicht erreichbar und auch nicht auffindbar [15] und stellen Dienste im Darknet, wie zum Beispiel anonyme E-Mail-Services, zur Verfügung. Will man einen solchen Dienst nutzen, benötigt man die Tor-Software und das Wissen um die Adresse des Servers. Diese hat die Form `x.y.onion` und ist abgesehen von der .onion-Pseudodomain nicht von einer Domain üblicher Webseiten und Dienste zu unterscheiden. Im Gegensatz zu einer IP-Adresse oder einer Domain, lassen sich jedoch von .oinon-Adressen keine Rückschlüsse auf den Ort, an dem der Server betrieben wird, ziehen.
Die Bereitstellung von Hidden Services funktioniert analog zur Nutzung eines Proxies: Die Tor-Software fungiert in diesem Fall als *Reverse Proxy*, leitet also aus dem Tor-Netzwerk eingehenden Verkehr an einen lokal betriebenen Server weiter. Da in diesem Fall der gesamte Netzwerkverkehr innerhalb des Tor-Netzwerk stattfindet, ist es nicht möglich von außen festzustellen, wo ein Hidden Service betrieben wird und wer diesen nutzt. Dank der Eigenschaften von Mix-Netzen ist es auch innerhalb des Netzwerks nicht möglich, einen Hidden Service oder die Nutzer eines solchen Dienstes aufzuspüren. Der Verbindungsaufbau zu einem Hidden Service wird nicht direkt durchgeführt, sondern über einen Mittelsmann. Ausgehend vom Server wird ein Pfad (wie in Abschnitt 3.2 beschrieben) zu einem Mittelsmann (welcher als *Rendezvous Point* bezeichnet wird) aufgebaut. Auch Nutzer können sich lediglich zu diesem Rendezvous Point verbinden. Wenn diese beiden partiellen Pfade durch das Netzwerk gebildet sind, werden sie vom Mittelsmann zu einer durchgehenden Verbindung verbunden [15]. Nutzer und Server wissen somit lediglich um den Pfad zum Mittelsmann, dieser aber weder den Pfad zum Server noch zum Nutzer. Folglich ist garantiert, dass Server, Nutzer und Rendezvous Point einander nicht aufspüren können. Auch hier gilt jedoch zu beachten, dass die benutzten Server und Applikationen so konfiguriert werden müssen, dass eine Deanonymisierung nicht durch Informationen, welche auf der Applikationsschicht ausgetauscht werden, ermöglicht wird.
Hidden Services werden beispielsweise für Umschlagplätze illegaler Waren und Dienstleistungen aller Art verwendet [17]. Jedoch gibt es auch zahllose legitime Einsatzmöglichkeiten für anonyme Dienste. Beispielsweise kann Tor genutzt werden, um derart abhöhrsichere Kommunikation umzusetzen, dass es sogar unmöglich ist, festzustellen, dass überhaupt kommuniziert wird. Ein Beispiel hierfür ist *OnionCat* [18], welches ein VPN-Service über Tor umsetzt. Auch Whistleblower-Plattformen können aus Gründen der Zensurresistenz in Form von Hidden Services innerhalb des Tor-Darknets betrieben werden. Des Weiteren bietet Tor auf Grund seiner Robustheit die Möglichkeit Services selbst hinter Firewalls zu betreiben, welche Serverbetrieb eigentlich verhindern sollten. Zusätzlich ist es möglich, mobile Hidden Services zu betreiben: Da ein Hidden Service von der physischen Position unabhängig ist, ist es z.B. möglich, Server auf Smartphones zu betreiben, während man in Bewegung ist um somit trotzdem erreichbar zu bleiben.
Im folgenden Abschnitt wird das System *I2P* beschrieben, welches ähnliche Grundsätze wie Tor verfolgt, jedoch andere Ziele anstrebt.

4.2 I2P

Das Onion-Routing-Netzwerk *I2P* (auch: *The Invisible Internet Project* [19]) basiert auf einem ähnlichen Konzept wie Tor. Allerdings strebt I2P nicht an, beliebigen Netzwerkverkehr und möglichst alle bestehenden Dienste zu anonymisieren. Im Gegensatz zu Tor, welches einen applikations- und protokollneutralen, generischen SOCKS5-Proxy zur Verfügung stellt, strebt I2P eine möglichst tiefgreifende Integration zwischen Applikation bzw. Dienst und dem I2P-Netzwerk an. Daher werden protokollspezifische Proxies eingesetzt, die Informationen auf Applikationsebene entfernen, welche

