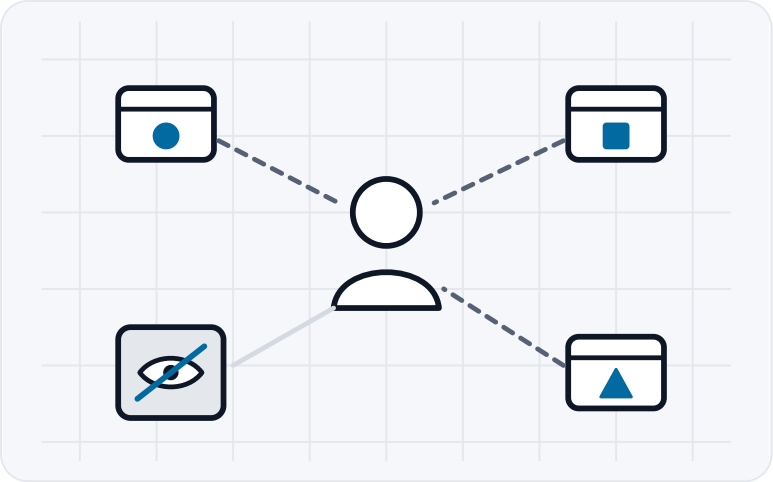
A=SIT Technology
E-Government Innovationszentrum

4. September 2024 Überblick · Privatsphäre & Krypto · Digitale Identität · 17 Min. Lesezeit

Privatsphäre-bewahrende Identifikatoren

Stefan More



Autorinnen und Autoren	Stefan More
Quelle	https://technology2.a-sit.at/veroeffentlichungen/privatsphaere-bewahrende-identifikatoren/
Hinweis	Maschinell übersetzt.

Kernaussagen

- Globale Kennungen erlauben es Diensteanbietern und Identitätsanbietern, Aktivitäten einer Person über Plattformen hinweg zu verknüpfen.
- Herkömmliche PPID-Ableitung und Unobservability gegenüber dem Identitätsanbieter schließen einander aus.
- OPRF-basierte Verfahren wie BISON erreichen Unlinkability und Unobservability ohne dauerhaften Zustand auf dem Gerät.
- Für Unobservability, Unlinkability und Sybil-Resistenz gleichzeitig ist ein protokollkundiger Browser nötig.
- MPC-basierte Kennungen könnten flexibler als OPRF sein, werfen aber die Frage auf, wer die Knoten betreibt.

In föderierten Authentifizierungsprotokollen (z. B. OpenID Connect) werden Benutzer:innen oft eindeutige Kennungen zugewiesen, um von Diensten (SPs) identifiziert zu werden. Dies wird beispielsweise verwendet, um eine zurückkehrende Benutzerin bei einem Onlinedienst erneut zu authentifizieren. Ein daraus resultierender Nachteil ist, dass das Verhalten der Benutzerin über mehrere SPs hinweg verknüpft werden kann, was die Privatsphäre der Benutzerin beeinträchtigt (Linkability). Alternativ kann der Identitätsanbieter (IDP) für jeden SP eine separate Kennung ableiten (SP-bezogene Kennung). Bei den aktuellen Methoden geschieht dies jedoch immer direkt durch den IDP, wodurch es dem IDP ermöglicht wird, zu beobachten, bei welchen SPs sich eine Benutzerin authentifiziert (Beobachtbarkeit, Observability).

Das Ziel einer datenschutzfreundlichen Kennung ist es daher, sowohl Unlinkability als auch Unbeobachtbarkeit (Unobservability) zu ermöglichen. In einem System, das solche Kennungen verwendet, sind Benutzer:innen zwischen verschiedenen Diensten nicht nachverfolgbar, und Identitätsanbieter können das Verhalten der Benutzer:innen nicht beobachten.

Dieses Projekt zielt darauf ab, verschiedene Methoden für datenschutzfreundliche, SP-bezogene Benutzerkennungen zu analysieren und zu diskutieren. Systeme/Methoden, die auf kryptografischen Techniken wie OPRF, MPC und ZKP basieren, dienen als Beispiele.

1. Einleitung

In der heutigen digitalen Welt gewinnt der Schutz der Privatsphäre zunehmend an Bedeutung, insbesondere bei Online-Identitäten. Föderierte Authentifizierungssysteme wie OpenID Connect (OIDC) ermöglichen es Benutzer:innen, sich mit einem einzigen, von einem Identitätsanbieter (IdP) verwalteten Konto bei mehreren Diensten anzumelden.

