Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Entwurf vom
9. September 2024

Bachelorarbeit

Privatsphärewahrendes Anreiz- und Betrugserkennungssystem im Datentreuhandmodell

vorgelegt von

Knut Hoffmeister
Matrikelnummer 7509085
Studiengang Software System Entwicklung

MIN-Fakultät
Fachbereich Informatik

eingereicht am 9. September 2024

Betreuer: Kevin Röbert

Inhaltsverzeichnis

Abstract

1 Einleitung

In der heutigen Zeit wird der fachgerechte Umgang mit Daten jeglicher Form, zunehmend wichtiger. Viele der großen Player wie Facebook oder Google, machen ihr Hauptgeschäft mit dem Verwenden von Nutzerdaten zu gewerblichen Zwecken, wie bspw. targeted Ads [**facebookad**] [**googlead**]. Und obwohl das Misstrauen eines Nutzers, solch eines großes Unternehmens gegenüber berechtigt ist, benötigen diese die gesammelten Daten auch, um neue Technologien zu entwickeln. Jedoch hat der Nutzer, der diese Daten generiert, meist keine Einsicht darüber, wer seine Daten verwendet und wofür diese zum Einsatz kommen. Ein potentieller Lösungsansatz für dieses Problem, ist die Verwendung von Datentreuhändersystemen. In einem Datentreuhändersystem, kann ein Endnutzer der Daten generiert zum einen seine Daten verschlüsselt bei einem Treuhänder lagert. Zum anderen ist der Treuhänder ein Mediator zwischen dem Endnutzer und einem Unternehmen, welches an den Daten interessiert ist. Möchte nun ein Unternehmen die Daten für ihre Zwecke verwendet, so kann dieses bei dem Treuhänder Daten anfordern. Daraufhin kontaktiert der Treuhänder den Nutzer, von dem die Daten stammen und fragt diesen, nach seiner Zustimmung. So kann der Nutzer genau einsehen, wer seine Daten verwenden möchte und kann unerwünschte Benutzung unterbinden. Das Ziel liegt hierbei vor allem darauf den Nutzer und seine Privatsphäre so gut wie möglich zu schützen.

In dieser Arbeit sollen, um die Motivation zur Benutzung eines Datentreuhänder zu erhöhen, zwei Forschungsfragen beantwortet werden: Wie kann ein privatsphäreschützender Anreiz zur Benutzung eines Datentreuhändermodells geschaffen werden? Und: Wie kann dieser Anreiz gegen Missbrauch geschützt werden? Aktuell ist der einzige Anreiz zur Benutzung eines Datentreuhändermodells für den Nutzer, der Schutz der Privatsphäre durch die Kommunikation über den Treuhänder. Um einen weiteren Anreiz zu bieten, soll hier ein Ansatz vorgestellt werden, welcher den Datengebenden für seine zur Verfügung gestellten Daten, angemessen entlohnt. Dies geschieht durch eine Transaktion von dem Datennutzenden an den Datengebenden, die trotz des Austausches von Zahlungsinformationen, die Privatsphäre des Datengebenden und dessen Identität so weit wie möglich schützt. Hierfür könnte bspw. eine leicht erweiterte Version des GNU Taler Zahlungssystem verwendet werden **??**, um den Zahlungsverkehr zwischen den Parteien zu ermöglichen. Um dieses System vor Missbrauch zu schützen, wird ein Reputationssystem eingesetzt, um zu verhindern, dass ein Datengebender wertlose Daten oder qualitativ niedrig wertige Daten, in großer Masse bereitstellen kann, um so das System auszunutzen. Es soll zum einen, gleich wie der Bezahlvorgang, die Identität des Datengebenden schützen. Gleichzeitig soll es den Datennutzenden, der die Bewertung ausstellt, davon abhalten, das System durch mehrfaches Bewerten auszunutzen. Auch Betrugsversuche des Datengebenden sind nicht zu vernachlässigen, wie bspw. die Neuanmeldung eines Nutzers mit schlechter Reputation, um diesen Wert zurückzusetzen.

Dafür wird in dieser Arbeit der aktuelle Stand der Forschung in dem vorliegenden Kontext evaluiert und geprüft was bereits anwendbar und wo noch Lücken zur gewünschten Verwendung bestehen. Mit den gesammelten Forschungsergebnissen werden daraufhin mehrere Konzepte präsentiert, um das gerade genannte Ziel in das Treuhändermodell zu integrieren. Diese Konzepte werden des weiteren in ein bestehendes Treuhändersystem eingebaut, um konkrete Vergleichsgrundlagen zu erhalten. Bei dem hierfür vorgesehenen System, handelt es sich um das Tresor-Projekt der Universität Hamburg [**TRESOR**].