zur ungewollten Benutzeridentifikation herangezogen werden können (siehe Abschnitt 3.3.2). Außerdem werden Programmbibliotheken zu Verfügung gestellt, welche direkt in die Implementierung eines Dienstes bzw. Servers integriert werden können. Der Fokus des Projekts liegt im Gegensatz zu Tor auf Hidden Services (welche über .ip2-Adressen erreichbar sind), und nicht darin den Internetverkehr der Nutzer im großen Stil zu anonymisieren. Als Hauptziele können daher das zur Verfügung stellen I2P-basierter Applikationen und Bereitstellen einer Applikationsplattform genannt werden [20]. Dabei kommt ein eigenes UDP-ähnliches Protokoll namens *NTCP* zum Einsatz, welches (wie TCP und UDP) auf der Transportschicht angesiedelt ist. Grund für ein spezielles Protokoll ist unter anderem, dass auch Dienste wie File Sharing mittels *BitTorrent*, welche klassischerweise über UDP umgesetzt werden, über I2P realisierbar sein sollen [20]. Beispiele für Anwendungen, welche über protokollspezifische I2P-Proxies ins I2P-Netz eingebunden werden können, sind HTTP und IRC [19].

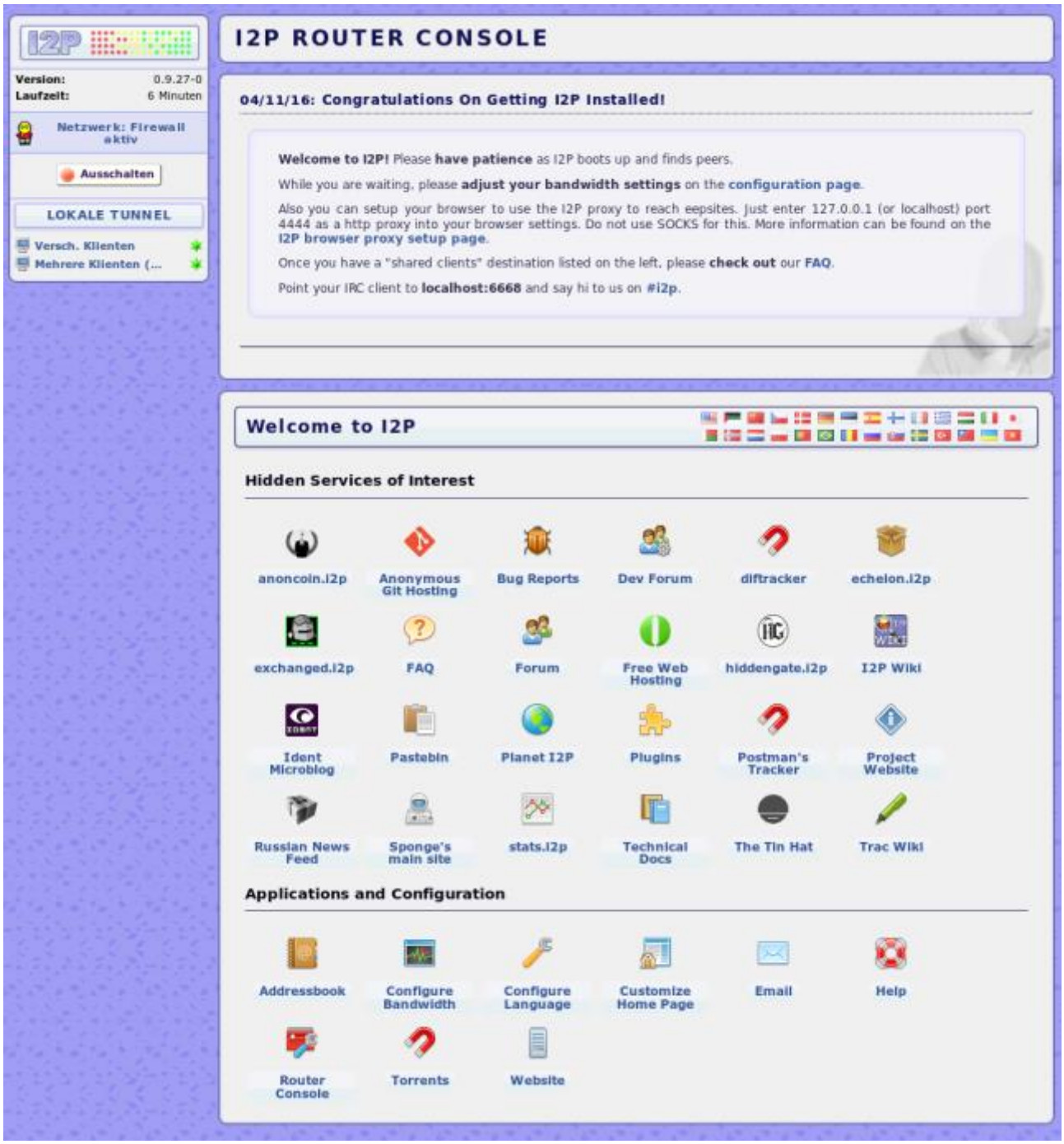

Abbildung 6: I2P Konsole

Die angestrebte direkte Integration von Diensten spiegelt sich auch im Gesamteindruck wider, der vermittelt wird, wenn man I2P installiert und zum ersten Mal in Betreib nimmt. Die Webseite, die dem Nutzer nach dem Starten des I2P-Dienstes präsentiert wird (siehe Abbildung 6), listet eine Reihe von Hidden Services auf, welche vom restlichen Internet – durch den Einsatz von NTCP sogar auf Netzwerkebene – abgeschottet sind. Das Resultat ist ein gut ausgebautes Darknet, welches eine umfangreiche Auswahl an hochintegrierten Diensten bereitstellt und auch direkt von den am Projekt beteiligten Personen empfohlen wird.
Das I2P-Netz ist auch intern anders organisiert als Tor, da Einstiegspunkte weniger strikt definiert sind. Dadurch ergibt sich ein gewisses Maß an Zensurresistenz, da blockierte Einstiegspunkte (ähnlich wie Tor Bridges) schlicht gegen andere ersetzt werden können [21]. Weiters verwendet I2P längere Pfade, um eine Verbindung herzustellen und benutzt außerdem für eingehende und ausgehende Verbindungen getrennte Pfade, welche als *Tunnel* bezeichnet werden [22]. Dadurch ergeben sich aus technischer Sicht unterschiedliche Sicherheitsmodelle. Eine Evaluierung, bzw. ein Vergleich, welcher auf Grund dieser Aspekte darlegt, ob Tor oder I2P sicherer und/oder anonymer nutzbar sind, ist leider nur schwer durchführbar. Konzeptionell sind Vergleiche zwar durchaus möglich, allerdings sind vor allem Implementierungsfehler oft der Grund dafür, dass formal sichere Systeme in der Praxis versagen und