Das vereinfacht zwar den Zugang für Benutzer:innen, schafft aber auch Herausforderungen für die Privatsphäre. Eines der Hauptprobleme ist die Verknüpfbarkeit (Linkability): Die Aktivitäten einer Benutzerin oder eines Benutzers können über verschiedene Diensteanbieter (SPs) hinweg anhand einer eindeutigen Kennung nachverfolgt werden. Zusätzlich kann der IdP beobachten, auf welche Dienste Benutzer:innen zugreifen, was zu Bedenken hinsichtlich der Beobachtbarkeit (Observability) führt.

Datenschutzfreundliche Kennungen sollen diese Probleme lösen, indem sie Lösungen bieten, die die Privatsphäre der Benutzer:innen stärken. Konkret sollen sie sicherstellen, dass Benutzer:innen zwischen verschiedenen Diensten nicht nachverfolgt werden können (Unlinkability) und dass Identitätsanbieter die Aktivitäten der Benutzer:innen nicht überwachen können (Unobservability). Mithilfe fortgeschrittener kryptografischer Techniken wie Oblivious Pseudo-Random Functions (OPRF), Multi-Party Computation (MPC) und Zero-Knowledge Proofs (ZKP) lassen sich Systeme schaffen, in denen sich Benutzer:innen sicher authentifizieren können und dabei ihre Privatsphäre wahren.

Als Beispiel soll eine Benutzerin dienen, die sich über denselben IdP bei mehreren Onlinediensten anmeldet, etwa bei einer Social-Media-Plattform, einem Onlineshop und einer Nachrichtenwebsite. In einem typischen föderierten Authentifizierungssystem wird bei allen Diensten dieselbe eindeutige Kennung verwendet. Dadurch können sowohl die Diensteanbieter als auch der IdP das Verhalten der Benutzerin nachverfolgen und ihre Aktivitäten über diese verschiedenen Plattformen hinweg verknüpfen. Dabei entsteht ein detailliertes Profil der Benutzerin, das ihre Privatsphäre beeinträchtigt.

Bei einem System mit datenschutzfreundlichen Kennungen erhielte hingegen jeder Diensteanbieter für dieselbe Benutzerin eine eigene, SP-bezogene Kennung. So ist sichergestellt, dass weder die Diensteanbieter noch der IdP die Aktivitäten der Benutzerin über Plattformen hinweg verknüpfen können, was die Privatsphäre erheblich verbessert. Meldet sich die Benutzerin beispielsweise bei der Social-Media-Plattform, dem Onlineshop und der Nachrichtenwebsite an, bleibt jede Interaktion für sich isoliert, und keine Stelle kann sich ein vollständiges Bild ihres Online-Verhaltens machen.

1.1. Beispiel

Beispiel für Optionen der föderierten Authentifizierung („Social Login“) bei einem Diensteanbieter.

Angenommen, Alice meldet sich häufig mit ihrem Google-Konto bei verschiedenen Websites an. Google ist ein gängiger Identitätsanbieter (IdP) in föderierten Authentifizierungssystemen. Sie nutzt ihr Google-Login für einen Streamingdienst, eine Onlineshopping-Plattform und eine Fitness-App. Im herkömmlichen föderierten Modell erhält jeder dieser Dienste – nennen wir sie StreamNow, ShopEasy und FitTrack – von Google dieselbe eindeutige Kennung für Alice, z. B. ihre GMail-Adresse.

Mit dieser Kennung können StreamNow, ShopEasy und FitTrack Alice bei jeder Anmeldung wiedererkennen. Da dieselbe Kennung bei allen drei Plattformen verwendet wird, können diese Dienste jedoch Informationen über Alices Verhalten zusammenführen, etwa welche Sendungen sie ansieht, welche Produkte sie kauft und wie sie trainiert. Schlimmer noch: Google als IdP sieht ebenfalls, wann und wo sich Alice anmeldet, und erhält so ein vollständiges Bild ihrer Online-Aktivitäten.

In einem datenschutzfreundlichen System würde Google hingegen für jeden Dienst eine eigene Kennung erzeugen, wenn sich Alice bei StreamNow, ShopEasy und FitTrack anmeldet. StreamNow würde eine Kennung erhalten, ShopEasy eine andere und FitTrack eine dritte. So bleibt Alices Aktivität auf diesen Plattformen voneinander getrennt – kein Dienst kann ihr Verhalten über Plattformen hinweg verknüpfen. Bietet das System zusätzlichen Schutz, kann Google nicht nachverfolgen, welche Dienste sie nutzt; das gilt allerdings nicht für alle Ansätze datenschutzfreundlicher Kennungen (siehe unten). Bei vollständiger Anwendung bieten solche Systeme Alice ein deutlich höheres Maß an Privatsphäre.



Allgemeiner dienen Pseudonyme als Mechanismus zur Stärkung der Privatsphäre, da sie es Benutzer:innen ermöglichen, mit Diensteanbietern zu interagieren, ohne ihre wahre Identität oder eine dauerhafte Kennung preiszugeben.

