Spuren im Speicher

Webbrowser als Informationslecks

> Gegen Ende unseres Artikels zum „Prinzip der Verfügbarkeit“ bei der Sicherheit der EU-Staaten (RHZ 2/2010) hatten wir ein wenig darüber philosophiert, ob Leute Google nach einem spannenden Suchwort befragt haben oder nicht und ob das BKA das war. Wir haben daraufhin etliche Rückfragen bekommen, wie wir darauf kommen. Um das zu beantworten, müssen wir etwas weiter ausholen auf die Spuren, die mensch beim Web-Browsen so im Netz hinterlässt – und auch den Mitteln, diese zu verwischen. Wer nun die Nase rümpft und sich fragt, seit wann wir hier VerbraucherInnenschutz diskutieren: Was dem Verbraucher die werbetreibende Wirtschaft, ist der Linken die Staatsgewalt...

Wenn ihr einen Link im Web anklickt, beginnt eine hektische Kommunikation zwischen eurem Web-Browser und mindestens einem Rechner irgendwo in der Welt, nicht selten gleich einer ganzen Handvoll davon. Wer sich das ansehen will, kann das beispielsweise mit der Firefox-Erweiterung LiveHTTPHeaders machen, die in unserem Werkzeug-Kasten erwähnt ist.

Das Ergebnis eines solches solchen Experiments: Eine typische kommerzielle Webseite wie zeit-online.de etwa überträgt im April 2010 für eine einzige „Webseite“ über 70 Dateien von 10 verschiedenen Servern, darunter Maschinen von Google, vom Werbewirtschaftsverein IVW, dem Holtzbrinck-Schwesterunternehmen Handelsblatt sowie dem Berliner „Webcontrolling“-Unternehmen Webtrekk. Bei all diesen Läden hinterlasst ihr Spuren. Welche?

Was ein Webserver loggt

Wenn ein Webserver euch eine Datei gibt – Text, Bild oder was immer –, schreibt er bei den meisten Seiten in eine Datei (ein so genanntes Log):

- Zeit und Datum des Zugriffs
- Die IP-Adresse des anfragenden Rechners
- Die angefragte Seite
- Einen „User Agent“, aus dem in der Regel hervorgeht, welcher Browser auf welchem Betriebssystem benutzt wurde
- Einen Referrer, der sagt, auf welcher Seite der Link stand, der zu der Datei geführt hat
- Etliche eher technische Daten, die eher weniger kritisch sind.

Es ist denkbar, dass Server noch mehr loggen, etwa, welche Sprachen der/die NutzerIn bevorzugt und ähnliches. Das erwähnte LiveHTTPHeaders ist wahrscheinlich der einfachste Weg, sich einen Einblick zu verschaffen in das, was der Server sieht. Nett ist auch http://panopticlick.eff.org, eine Webseite, die euch allerhand Informationen über euch selbst anzeigt – erschlossen einfach aus der Anfrage selbst.

Das kritischste Datum dabei ist sicher die IP-Adresse, denn sie identifiziert euch fast wie eine Telefonnummer. Im Gegensatz zur Telefonnummer ändert sich zwar eure IP-Adresse zwecks Extraprofit der Telekoms derzeit noch ständig. Die Zuordnung von Zeit und IP zu realen Personen ist aber für die Staatsgewalt für einige Zeit verfügbar; wie zuverlässig und lang, hängt ohne Vorratsdatenspeicherung von vielen Details ab. Das Bundesverfassungsgericht hat aber schon signalisiert, ein guter Rechtsstaat habe ein Anrecht auf solche Daten. Unterhalb eines mittleren Aufstands wird es also absehbar wieder für Monate garantierte Auflösung von IPs zu Personen geben.

Wie ihr um die Speicherung eurer IP-Adresse rumkommt, haben wir in RHZ 3/2009f in unserem Artikel zu Tor erläutert; viel einfacher geht das auch nicht. Bei den übrigen Daten gehts leichter, denn sie werden von eurem Browser im Groben freiwillig herausgerückt.

Identifiziert und Verwiesen

Die Fälschung des User Agent (im Groben Browser und Betriebssystem) ist vor allem interessant, wenn mensch mit eher exotischen Programmen unterwegs ist. Dann nämlich lässt der User Agent schon Schlüsse auf eure Identität zu („Aha, hier ist wieder der mit Opera auf MacOS“). Privoxy oder diverse Firefox-Erweiterungen lassen euch den User Agent auf unverfängliche Werte setzen. Das kann dann zwar dazu führen, dass Download-Seiten euch Software anbieten, die bei euch nicht läuft – aber Programme mit dem Browser runterladen ist eh keine empfehlenswerte Praxis, wenn euch sowas aus der Spur wirft.