angreifbar sind. Da I2P um ein vielfaches weniger Ressourcen als dem Tor Projekt zur Verfügung stehen, ist speziell die Sicherheit der I2P-Implementierung in Frage zu stellen. Dieser Umstand zeigt sich auch, wenn man die Zahl und Qualität wissenschaftlicher Publikationen von Tor und I2P vergleicht. Quantitativ lässt sich lediglich feststellen, dass Tor weitaus mehr Nutzer hat als I2P und daher die Masse an Nutzern in der man selbst als Nutzer „untertauchen“ kann im Tor-Netzwerk um ein vielfaches höher ist [23] [24]. Hierbei gilt es jedoch zu beachten, dass weder für Tor, noch für I2P definitive Nutzerzahlen vorliegen, da eine genaue Erhebung technisch nicht möglich ist.

4.3 Beispiele für Proxies

Für die Nutzung von Proxies gibt es mehrere Möglichkeiten. Die einfachste Variante dabei ist die Verwendung von (kostenlosen) Anonymisierungsdiensten. Anbieter wie *HIDEme*[13] bieten dabei die Möglichkeit, Webseiten die durch Firewalls gesperrt oder die anonym genutzt werden sollen, ohne zusätzlichen Konfigurationsaufwand abzurufen. Dabei wird die Ressource vom jeweiligen Dienst abgerufen und anschließend dem Benutzer oder der Benutzerin angezeigt. Da diese Anonymisierungsdienste in den meisten Fällen keine direkte Verbindung (ein Tunnel) für den Benutzer oder die Benutzerin aufbauen, kann der Datenverkehr somit selbst bei der Verwendung von HTTPS vom Anbieter mitgelesen oder gegebenenfalls modifiziert werden. Speziell bei sensiblen Inhalten, wie bei der Übertragung von Passwörtern, ist somit Vorsicht geboten. Die meisten freien Anonymisierungsdienste werden durch die Einblendung von Werbung finanziert, die durch Modifikation der angeforderten Ressource eingeblendet wird. Dies erlaubt es außerdem, die Aktivitäten von Benutzern und Benutzern über verschiedene Sitzungen bzw. Seiten zu verfolgen.
Zusätzlich zu Anonymisierungsdiensten gibt es außerdem die Möglichkeit Proxies direkt im Browser, oder auf Betriebssystemebene festzulegen. Damit wird automatisch der entsprechende Verkehr über den Proxy geroutet. Falls unterstützt, kann der Proxy somit auch automatisch Filter für den Schutz der Privatsphäre anwenden. Als Beispiel sei hier *Privoxy*[14] genannt. Im Vergleich zu Anonymisierungsdiensten kann die Nutzung von betriebssystemweiten Proxies vor allem im Fall von HTTPS Anfragen entscheidende Vorteile bringen. Während ungesicherte Anfragen zwar nach wie vor geloggt werden können, wird bei HTTPS-Anfragen ein Tunnel zwischen Client und Server aufgebaut. Der Datenverkehr kann somit vom Proxy nicht mehr mitgelesen oder modifiziert werden. Eine Liste verfügbarer Proxies ist beispielsweise unter Netzwelt.de[15] zu finden.