Ein gängiger Ansatz für Pseudonymität sind Pairwise Pseudonymous Identifiers. Dabei wird für jede paarweise Beziehung zwischen einer Benutzerin oder einem Benutzer und einem Diensteanbieter eine eigene Kennung erzeugt. Das bedeutet: Jede Benutzerin und jeder Benutzer erzeugt für jeden Diensteanbieter, mit dem interagiert wird, ein neues Pseudonym und verwendet das Pseudonym eines SP

nicht bei einem anderen SP wieder. Diese Kennungen dienen bei den Interaktionen als temporäre oder einmalige Pseudonyme und verhindern, dass SPs Benutzer:innen über verschiedene Dienste oder Sitzungen hinweg nachverfolgen (Linkability).

2. Föderierte Authentifizierung im Web

Der Prozess der Authentifizierung („Login“) im Web umfasst typischerweise die folgenden Akteure [1]:

Benutzer:in Die Benutzer:in ist typischerweise eine Person, die sich bei einem Dienst anmelden möchte. Allgemeiner ist die Benutzer:in eine Partei, deren primäres Ziel es ist, ihre Identität gegenüber einem Dienst sicher und effizient nachzuweisen, idealerweise auf datenschutzfreundliche Weise. Im Web nutzt die Benutzer:in dazu einen Webbrowser, um auf verschiedene Dienste zuzugreifen.

Diensteanbieter (Service Provider, SP) Der Diensteanbieter ist eine Partei, die einen Dienst betreibt. Er verlangt von der Benutzer:in eine Authentifizierung als Teil einer Autorisierungsentscheidung für den Zugriff auf einen Dienst oder eine Ressource. Es gibt viele verschiedene Diensteanbieter, die ganz unterschiedliche Dienste anbieten können. Ein Internetforum, ein virtuelles Ladengeschäft, die Terminsoftware eines Gesundheitsdienstleisters, die Onlinebanking-Anwendung einer Bank und das Webportal einer Behörde sind sehr verschiedene Arten von Diensteanbietern mit unterschiedlichen Anforderungen.

Fast alle Diensteanbieter müssen Benutzer:innen (wieder)erkennen. Besucht eine Benutzer:in ein Forum, sollte sie zuvor gesendete Nachrichten bearbeiten können. Besucht sie ein Ladengeschäft, sollte sie ihre offenen Bestellungen sehen können. Besucht sie ihren Gesundheitsdienstleister, sollte sie ihre Termine sehen und absagen können, und so weiter. Es ist daher ein Authentifizierungsprozess irgendeiner Art nötig, der zu einer vertrauenswürdigen, dauerhaften Kennung für die Benutzer:in führt.

Je nach Methode muss der Server für jede Benutzer:in dauerhafte Informationen – Passwörter, geteilte Geheimnisse oder Wiederherstellungscodes – sicher speichern. Umgekehrt müssen sich Benutzer:innen dann auch viele verschiedene Zugangsdaten für verschiedene Diensteanbieter merken.

Da dies eine mühsame Aufgabe ist, entstand der Wunsch, den eigentlichen Authentifizierungsprozess an eine spezialisierte dritte Partei zu delegieren. Das resultierende Modell heißt föderierte Authentifizierung.

Identitätsanbieter (Identity Provider, IdP) Der Identitätsanbieter ist eine Partei, der zugetraut wird, die Benutzerauthentifizierung durchzuführen. Er entscheidet über die zu verwendenden Authentifizierungsfaktoren und speichert alle Informationen, die nötig sind, um die Benutzer:in mit diesen Faktoren (erneut) zu authentifizieren, etwa Passwörter sowie verschiedene zweite Faktoren oder Passkeys.

Der Identitätsanbieter weist jedem Benutzerkonto dann eine „globale“ Kennung zu, die nur im Kontext dieses bestimmten IdP eindeutig ist. Manche Identitätsanbieter verwenden wirklich globale Kennungen wie E-Mail-Adressen, aber nicht alle.

Je nach Kontext kann der Besitz eines Kontos bei einem bestimmten Identitätsanbieter besondere Implikationen haben, etwa die Mitgliedschaft in einer bestimmten Organisation, die vom Identitätsanbieter überprüft wird. In diesem Fall delegiert der Diensteanbieter nicht nur die Authentifizierung, sondern implizit auch die Überprüfung dieser Eigenschaften an den Identitätsanbieter.

Dieser Ansatz erlaubt es Diensteanbietern, Benutzer:innen sicher erneut zu authentifizieren, ohne den eigentlichen Authentifizierungs- (und Identifizierungs-)Prozess selbst abwickeln zu müssen. Er ist im heutigen Web weit verbreitet, z. B. in Form von Social Logins („Login with Google“, „Login with Facebook“) oder einigen E-ID-Systemen wie ID Austria.