Richtig dramatisch ist der Referrer. Klickt ihr beispielsweise einen Link bei Google an, so enthält der Referrer euer Suchwort – es dürfte sicher sein, dass z.B. das BKA Informationen dieser Art auswertet. Ebenfalls auf Referrern basieren im einfachsten Fall so genannte Web Bugs. Das sind kleine Bilder (z.B. die bekannten Buttons für „soziale Netzwerke“, aber auch unsichtbare Pixel), die von anderen Servern kommen. Euer Browser wird die Bildchen in der Regel brav von diesen fremden Servern ziehen und dabei die Referrer mit übertragen.

Nettoeffekt ist, dass für solche Seiten – praktisch alle Zeitungsseiten haben sowas, aber auch viele andere – noch an anderer Stelle nachvollziehbar wird, auf welchen Seiten ihr wart, und zwar potenziell für längere Zeit. Zusammen mit Cookies geht das nicht selten auch ohne die fragile Zuordnung von IP-Adressen und Personen (vgl. unten). Ist die Polizei schlau, fragt sie so nicht fünfzig Verlage, sondern zentral die Auflagenkontrolle von IVW („wer hat mehr als vier Artikel über das brennende Auto vom 20.2. gelesen?“).

Die Unterdrückung von Referrern stört fast nie; in dem Sinn ist die Installation der Firefox-Erweiterung Refcontrol (bzw. ähnlicher Mechanismen, die die Übermittlung von Referrern unterdrücken) uneingeschränkt empfehlenswert. Web Bugs könnten allerdings eine äquivalente Information auch direkt in der URL übertragen, vor allem im Zusammenhang mit Javascript. Um damit fertig zu werden, gibt es Erweiterungen wie Adblock und Counterpixel.

Cookies

An sich sind Cookies nichts Magisches: Es handelt sich einfach um ein paar Byte, die ein Server in eurem Browser speichern kann. Bei jedem Zugriff auf eine Seite von diesem Server schickt euer Browser die paar Byte an den Server zurück.

So harmlos das klingt, der Effekt ist fatal im Hinblick auf die Identifizierbarkeit eurer Spuren im Netz. Die diversen Google-Seiten z.B. geben euch beim Erstkontakt einen Cookie mit einer Identifikationsnummer, die euer Browser wieder an Google schickt, wenn ihr auf ihren Seiten unterwegs seid. Auf diese Weise können alle Zugriffe, die von diesem Browser aus gemacht werden, zusammengeführt werden, ob ihr nun irgendwie eingeloggt seid oder nicht. Wenn Google dann noch irgendwo Realnamen oder wenigstens eine Mailadresse bekommt, kann z.B. die gesamte Suchgeschichte personalisiert werden. Über Cookies korrelieren die Herausgeber der erwähnten Web Bugs auch die verschiedenen Zugriffe.

In dem Sinn könnte mensch einfach einem der Rezepte zum kompletten Abschalten von Cookies folgen. Es gibt aber durchaus sinnvolle Anwendungen für Cookies, und manches geht ohne sie gar nicht. Ein ganz guter Kompromiss ist oft, Cookies nach jeder „Sitzung“ (d.h., wenn ihr euren Browser verlasst), wegzuwerfen. Der Firefox hat dazu unter eine Einstellung unter Datenschutz („Behalten bis: Firefox geschlossen wird“). „Cookies von Drittanbietern akzeptieren“ solltet ihr vermutlich immer aus haben, dabei geht es gerade um die Web Bugs.

Wenn ihr ohnehin in dem Dialog seid, könnt ihr auch gleich sehen, wer alles Cookies bei euch hinterlassen hat und löschen, was euch zuwider ist – das ist hinter „Cookies anzeigen“, sortiert nach dem Rechner, der den Cookie sieht.

Javascript

Das größte Einfallstor für allerlei Überwachungsmöglichkeiten sind in Webseiten eingebettete Programme, die auf *eurem* Rechner laufen; nach der Sprache, die Webbrowser für diese Zwecke verstehen läuft der Problembereich unter dem Titel Javascript. Nun ist klar, dass ein Programm, das von einer mehr oder minder beliebigen Webseite kommt, nicht z.B. eure Mails oder Flugblattentwürfe lesen und naheliegenderweise an den Verfassungsschutz verhökern können sollte. Die Browserhersteller bemühen sich auch, Javascript-Programmen so etwas zu verbieten. Fehler passieren natürlich immer wieder.

Aber auch ohne solche Fehler erlaubt Javascript Seitenbetreibern, so gut wie alle Blockaden auszuhebeln, die ihr durch die erwähnten Erweiterungen aufbauen

könnt. Ein sehr verbreitetes Beispiel für Schnüffelei mit Javascript ist Google Analytics, das auf jeder zweiten Seite spannende Daten zum Verhalten im Web an Google und mithin potenziell an die Staatsgewalt liefert. In dem Sinn wäre es also richtig, Javascript einfach komplett abzuschalten, im Firefox etwa unter „Inhalt“ in den Einstellungen.