[13] https://hide.me/de/proxy
[14] https://www.privoxy.org
[15] https://www.netzwelt.de/proxy/index.html

4.4 Anonyme Bezahlmöglichkeiten

Es gibt verschiedene Möglichkeiten zur anonymen Bezahlung im Internet. Zu den wichtigsten gehören *Bitcoin*[16], *cash4web*[17], *mywirecard*[18] und *paysafecard*[19]. Bitcoins können beispielsweise anonym an Automaten gekauft werden [25]. Eine Ausweispflicht herrscht erst ab einem gewissen Betrag. Im Falle des österreichischen Anbieters *Coinfinity*[20] handelt es sich um 250€ pro Tag und Person. Nach dem Kauf können die Bitcoins zur anonymen Bezahlung benutzt werden. Eine andere Möglichkeit an Bitcoins zu kommen, wird als *minen* bezeichnet. Verfügt man über entsprechend viel Rechenleistung, um ein kryptografisches Rätsel als erster zu lösen bekommt man als Belohnung Bitcoins ausbezahlt. Verfügt man nicht über die entsprechende Rechenleistung kann man sich einem Mining-Pool anschließen und bekommt entsprechend der zur Verfügung gestellten Rechenleistung einen Anteil der Belohnung. Im Gegensatz zu Bitcoin kann Guthaben für cash4web, mywirecard und paysafecard nicht vom Benutzer bzw. der Benutzerin generiert werden. Stattdessen kann bei verschiedenen Vertriebstellen wie beispielsweise Trafiken Guthaben gekauft werden. Beim Kauf bekommt man allerdings nicht direkt einen Voucher sondern einen Einmalcode um auf dieses Guthaben zuzugreifen. Im Falle von *paysafecard* handelt es sich beispielsweise um einen PIN. Anschließend kann das Guthaben bei allen Partnern eingelöst werden.

4.5 Ghostery

Bei *Ghostery* handelt es sich um eine Browser-Erweiterung, die gezielt Skripte blockiert, welche die Privatsphäre oder Anonymität der Nutzerin bzw. des Nutzers gefährden. Ghostery verwendet hierfür eine Datenbank von bekannten Trackern[21]. Diese umfasst derzeit über 2000 Einträge.

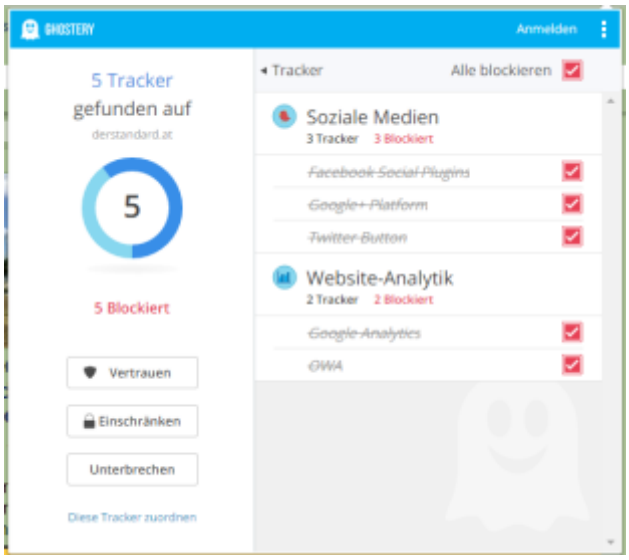

Abbildung 7: Ghostery auf derstandard.at

Die Tracker werden in verschiedene Kategorien eingeteilt, welche individuell geblockt bzw. erlaubt werden können. Diese Kategorien umfassen Werbung, Website-Analytik, Kundeninterkation, Soziale Medien, Essential, Audio/Video-Player, Werbung mit nicht jungendfreien Inhalten sowie Kommentare. Die Datenbank wird regelmäßig aktualisiert. Neben Skripten werden auch andere Verfahren wie bestimmte Cookies, Bilder, Canvas-Fingerprinting oder andere Webseiten-Elemente, welche die Privatsphäre oder die Anonymität gefährden, identifiziert. Auf Wunsch kann Ghostery zudem bestimmte Social-Media-Schaltflächen wie beispielsweise den „Like"-Button von Facebook

[16] http://bitcoin.org
[17] https://www.paylife.at/de/home/private/prepaid-cards.html
[18] http://www.mywirecard.com/
[19] https://www.paysafecard.com/de-at/
[20] https://coinfinity.co/
[21] Software mit dem Ziel Benutzer eindeutig zu identifizieren, und deren Aktivitäten über verschiedene Webseiten zu verfolgen

durch eigene „Click to Play“-Schaltflächen ersetzen. Diese erlauben die einfache Aktivierung ebendieser Schaltflächen durch einen simplen Klick, falls sie benötigt werden.