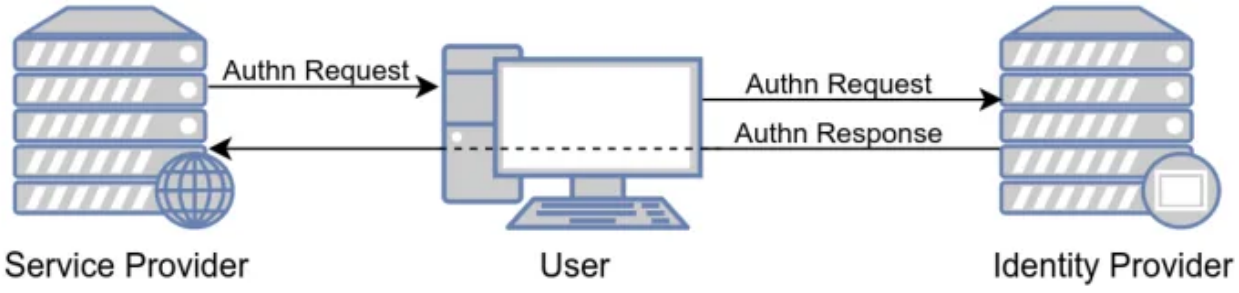
Grundlegender Ablauf der Authentifizierung bei föderierter Authentifizierung.

2.1. Federated Credential Management (FedCM)

Herkömmliche OpenID-Connect-Authentifizierungsabläufe, wie oben beschrieben, setzen ein protokollunkundiges Benutzergerät voraus und umgehen diese Einschränkung. Dies gelingt durch den Einsatz allgemeiner Technologien wie HTTP-Weiterleitungen (Redirects) und HTML-Formularübermittlungen, die den Browser der Benutzer:in dazu bringen, undurchsichtige Informationen zwischen logisch nicht zusammengehörigen Servern weiterzuleiten.

Diese Technologien stehen seit Kurzem in der Kritik. Dieser Schritt folgt darauf, dass Webbrowser die herkömmlichen Wege einschränken, das Benutzerverhalten über verschiedene Web-Ursprünge hinweg zu verfolgen, etwa Drittanbieter-Cookies, die von der Online-Werbebranche lange missbraucht wurden. Als Reaktion darauf haben Angreifer auf die Privatsphäre neue Wege erkundet, Benutzerinteraktionen einander zuzuordnen; dazu gehört die Verwendung nachrichtentragender Weiterleitungen (Bounce Redirects), mit denen der Browser sich gegenüber einem Werbe-Tracker identifiziert. Browserhersteller schlagen deshalb vor, solche zustandsbehafteten Weiterleitungen zu beschränken. Das verbessert zwar die Privatsphäre der Benutzer:in im Hinblick auf Tracking, würde aber genau die Technologie beseitigen, von der OpenID Connect und ähnliche Anwendungsfälle abhängen.

Der aktuelle W3C-Vorschlag zur Lösung dieses Dilemmas ist die Federated Credential Management (FedCM) API [7]. FedCM sieht vor, dass Browser Authentifizierungsprozesse konzeptionell kennen und eine aktive Rolle bei der Aushandlung der Benutzerauthentifizierung übernehmen. Das bedeutet einen Paradigmenwechsel weg vom herkömmlichen, OIDC-unkundigen Benutzergerät. Ein zusätzlicher Vorteil OIDC-kundiger Browser ist, dass der Browser nun eine aktive Rolle bei der Authentifizierung spielen kann, statt Daten lediglich zwischen SP und IdP weiterzureichen.



Mehrere Ansätze nutzen diesen Vorteil, um die Privatsphäre der Benutzer:in bei der föderierten Authentifizierung zu verbessern. [1]

3. Funktionale Anforderungen und Datenschutzziele

3.1. Funktionale Anforderungen

Die folgenden Anforderungen sind die Grundlage eines sicheren Authentifizierungssystems [1]:

- Soundness (Korrektheit): Schließt eine Benutzer:in einen Authentifizierungsprozess erfolgreich ab, soll dies garantieren, dass sie sich beim Identitätsanbieter erfolgreich als die zum resultierenden Pseudonym gehörende Benutzer:in authentifiziert hat. Mit anderen Worten: Eine böswillige Benutzer:in soll sich nicht – ohne Mitwirkung des Identitätsanbieters – gegenüber einem Diensteanbieter als harmlose Benutzer:in ausgeben können.

- Validity (Gültigkeit): Der Identitätsanbieter soll ein Benutzerkonto sperren oder löschen können, sodass die Benutzer:in sich nicht mehr authentifizieren kann.

