Fachhochschule Brandenburg
Fachbereich Informatik und Medien

Nutzeridentifikation im Internet: Evaluation von Abwehrstrategien

Masterarbeit, vorgelegt von Christian Labuda

Motivation:

Diese Arbeit soll dem Nutzer beim Surfen im Internet mehr Kontrolle darüber geben, wer Daten über ihn sammeln darf. Auslöser für die Arbeit ist der Umstand, dass Anbieter im Internet immer mehr Daten über die Surfgewohnheiten sammeln, ohne dass der Nutzer dies weiß. Das Recht des Nutzers auf informationelle Selbstbestimmung wird dadurch eingeschränkt.

Ziel:

Das Ziel der Arbeit ist es den aktuellen Stand der Technik im Bereich der Nutzeridentifikation zu untersuchen und dabei weitere Technologien, die in Zukunft eine Rolle spielen könnten, in Betracht zu ziehen.

Außerdem werden bestehende Abwehrstrategien zur Nutzeridentifikation vorgestellt und untersucht. Der Focus liegt dabei auf der Untersuchung der Cookieabwehr aktueller Browser.

Nutzeridentifikation:

Nutzeridentifikation ist ein Prozess, der das Ziel hat, die bei der Nutzung einer Webseite anfallenden Daten eines einzelnen Nutzers, genau diesem Nutzer zuzuordnen. Der Nutzer ist dabei die Person, die die Webseite im Browser aufruft und die Eingaben macht.

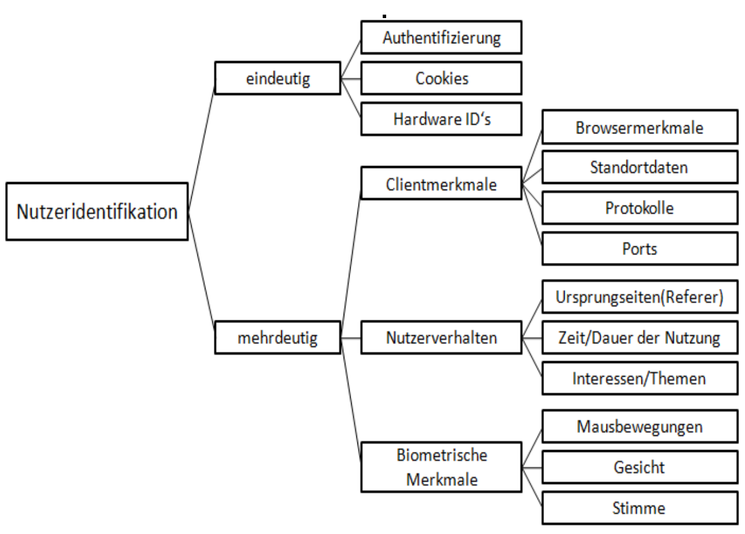
Der Prozess der Nutzeridentifikation besteht aus folgenden Schritten:

1. dem Auslesen der Daten,
2. der Zuordnung der Daten zu einer Session,
3. der Zuordnung der Sessiondaten zu einem abstrakten Profil des Nutzers und
4. der Zuordnung eines Profils zur realen Person.

Methoden:

Es ist möglich die Methoden zur Nutzeridentifikation in eindeutige und ähnlichkeitsbasierte Verfahren aufzuteilen. Bei ähnlichkeitsbasierten Verfahren wird der Nutzer nur mit einer bestimmten Wahrscheinlichkeit dem richtigen Profil zugeordnet.

Verschiedenste Methoden wurden im Rahmen der Arbeit untersucht, um Kenntnisse über deren Funktion und mögliche Schwachstellen zu erlangen.



Methodenübersicht: Nutzeridentifikation im Internet

Konzeption einer neuen Cookietechnologie:

Im Rahmen der Arbeit wurde ein Konzept für eine neue Cookieart auf der Basis von Java erstellt und implementiert. Dieser Cookie nutzt das Verhalten der Java Virtual Machine Java-Applets und JAR-Archive zum schnelleren Laden zwischenzuspeichern. Der eigentliche Cookie ist dabei ein JAR-Archiv. Die Menge, der durch diese Technik speicherbaren Daten, beträgt standardmäßig 1GB.

Der neue Java-Cache-Cookie wurde außerdem in die Testseite zur Evaluation der Browser integriert.

Abwehr der Identifikationsverfahren:

Aktuelle Browser enthalten bereits eine Vielzahl von Mechanismen, die verschiedene Identifikationsverfahren abwehren oder behindern können. Die Evaluation der Abwehrmechanismen im Rahmen der Arbeit sollte zeigen, ob diese die Erwartungen eines Nutzers erfüllen können.

Zusätzlich wurden weitere Möglichkeiten zur Abwehr von Nutzeridentifikationsverfahren analysiert.

Erstellung einer Testseite:

Um die Cookieabwehr verschiedener aktueller Browser zu untersuchen, wurde zunächst eine Testseite implementiert. Die Testseite bietet die Möglichkeit die Abwehr eines Browser gegen 12 verschiedene Cookiearten zu testen.

Evaluation der Cookieabwehr:

Eine erster Test ergab, dass die Abwehrmechanismen der Browser hauptsächlich gegen Cookies gerichtet sind und keinen Schutz vor ähnlichkeitsbasierenden Verfahren bieten.