4.6 NoScript

NoScript [26] ist Erweiterung für Firefox, welche die Ausführung von JavaScript, Java, Flash und anderen ausführbaren Inhalten verhindert. Wahlweise können einzelne oder alle Inhalte erlaubt werden. Dadurch bietet NoScript einen effektiven Schutz gegen die meisten Fingerprinting-Verfahren, da deren Programmcode nicht ausgeführt wird. Zudem können Tracker keine Java- oder Flash-Anwendungen ausführen und daher auch keine Daten in deren Speichern abzulegen. Zusätzlich enthält NoScript auch einen Schutz geben Cross-Site-Scripting. Dieser verhindert, dass fremde Inhalte auf der besuchten Seite ausgeführt werden [27].

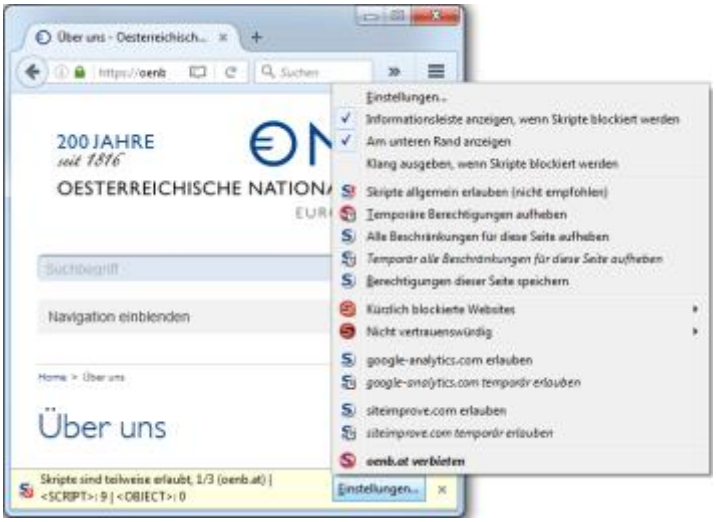

Abbildung 8: Firefox mit aktivierter NoScript-Erweiterung

4.7 Tor-Browser

Beim Tor-Browser, bzw. dem *Tor Browser Bundle* handelt es sich um ein vorkonfiguriertes, portables Paket, bestehend aus mehreren Komponenten. Dieses Paket ist für Windows, Mac OS X und Linux verfügbar und erlaubt auch Laien einen einfachen Einstieg ins Tor-Netzwerk. Neben Tor enthält das Paket eine speziell auf Sicherheit und Privatsphäre optimierte Version von Firefox mit erweitertem Support Zeitraum (*Extended Support Release*, kurz ESR). Angesehen von Änderungen am Quelltext wurden dazu auch einige Add-ons installiert. Zusätzlich wird *Torbutton,* welcher unter anderem Cookie-Management, User-Agent-, Locale- und Zeitzonen-Spoofing zur Verfügung stellt, verwendet. Der integrierte *TorLauchner* startet zuerst Tor und, sobald sich dieser zum Tor-Netz verbunden hat, Firefox. *NoScript* (siehe Abschnitt 4.6) ist ebenfalls im Tor Browser Bundle enthalten und verhindert die Ausführung von aktiven Inhalten. Um Verbindungen zu Webseiten automatisch verschlüsselt anzufordern, wird außerdem das Add-on *HTTPS-Everywhere* [28] eingesetzt. Dafür greift das Add-on auf eine Whitelist zurück. Wird eine Webseite von der Whitelist über eine unverschlüsselte Verbindung abgerufen, greift das Add-on ein und fordert stattdessen dieselbe Webseite über eine verschlüsselte Verbindung an. Im einfachsten Fall muss dazu nur „http“ durch „https“ ersetzt werden, das Add-on unterstützt aber auch komplexere URL-Umwandlungen. Durch die Kombination von Tor und Browser-Plugins, welche auf den Schutz der Privatsphäre ausgelegt sind, kann bereits durch den Einsatz des Tor-Browsers ein hohes Maß an Anonymität sowohl gegenüber dem Netzwerk als auch gegenüber Webseiten erreicht werden. Allerdings muss man als Nutzer abwägen, wann bestimmte Funktionen, wie z.B. JavaScript explizit aktiviert werden können, um z.B. bestimmte Funktionen einer Webseite nutzen zu können – ein zu sorgloser Umgang mit NoScript kann bereits ausreichen um die eigene Identität (unfreiwillig) preiszugeben.