- Kein clientseitiger Zustand: Geheime Informationen auf dem Gerät der Benutzer:in zu speichern, kann für den Protokollentwurf sehr bequem sein, bringt aber viele praktische Herausforderungen mit sich. Geräte können kaputtgehen oder verloren werden, was zum unwiederbringlichen Verlust aller gespeicherten kryptografischen Informationen führt. Solche Informationen zu sichern und dabei wirklich privat zu halten, ist eine offene Herausforderung. Benutzer:innen möchten sich möglicherweise auch auf einem gemeinsam genutzten Computer oder dem Gerät einer Freundin oder eines Freundes anmelden. Die Übertragung kryptografischen Schlüsselmaterials auf ein solches Gerät ist oft komplex, und das Schlüsselmaterial einem möglicherweise nicht vertrauenswürdigen Gerät auszusetzen, ist unerwünscht. Wir konzentrieren uns daher auf Ansätze, die keinen dauerhaften Zustand auf dem Benutzergerät benötigen. Jeder auf dem Benutzergerät gehaltene Zustand soll auf einen einzelnen Authentifizierungsprozess beschränkt sein und nach Abschluss des Prozesses verworfen werden können. Das entspricht bestehenden, in der Praxis eingesetzten Protokollen und vermeidet Komplikationen bei der Integration unseres Verfahrens.

3.2. Datenschutzziele:

Im Kontext der (föderierten) Authentifizierung im Web sind die folgenden Datenschutzziele relevant [3,6]:

- Unlinkability (Unverknüpfbarkeit): Benutzer:innen erwarten, dass sie eine Zugangsberechtigung oder mehrere Zugangsberechtigungen (Credentials) bei verschiedenen Diensteanbietern verwenden können, ohne ihre Privatsphäre zu verringern. Konkret sollen die beteiligten Diensteanbieter die Interaktionen nicht miteinander verknüpfen können, was ein Profiling der Benutzer:innen ermöglichen würde. Bei anonymen oder pseudonymen Interaktionen möchten Benutzer:innen dasselbe Credential zweimal (z. B. für eine Altersverifikation) in zwei Interaktionen verwenden. Erwartet wird, dass der Dienst die beiden Interaktionen nicht miteinander verknüpfen kann, also nicht weiß, dass dieselbe Benutzer:in zweimal mit ihm interagiert hat (Multishow-Unlinkability).

- Unobservability (Unbeobachtbarkeit): Ebenso erwarten Benutzer:innen, dass keine Partei außer den unmittelbar an einer Interaktion Beteiligten (also Benutzer:in und Diensteanbieter) erfährt, dass die Interaktion stattfindet. Konkret erwarten Benutzer:innen, dass der Aussteller oder eine andere Stelle (z. B. der Staat) nicht erfährt, bei welchem SP die Benutzer:in ihre Wallet verwendet, oder auch nur, dass die Benutzer:in ihre Wallet verwendet.[1]

4. Ansätze

4.1. Nicht-pseudonyme Kennungen

Die Grundlage der föderierten Authentifizierung im Web ist eine Kennung ohne zusätzliche Datenschutzmaßnahmen: Die Benutzer:in meldet sich beim Identitätsanbieter (IdP) an, und der IdP gibt eine statische Kennung an den Diensteanbieter (SP) zurück, z. B. die E-Mail-Adresse der Benutzer:in. Diese Kennung ist dieselbe, unabhängig davon, auf welchen SP die Benutzer:in zugreifen möchte, und wird daher als globale Kennung bezeichnet.

Eine globale Kennung ist eine eindeutige Kennung, die ein Identitätsanbieter einer Benutzer:in zuweist und die bei allen Diensteanbietern gleich bleibt. Diese Art von Kennung wird typischerweise in föderierten Authentifizierungssystemen verwendet, in denen sich Benutzer:innen mit denselben, von einem einzigen IdP verwalteten Zugangsdaten bei mehreren Diensten anmelden. Die globale Kennung erlaubt es SPs, die Benutzer:in plattformübergreifend zu erkennen, und sorgt für ein nahtloses Erlebnis, bei dem die Identität der Benutzer:in unabhängig vom genutzten Dienst einheitlich anerkannt wird.

So bequem globale Kennungen sind, bergen sie jedoch erhebliche Risiken für die Privatsphäre. Da dieselbe Kennung bei mehreren Diensten verwendet wird, können SPs und andere Stellen die Aktivitäten einer Benutzer:in plattformübergreifend verknüpfen, etwas über ihr Verhalten erfahren und ihre Privatsphäre beeinträchtigen.