Das allerdings wird auf vielen Webseiten dazu führen, dass fast nichts mehr geht, insbesondere im „Web 2.0“. Typische Webmail-Anwendungenn etwa, auf die leider kaum mehr jemand verzichten will, tun ohne Javascript nicht viel. Deshalb sind Browser-Erweiterungen wie NoScript für Firefox dem kaum durchzuhaltenden globalen Abstellen von Javascript vorzuziehen. Dabei dürfen Webseiten erstmal kein Javascript ausführen. Der Browser sagt dann etwas wie „Die Seite wollte 7 Skripten laufen lassen“. Wer der Seite traut und etwas braucht, das ohne Javascript nicht geht, kann ihr die Skripterei dann vorübergehend oder permanent erlauben.

Das für Javascript Gesagte gilt doppelt für Flash – bei Javascript kommt immer noch die Quelle mit, so dass mit vertretbarer Mühe noch rauszukriegen ist, was so eine Seite gern tun würde. Flash hingegen kommt als undurchsichtiger Bytehaufen, in dem nichts mehr erkennbar ist und der nochmal ein Extra-Cookie-System mitbringt. Abgesehen davon ist Flash auch fies, weil es keine brauchbare freie (im Sinne von freier Rede) Laufzeitumgebung dafür gibt – das Flash-Plugin gibt es nur für Systeme, denen der Hersteller Adobe das erlaubt.

Dank neuer Standards im Web („HTML 5“) dürfte der Schrecken, auf Flash und also all die tollen Spielereien, die damit gemacht wurden/werden (angefangen bei youtube) verzichten zu müssen, in absehbarer Zeit recht überschaubar werden. Dann kommt einfach kein Flash-Plugin auf die Maschine und fertig ist das Moped. Bis dahin ist die zu NoScript analoge Erweiterung Flashblock ist ein Muss für Nerven und Privatsphäre.

Fazit

Der Datensammelwut von Seitenbetreibern kommen Webbrowser im Auslieferungszustand recht weit entgegen. Was aus VerbraucherInnensicht nur lästig ist, bedeutet für Kontaktpersonen und GefährderInnen in Zeiten staatlichen Zugriffs auf praktisch alle gesammelten Daten – meist per freundlichem Ersuchen, zur Not per Beschlagnahme – Ärger vor und hinter der Haustür. Drum: überlegt, was ihr im Netz tut und gönnt eurem Browser ein paar der Erweiterungen, die wir hier diskutiert haben.

Oh, und: Wenn ihr selbst eine Webseite betreibt: Loggt doch einfach nicht, oder jedenfalls keine IP-Adressen. Im Zweifel fragt den Serverbetreiber.

Datenschutzgruppe der Roten Hilfe Heidelberg

datenschutzgruppe@rotehilfe.de

PGP Fingerprint: a3d8 4454 2e04 6860 0a38 a35e d1ea ecce f2bd 132a

http://www.datenschmutz.de

Der Werkzeugkasten

Für den Browser Firefox gibt es unzählige Erweiterungen; hier wollen wir ein paar empfehlen, die der Datensammelwut entgegenwirken. Die Erweiterungen könnt ihr unter Extras/Add-ons suchen, herunterladen und installieren. Für andere Browser gibt es meist analoge Mechanismen.

NoScript Ein Muss – mit NoScript wird Seiten erstmal verboten, Javascript-Programme in eurem Browser auszuführen. Ihr könnt vertrauenswürdigen Seiten bequem erlauben, es doch zu tun, wenn ihr das wollt

Flashblock Noch ein Muss – Flash-Geblinke wird erstmal nur als Icon dargestellt. Wollt ihr es sehen, könnt ihr auf das Icon klicken und den Kram sehen. Schont nicht nur die Nerven, sondern lässt auch manchen Schnüffelangriff ins Leere laufen.

Refcontrol Unterdrückt die Übermittlung der Information, von welcher Seite ihr kommt – wärmstens empfohlen.

Adblock Eine etwas komplexe Erweiterung, die es erlaubt, eingeblendete Werbung zu blockieren.

LiveHTTPHeaders Wer sehen will, was euer Browser so ins Netz posaunt, kann sich das mit dieser Erweiterung ansehen. Der Kram sieht für ungeübte Augen erstmal wie völlige Wirrnis aus, aber mit etwas Geduld sieht mensch auch als Nicht-Nerd, was da so verhandelt wird

Privoxy Das ist keine Firefox-Erweiterung, sondern ein „Proxy“, eine Software, die zwischen den Browser und das Netz geht. Privoxy kann fast alles, was die erwähnten Firefox-Erweiterungen können, für alle Browser und ggf. auch für mehrere Rechner – aber es ist weit weniger komfortabel zu bedienen und komplizierter aufzusetzen. http://www.privoxy.org