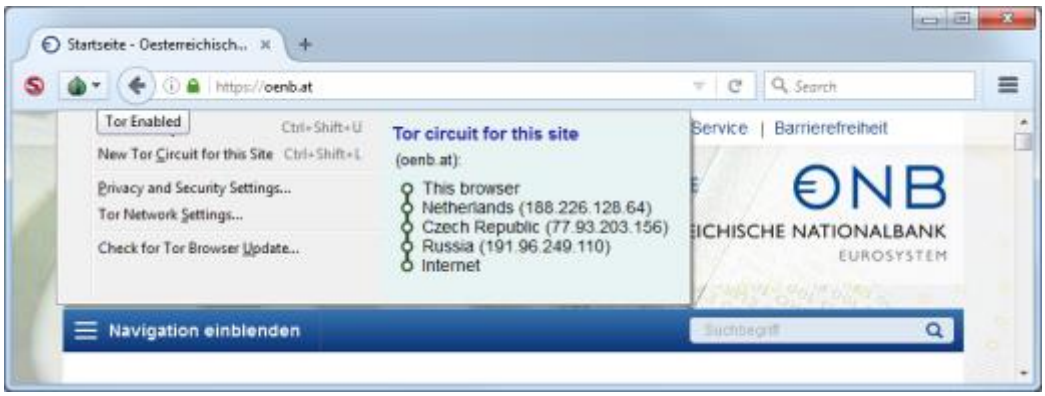

Abbildung 9: Tor-Browser mit Tor-Circuit

4.8 Tails

Bei *Tails* [29] handelt es sich um Linux-Distribution, welche auf Privatsphäre und Anonymität optimiert wurde. Tails steht für ***T**he **A**mnesic **I**ncognito **L**ive **S**ystem*. Wie der Name andeutet, handelt es sich um ein Live-Betriebssystem, welche direkt von DVD, USB-Stick oder SD-Karte aus gestartet werden kann. Da keine Installation notwendig ist und auch keine Schreibzugriffe auf die Festplatte erforderlich sind, werden am lokalen Rechner keine Spuren hinterlassen. Alle Verbindungen ins Internet werden durch den standardmäßigen Einsatz des Tor-Browsers, sowie einer systemweiten Tor-Installation automatische über das Tor-Netzwerk geroutet. Zusätzlich stellt Tails eine Reihe von nützlichen Tools zur Verfügung, wie *GnuPG* zur Verschlüsselung von Daten und Nachrichten, *KeePassX* zur sicheren Speicherung von Passwörtern oder *Florence*, eine virtuelle Tastatur, die versucht, Hardware-Keylogger zu umgehen. Doch auch hier gilt, dass die mitgelieferten Tools sehr bewusst eingesetzt werden müssen, um auch tatsächlich das höchstmögliche Maß an Anonymität zu gewährleisten und tatsächliche keine Spuren zu hinterlassen.