4.2. Pairwise Pseudonymous Identifier (PPID)

Um Benutzer:innen besseren Datenschutz zu bieten, erlaubt die OpenID-Connect-Spezifikation [8] die Verwendung von Pairwise Pseudonymous Identifiers (PPIDs) [1]. Dabei leitet der Identitätsanbieter für eine bestimmte Zielgruppe (Audience) eine pseudonyme Kennung ab und setzt sie an die Stelle der globalen Kennung der Benutzer:in. So erzeugt etwa der OpenID-Anbieter „Sign in with Apple“ von Apple eine unterschiedliche pseudonyme Kennung, die auf dem zugehörigen Apple Developer Account des Diensteanbieters beruht [9].

Die konkrete Ableitungsmethode und die Abgrenzung der Audience bleiben dem Identitätsanbieter überlassen. Üblicherweise wird ein Hashwert aus der globalen Kennung der Benutzer:in, verkettet mit einer Audience-Kennung, berechnet und als PPID der Benutzer:in verwendet [8, Abschnitt 8.1]. Die herkömmliche PPID-Ableitung geht davon aus, dass der IdP weiß, bei welcher Audience sich die Benutzer:in authentifiziert. Sie verankert damit die Notwendigkeit, dass der IdP bei jedem Login sensible Zuordnungsdaten erfährt. Herkömmliche PPID-Ableitung und Unobservability gegenüber dem Identitätsanbieter schließen einander aus.

4.3. Bereichsspezifische Personenkennzeichen (ssPIN, bPK)

Eine weitere Form datenschutzfreundlicher Kennungen sind Sector-specific Personal Identifiers (ssPINs), in Österreich auch Bereichsspezifisches Personenkennzeichen (bPK) genannt. Sie sind eine Methode zur Stärkung der Privatsphäre, mit der für verschiedene Bereiche oder Diensteanbieter eindeutige Benutzerkennungen erzeugt werden.



Beispiel für ein Login mit der österreichischen E-ID. Statt einer globalen Kennung wird nur die bereichsspezifische Kennung („bPK“) für den Bereich „Steuern und Abgaben (SA)“ an den SP der Finanzverwaltung übermittelt.

Die Idee hinter der ssPIN ist, jedem Bereich oder Diensteanbieter für dieselbe Benutzer:in eine eigene Kennung zu geben und so sicherzustellen, dass Aktivitäten in verschiedenen Bereichen nicht verknüpft werden können. Eine Benutzer:in, die Gesundheitsdienste nutzt, hätte beispielsweise eine Kennung, bei der Nutzung von Bankdiensten würde dagegen eine völlig andere Kennung entstehen, was bereichsübergreifendes Tracking unmöglich macht. Diese Kennungen werden typischerweise mit kryptografischen Methoden aus der primären Kennung der Benutzer:in (etwa einer nationalen ID) abgeleitet. So bleibt die Identität der Benutzer:in innerhalb eines Bereichs konsistent, ist aber über Bereiche hinweg nicht verknüpfbar. Durch diese Isolierung der Benutzeridentitäten stärkt die ssPIN die Unlinkability und bewahrt die Privatsphäre über verschiedene Diensteklassen hinweg.

4.4. OPRF-basierte Blind Identifiers (BISON, OPPID)

Ein weiterer Vorschlag für datenschutzfreundliche Authentifizierung in föderierten Systemen ist das BISON-Protokoll [1]. BISON steht für Blind Identification with Stateless scOped pseudoNyms und adressiert einige der größten Datenschutzrisiken herkömmlicher Systeme wie OpenID Connect.

Das zentrale Merkmal von BISON ist, dass es eine Pseudonymableitung ermöglicht, die sowohl unverknüpfbar als auch unbeobachtbar ist. Das bedeutet: Identitätsanbieter (IdPs) können nicht nachverfolgen, mit welchen Diensteanbietern (SPs) eine Benutzer:in interagiert, und Diensteanbieter können die Aktivitäten der Benutzer:in nicht über verschiedene Dienste hinweg verknüpfen. BISON erreicht dies mit einer kryptografischen Technik namens Oblivious Pseudorandom Functions (OPRF). Sie erlaubt es dem IdP, bei der Erzeugung von Pseudonymen mitzuwirken, ohne den genauen SP zu kennen, für den das Pseudonym erzeugt wird. Dies geschieht mit Hilfe des Browsers (vgl. Abschnitt 4.5 unten) und bietet damit einen deutlich besseren Schutz der Privatsphäre als herkömmliche Pairwise Pseudonymous Identifiers, bei denen der IdP die Interaktionen der Benutzer:innen weiterhin beobachten kann.