Cookie	IE 9.0.8112 (8.0.6)	Firefox 5.0 (3.6.8)	Chrome 12.0 (9.05)	Safari 5.0.5 (5.0.2)	Opera 11.11 (11.01)
HTTP-Cookie	c (c)	- (-)	- (-)	- (-)	c (-)
Session-Storage	c (c)	c (c*)	c (c)	c (c)	c (c)
Local-Storage	c (c)	c* (c*)	- (-)	c (c)	c (c)
Window.Name	c (c)	c (c)	c (c)	c (c)	c (c)
IE-UserData	- (-)	- (-)	- (-)	- (-)	- (-)
Global-Storage	- (-)	c* (c*)	- (-)	- (-)	- (-)
Database-Storage	- (-)	- (-)	- (-)	c (c)	- (-)
WebCache	c (c)	c (c)	- (-)	- (-)	c (c)
CSS-History-Cookie	- (c)	- (c)	- (-)	- (-)	c (c)
Flash-Cookie	- (c)	- (c)	- (c)	c (c)	c (c)
Java-Cache-Cookie	c (c)	c (c)	c (c)	c (c)	c (c)
Google Gears DB	-(-)	- (-)	- (c)	- (-)	(-)

Die Tabelle (oben) zeigt das Ergebnis nach dem Löschen der Cookies. Bei keinem der Browser wird darauf hingewiesen, dass die Cookies erst nach einem Neustart des Browsers gelöscht werden. Das Verhalten der Browser entspricht in diesem Fall nicht den Erwartungen des Nutzers. Ein Neustart bewirkt, dass die meisten der Cookies verschwinden. Übrig bleiben in diesem Fall die Cookies, die sich im Cache oder im Verlauf des Browser einnisten können.

Allein der Private Modus konnte bei allen Browsern überzeugen. Der Java-Cache-Cookie, der erst in dieser Arbeit entwickelt wurde, konnte jedoch nur im Firefox, im Safari und im Opera abgewehrt werden.

Evaluation Jondonym & Vidalia:

Diese beiden Anonymisierungspakete wurden im Rahmen der Arbeit genauer untersucht. Dabei fiel auf, dass beide den gleichen Ansatz zur Anonymisierung der Nutzer verfolgen. Beide verbergen den Nutzer durch ein besonders häufig genutztes Identifikationsprofil, wodurch die Unterscheidung verschiedener Nutzer erschwert werden soll.

Das führt jedoch dazu, dass Nutzer plötzlich mit einem UserAgent aus den USA surfen und auf deutschen Internetseiten eher auffallen als in der Masse unterzutauchen.

Ausblick:

In Zukunft werden zunehmend auch biometrische Merkmale bei der Nutzeridentifikation eine Rolle spielen. Das soziale Netzwerk Facebook setzt schon jetzt Gesichtserkennungsverfahren ein, um Nutzer auf Bildern zu identifizieren. Auch Verfahren, die die Mausbewegungen des Nutzers analysieren, um diesen zu erkennen, werden in Zukunft eine Rolle spielen.

Die neue WebGL-Schnittstelle, die in einigen Browsern bereits integriert ist, könnte in Zukunft ein wertvolles Merkmal für eine ähnlichkeitsbasierte Nutzeridentifikation werden.

Fazit:

Die im Rahmen dieser Arbeit untersuchten Cookies demonstrieren die Vielzahl von Möglichkeiten, die existieren, um einen Nutzer eindeutig einem Profil zuzuordnen. Durch die Entwicklung des Java-Cache-Cookies im Rahmen dieser Arbeit ist die Annahme, dass es in Zukunft noch weitere Cookiearten geben wird, durchaus berechtig. Sehr wahrscheinlich ist, dass die neue Möglichkeit Flash-Cookies zu löschen umgangen wird, indem ein Flash-Cache-Cookie nach dem Vorbild des Java-Cache-Cookies entwickelt wird.

Die Abwehrmaßnahmen der Browser weisen im Moment noch Lücken gegen die eine oder andere Cookieart auf, jedoch sind bei allen Browserherstellern Bestrebungen zu erkennen, diese Lücken zu schließen.

Der Private Modus ist eine vorbildliche Lösung zum Schutz der Privatsphäre des Nutzers, da dieser einfach durch einen Knopfdruck zum anonymen Surfen umschalten kann. Ähnlichkeitsbasierte Verfahren werden momentan jedoch von keinem Browserhersteller beachtet. Diese stellen eine Möglichkeit der Profilzuordnung dar, die auch dann funktioniert, wenn eindeutige Verfahren nicht anwendbar sind.

Zusammenfassend kann man sagen, dass es möglich ist das Recht auf informationelle Selbstbestimmung wahrzunehmen. Um dieses Ziel jedoch zu erreichen, ist der Nutzer gezwungen geeignete Maßnahmen zum Schutz der eigenen Daten durchzuführen. Die Hauptverantwortung für die im Internet preisgegebenen Daten trägt der Nutzer somit selbst.

Betreuer: Prof. Dr. phil Dipl. Math. Barbara Wiesner
Dipl.-Inform. Ingo Boersch

Abgabe: 24.08.2011