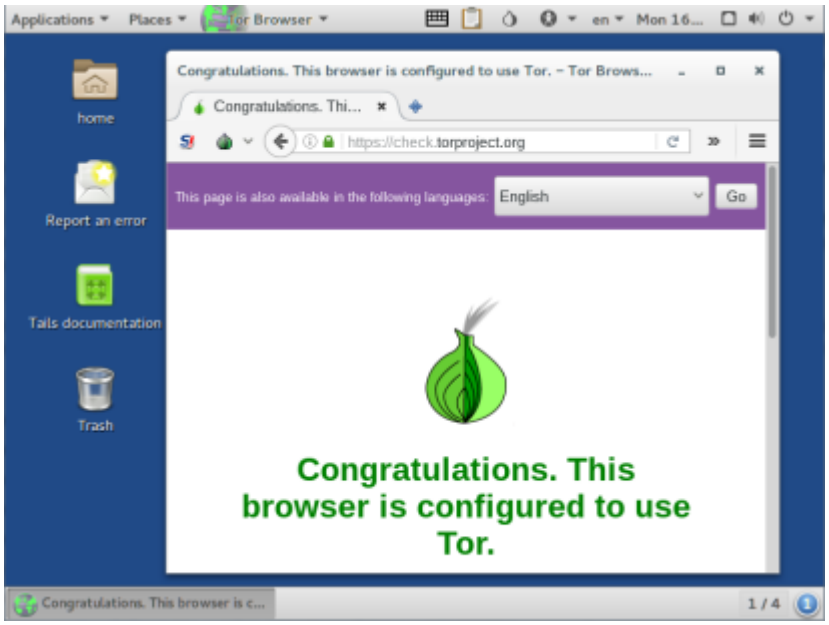

Abbildung 10: Tails mit geöffnetem Tor-Browser

5 Fazit

Bei allen Internet-Aktivitäten werden Spuren hinterlassen. Das ist einerseits dem Umstand geschuldet, dass die dem Internet zu Grunde liegende Infrastruktur in einer Zeit konzipiert wurde, als Datenschutz und Privatsphäre in diesem Zusammenhang schlichtweg irrelevant waren. Andererseits sind Nutzerdaten mittlerweile ein kostbares Gut, welches vor allem an werbetreibende Firmen gewinnbringend verkauft werden kann. Generell haben viele Parteien ein Interesse daran, Internetnutzer möglichst akkurat zu identifizieren und auch über unterschiedliche Geräte hinweg verfolgen zu können. Auch Aspekte der Redefreiheit und Zensur sind in diesem Zusammenhang relevant. Allerdings gibt es auch andere Motive, möglichst keine Spuren im Internet zu hinterlassen. Geht man beispielsweise illegalen Aktivitäten nach, möchte man ebenso wie bei der Umgehung von Zensurmaßnahmen unentdeckt bleiben.

Um dieses Ziel zu erreichen, bedarf es jedoch auf Grund der vielen unterschiedlichen Ebenen, welche zur Bereitstellung von Inhalten aus dem Internet involviert sind, einer Kombination unterschiedlicher Methoden. Einerseits können Onion-Routing-Netze wie Tor aber auch VPNs oder SSH-Tunnel verhindern, dass einzelne Nutzer oder Geräte auf Netzwerkebene identifiziert werden können. Gleichzeitig muss aber auch darauf geachtet werden, dass nicht auf Applikationsebene verräterische Informationen preisgegeben werden. Beispielsweise sind in HTTP-Headern Informationen über das Betriebssystem, die Systemsprache, und Zeitzone vorhanden, die vom Server ausgewertet werden können. Der Einsatz von Cookies und Referrern, erlaubt es, Nutzer und Nutzerinnen über Webseiten hinweg zu identifizieren und zu verfolgen. Um auch dieser Art von Tracking zu entgehen, können entsprechende Browser-Plugins und/oder applikationsspezifische Proxies eingesetzt werden. Bereits entsprechend vorkonfigurierte Softwarepakete wie der Tor-Browser sind frei verfügbar. Jedoch werden auch unter Einsatz einer Kombination ebengenannter Methoden Spuren am verwendeten Gerät hinterlassen. Um auch das zu verhindern, können Live-Systeme wie Tails verwendet werden, welche ohne Installation auf beliebigen Rechnen gestartet und ohne lokale Spuren zu interlassen wieder entfernt werden können. Allerdings garantiert auch der Einsatz sorgfältig vorkonfigurierter Software keine vollständige Anonymität, da diese Programme eine korrekte Nutzung voraussetzen. Tatsächlich ist vollständige Anonymität ein Ideal, das realistisch nicht erreichbar ist. Nutzer müssen sich darüber im Klaren sein, vor wem die eigene Identität geheim gehalten werden soll, und welche Aktionen bestimmte Spuren hinterlassen, um die eigenen Online-Aktivitäten gezielt vor ausgewählten Parteien zu verbergen. Ein Patentrezept, oder eine allumfassende Lösung zum Schutz der eignen Identität oder zur vollständigen Verschleierung der eigenen Spuren im Internet ist schon rein technisch nicht umsetzbar.

6 Abbildungsverzeichnis

7 Literaturverzeichnis

[1] P. Eckersley, „How Unique Is Your Browser?," *Proc. of the Privacy Enhancing Technologies Symposium (PETS),* pp. 1-18, 2010.

[2] S. Hayne und R. Rice, „Attribution Accuracy When Using Anonymity in Group Support Systems," *International Journal of Human-Computer Studies,* Bd. 47, pp. 429-452, 1997.

[3] P. Eckersley, „How Unique Is Your Web Browser?," in *Privacy Enhancing Technologies: 10th International Symposium, PETS 2010, Berlin, Germany, July 21-23, 2010. Proceedings*, Berlin, Springer Berlin Heidelberg, 2010, pp. 1-18.

[4] M. Zeppelzauer, „SoniControl," 2016. [Online]. Available: https://www.fhstp.ac.at/de/forschung/projekte/sonicontrol. [Zugriff am 7 12 2016].

[5] M. K. Bergman, „White Paper: The Deep Web: Surfacing Hidden Value," *Taking License,* Bd. 7, Nr. 1, 2001.