Ein weiterer wichtiger Aspekt von BISON ist, dass es keine langfristige Zustandsspeicherung auf dem Gerät der Benutzer:in erfordert. Dadurch ist es leichtgewichtig und lässt sich leicht in bestehende Systeme integrieren. Zusätzlich führt BISON eine kleine Erweiterung ein, mit der sich PPID-Pseudonyme über das BISON-Protokoll erzeugen lassen. Das macht es zu einer praxistauglichen Lösung für reale Datenschutzprobleme, zumal es starke Datenschutzgarantien bietet, ohne den Authentifizierungsprozess zu verkomplizieren.

Zwei weitere Vorschläge, die BISON ähneln, sind OPPID (Oblivious Pairwise Pseudonyms) [2] und UPPRESSO (Untraceable and Unlinkable Privacy-PREserving Single Sign-On Services) [10]. Mithilfe von Zero-Knowledge Proofs (ZKPs) erreicht OPPID außerdem SP Authentication, d. h., das System kann einschränken, welche Stellen als SP fungieren dürfen.

4.5. Exkurs: Clientseitige Änderungen

Ein entscheidender Vorteil von OpenID Connect, der zu seiner weiten Verbreitung beigetragen hat, ist, dass das Benutzergerät von seiner Existenz nichts wissen muss. Client-augmented OIDC bietet diesen Vorteil nicht. Es verlangt, dass der Webbrowser eine aktive Rolle übernimmt: die Kennung (den Scope) des SP überprüfen, Blinding-Operationen durchführen und die Zustimmung der Benutzer:in über integrierte Dialogfenster einholen. Dazu muss der Webbrowser von der Existenz von OIDC wissen.

Ein protokollkundiger Browser ist jedoch nötig, um Unobservability, Unlinkability und Sybil-Resistenz gleichzeitig zu erreichen [1].

Um das zu zeigen, sei die Authentifizierung als Zwei-Parteien-Protokoll zwischen einem Diensteanbieter und einem Identitätsanbieter betrachtet, bei dem ein protokollunkundiger Browser als Kommunikationskanal dient. Angenommen, es gibt zwei verschiedene Diensteanbieter, A und B, und eine Benutzer:in X.

Der Identitätsanbieter erhält nun einen Authentifizierungsrequest. Dieser Request darf es dem Identitätsanbieter nicht erlauben festzustellen, ob er von A oder B stammt; wäre das möglich, ginge die Unobservability verloren. Nach einer Benutzerinteraktion zur Authentifizierung von X werden Daten an den Urheber zurückgegeben, die in der Kennung P resultieren. Da Sybil-Resistenz vorausgesetzt wird, muss P über Authentifizierungsrequests hinweg stabil sein. Man bedenke jedoch, dass der Request sowohl von A als auch von B stammen könnte; daher kann er keine Unlinkability bieten.

4.6. MPC-basierte und ZKP-basierte Kennungen

Eine mögliche Forschungsrichtung sind MPC-basierte Kennungen. Bei MPC (Multi-Party Computation) berechnet eine Gruppe von Knoten (mindestens zwei) gemeinsam eine Funktion. Damit lassen sich Protokolle umsetzen, bei denen mehrere Parteien Daten in die Funktion einspeisen und sichergestellt ist, dass keine Partei die Eingaben der anderen erfährt. Am Ende erfährt einer oder erfahren alle Knoten das Ergebnis der Berechnung.

Damit lässt sich ein Protokoll zur Kennungsableitung konstruieren, das dem oben im Abschnitt zu BISON beschriebenen ähnelt: Zuerst authentifiziert sich die Benutzer:in beim IdP. Dann initiiert sie ein Login bei einem SP und gibt die Scope-Kennung des SP in die MPC-Berechnung ein. Auf der anderen Seite gibt der IdP die Konto-Kennung der Benutzer:in in die Berechnung ein. Als Ergebnis erfährt die Benutzer:in die SP-spezifische Kennung (das Login ist unverknüpfbar). Zugleich erfährt die Benutzer:in ihre globale Kennung nicht und, vor allem, der IdP erfährt die Kennung des SP nicht (die Anfrage bleibt unbeobachtbar). Eine

Herausforderung bei MPC-Protokollen ist die Frage, wer die Knoten pflegt und betreibt (Anreizstruktur, Vertrauensaufbau usw.). Ein möglicher Vorteil von MPC gegenüber OPRF-basierten Protokollen könnte sein, dass MPC-Protokolle flexibler sind und zusätzlich zur Kennungsberechnung weitere Prüfungen durchführen können (z. B. Prüfungen von Zugriffsrichtlinien [11]).

Eine weitere besondere Form datenschutzfreundlicher Kennungen sind kryptografische Pseudonyme auf Basis von Zero-Knowledge Proofs (ZKP), z. B. ABC4Trust-Pseudonyme [5]. Sie ermöglichen eine anonyme (oder pseudonyme) und unverknüpfbare Nutzung eines Dienstes, erlauben es aber einer vertrauenswürdigen Stelle (Inspector) oder einer Gruppe von Stellen, die Identität hinter einem Pseudonym später aufzudecken.

Mit auf dem Gerät der Benutzer:in (z. B. Telefon, Smartcard oder Browser) gespeichertem kryptografischem (Schlüssel-)Material können Benutzer:innen ihre eigenen Pseudonyme lokal ableiten. Dadurch ist der IdP nicht direkt am Authentifizierungsprozess beteiligt und kann das Verhalten der Benutzer:in nicht beobachten. Da eine Benutzer:in zudem für jeden Diensteanbieter ein neues Pseudonym erzeugen kann, besteht keine Verknüpfbarkeit zwischen den verschiedenen Authentifizierungsprozessen bei verschiedenen Diensteanbietern. Dieser Ansatz erfordert jedoch lokalen Zustand im Bereich der Benutzer:in und eignet sich daher nicht für alle Anwendungsfälle, insbesondere nicht im Kontext der Webauthentifizierung.

5. Schlussfolgerungen

In diesem Bericht wurden verschiedene Ansätze zur Umsetzung datenschutzfreundlicher Kennungen in föderierten Authentifizierungssystemen untersucht, mit Schwerpunkt auf Unlinkability und Unobservability. Herkömmliche föderierte Systeme, die globale Kennungen verwenden, bergen erhebliche Datenschutzrisiken, da sowohl Diensteanbieter (SPs) als auch Identitätsanbieter (IdPs) die Aktivitäten einer Benutzer:in über Plattformen hinweg verknüpfen können. Um diesen Bedenken zu begegnen, wurden mehrere fortgeschrittene kryptografische Techniken vorgeschlagen und diskutiert.

Mit wachsender Nachfrage nach datenschutzfreundlicher Authentifizierung dürften diese Techniken breitere Verbreitung finden, insbesondere wenn sich Webstandards weiterentwickeln, um stärker datenschutzorientierte Modelle zu unterstützen. Künftige Arbeiten sollten die Abwägungen zwischen Privatsphäre, Benutzbarkeit und Performance sowie das Potenzial für eine reale Umsetzung in verschiedenen Sektoren untersuchen, von sozialen Medien bis zu E-Government-Diensten.

Referenzen

[*] OpenAI. ChatGPT (4o) [Large language model] was used to edit parts of the text.

[1] Jakob Heher, Stefan More, Lena Heimberger, BISON: Blind Identification with Stateless scOped pseudoNyms. arXiv pre-print, 2024.

[2] Maximilian Kroschewski, Anja Lehmann, Cavit Özbay, OPPID: Single Sign-On with Oblivious Pairwise Pseudonyms. Cryptology ePrint Archive, 2024.

[3] Stefan More, Jakob Heher, Edona Fasllija, Maximilian Mathie, Service Provider Accreditation: Enabling and Enforcing Privacy-by-Design in Credential-based Authentication Systems. Proceedings of the 19th International Conference on Availability, Reliability and Security, 2024.

[4] Stefan More, Trust and Privacy in a Heterogeneous World. PhD Thesis, Graz University of Technology, 2023.

[5] K. Rannenberg, J. Camenisch, A. Sabouri, Attribute-based credentials for trust. Springer Cham, 2015.

[6] A. Pfitzmann, M. Hansen. Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology (Version 0.28), 2006.

[7] Nicolás Peña Moreno, Sam Goto, Federated Credential Management API (Working Draft). https://www.w3.org/TR/fedcm, August 2024.

[8] Nat Sakimura, John Bradley, Michael B. Jones, Breno de Medeiros, Chuck Mortimore, OpenID Connect Core 1.0. https://openid.net/specs/openid-connect-core-1_0.html, 2014.

[9] Apple, Authenticating users with “Sign in with Apple”. https://developer.apple.com/documentation/sign_in_with_apple/sign_in_with_apple_rest_api/authenticating_users_with_sign_in_with_apple, retrieved Sept 2023.

[10] Chengqian Guo et al., UPPRESSO: Untraceable and Unlinkable Privacy-PREserving Single Sign-On Services. arXiv pre-print, 2022.

[11] Stefan More, Lukas Alber, YOU SHALL NOT COMPUTE on my Data: Access Policies for Privacy Preserving Data Marketplaces and an Implementation for a Distributed Market using MPC. Proceedings of the 17th International Conference on Availability, Reliability and Security, 2022.

---

1. Wir argumentieren in Abschnitt 4.5, dass protokollkundige Clients nicht nur vorteilhaft, sondern notwendig sind, um Unlinkability und Unobservability gleichzeitig zu gewährleisten. ↩