[6] S. Aked, C. Bolan und B. Murray, „Determining What Characteristics Constitute a Darknet," *Proceedings of the 11th Australian Information Security Management Conference,* pp. 11-20, 2013.

[7] G. Owen und N. Savage, „The Tor Dark Net," *Global Commission on Internet Governance Paper Series,* Nr. 20, p. 9, 2015.

[8] J. Moy, „OSPF Version 2," *RFC 2328,* 1998.

[9] G. Malkin, „RIP Version 2," *RFC 2453,* 1998.

[10] P. Srisuresh und M. Holdrege, „IP Network Address Translator (NAT) Terminology and Considerations," *RFC 2663,* 1999.

[11] P. Syverson, „Re: OSI 1-3 attack on Tor? in it.wikipedia," 13 02 2008. [Online]. Available: http://archives.seul.org/or/talk/Feb-2008/msg00145.html. [Zugriff am 04 11 2011].

[12] J. McLachlan und N. Hopper, „On the Risks of Serving Whenever You Surf: Vulnerabilities in Tor's Blocking Resistance Design," *Proceedings of the 8th ACm Workshop on Privacy in the Electronic Society,* pp. 31-40, 2009.

[13] The Tor Project, „Tor: Pluggable Transports," [Online]. Available: https://www.torproject.org/docs/pluggable-transports.html.en. [Zugriff am 04 11 2011].

[14] P. Syverson, G. Tsudik, M. Reed und C. Landwehr, „Towards an Analysis of Onion Routing Security," in *Designing Privacy Enhancing Technologies*, Deutschland, Springer, 2009, pp. 96-114.

[15] R. Dingeldine, N. Mathewson und P. Syverson, „The Second Generation Onion Router," *Proceedings of the 13th USENIX Security Symposium,* 08 2004.

[16] The Tor Project, „Want Tor to Really Work?," [Online]. Available: https://www.torproject.org/download/download-easy.html.en#warning. [Zugriff am 04 11 2016].

[17] N. Christin, „Traveling the Silk Road: A Measurement Analysis of a Large Anonymous Online Marketplace," *Proceedings of the 22nd International Conference on World Wide Web,* pp. 213-224, 2013.

[18] OnionCat, „About OnionCat," 2016. [Online]. Available: http://www.onioncat.org/about-onioncat/. [Zugriff am 04 11 2016].

[19] The I2P Project, „The Invisible Internet Project (I2P)," [Online]. Available: https://geti2p.net/en/about/intro. [Zugriff am 04 11 2016].

[20] C. Egger, J. Schlumberger, C. Kruegel und G. Vigna, „Practical Attacks Against The I2P Network," in *Research in Attacks, Intrusions, and Defenses*, Berlin, Springer, 2013, pp. 432-451.

[21] L. Schimmer, Interviewee, *On I2P.* [Interview]. 16 10 2015.

[22] J. P. Timpanaro, I. Chrisment und O. Festo, „A Bird's Eye View onthe I2P Anonymous File-Sharing Environment," in *Network and System Security*, Berlin, Springer, 2012, pp. 135-148.

[23] The Tor Project, „Tor Metrics - Direct users by country," [Online]. Available: https://metrics.torproject.org/userstats-relay-country.html?start=2015-09-30&end=2016-09-01&country=all&events=off. [Zugriff am 04 11 2016].

[24] The I2P Project, „Three Month View for Total Routers," [Online]. Available: http://stats.i2p/cgi-bin/total_routers_3month.cgi. [Zugriff am 04 11 2016].

[25] Coinfinity, „Bitcoins kaufen, einfach und schnell am Automaten," [Online]. Available: https://coinfinity.co/bitcoin-kaufen/. [Zugriff am 07 11 2016].

[26] G. Maone, „NoScript - JavaScript/Java/Flash blocker for a safer Firefox experience!," [Online]. Available: https://noscript.net/.

[27] G. Maone, „NoScript - JavaScript/Java/Flash blocker for a safer Firefox experience! - features," [Online]. Available: https://noscript.net/features#xss. [Zugriff am 2016 09 28].

[28] EFF, „HTTPS Everywhere," [Online]. Available: https://www.eff.org/Https-Everywhere. [Zugriff am 26 09 2016].

[29] Tails, „Tails - Privacy for anyone anywhere," 2016. [Online]. Available: https://tails.boum.org. [Zugriff am 16 November 2016].

[30] V. Vasilyev, „Fingerprintjs2," [Online]. Available: https://github.com/Valve/fingerprintjs2/. [Zugriff am 29 09 2016].

[31] D. Chaum, „Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM,* Bd. 24, Nr. 2, pp. 84-90, 1981.