2 Hintergrund

Die zunehmende Digitalisierung unseres Alltags ist unabstreitbar [**dt-digitalisierung-stat**]. Viele der täglichen Aktivitäten hängen stark mit ihr zusammen. Sei es den schnellsten Weg zu Arbeit finden mit Google Maps, das kontaktlose Bezahlen an der Kasse mit GooglePay oder Applepay, die Benutzung von Social Media zur Unterhalten oder der Onlinehandel über Anbieter wie Amazon. Sie alle liefern Komfort der durch die zunehmende Verwendung von Computern ermöglicht wird, welche im Hintergrund Unmengen an Daten für ihre Berechnungen verwenden. Diese Daten stammen meist von den Benutzern selbst. Bspw. die Standortdaten für die Berechnung potentieller Staus im Straßennetz [**dt-googlemaps-staus**] oder die Unterhaltungsinteressen basierend auf Watchtime von bestimmten Social Media Inhalten.
Aufgrund des ständig wachsenden Markts für neue Digitaltechnologien ist auch die Nachfrage nach Daten über die letzten Jahr in Höhe gestiegen. Über das letzte Jahrzehnt haben Daten, Öl als wertvollste Ressource abgelöst [**dt-falck2020rohstoff**]. Während 2008 die vier weltweit wertvollsten Unternehmen Ölkonzerne waren, waren 2018 bereits die 7 wertvollsten Unternehmen Internet- und Technologiefirmen.

2.1 Herkömmliche Datenkommunikation

Unter anbetracht des hohen Wertes von Daten sind viele Unternehmen verständlicherweise sehr zurückhaltend was den Austausch betrifft. Schließlich beutet eine eigene Sammlung von Daten ein potentielles Verkaufsgut. Laut einer durch die Bundesregierung aufgegriffenen Studie von Fedkenhauer et al. geben zwar viele der befragten Unternehmen an, Aktivitäten im Bereich des Datenaustausches zu betreiben, allerdings umfasst das in 83% der Fällen den Austausch von Daten mit Kundinnen und Kunden. 53% der Unternehmen teilen ihre Daten mit Lieferantinnen und Lieferanten. Ein noch kleinerer Anteil von 21% teilen ihre Daten mit Unternehmen aus der gleichen oder anderen Branchen und nur 15% teilen sie mit Wettbewerbern [**dt-bundesregierung2021datenstrategie**].

Aus der kapitalgetriebenen Sicht eines Unternehmens besitzt das Teilen der eigenen Daten keinen direkten Nutzen. Da ein Unternehmen seine eigenen Gewinnmaximierung anstrebt, ist das Teilen von Daten eher ein Nachteil, da fremde Unternehmen mit den selbst gesammelten Daten ihre Produkte qualitativ erweitern können. Dadurch werden entweder andere Wettbewerben oder branchenfremde Unternehmen in ihrem Marktwert gefördert, was zu dem sinken des eigenen Marktwertes führt.
Diese protektive Herangehensweise kann allerdings auch der Gewinnmaximierung im Weg stehen. Im Fall von einem direkten Tausch an Daten können beide Parteien einen Profit aus der Interaktionen erwirtschaften. Die Bundesregierung selbst schreib in [**dt-bundesregierung2021datenstrategie**] das kaum Datenkooperationen zwischen staatlichen und wirtschaftlichen Akteurinnen bestehen, obwohl die staatlich gesammelten Daten eine Grundlage für wirtschaftliche Innovation sein könnten. Im Gegenzug könnten die Daten von Unternehmen dem Staat bei der Sicherstellung seines Versorgungsauftrages, der Daseinsvororge und der Wahrung öffentlicher Schutzgüter helfen. Dies ist eine optimale Situation für die Verwendung eines Datentreuhänders.

2.2 Datentreuhänder

Ein Datentreuhänder ist ein neutraler vertrauenswürdiger Vermittler von Daten eines Datengebenden zu einem Datennutzenden. Er hat selbst kein kommerzielles Interesse an der Verwertung der Daten und agiert vergleichbar zu einem Notar strickt für den Datengeber. Seine Hauptaufgaben umfassen meist

die Kontrolle von Zugriffsrechten, das Kontrollieren von Einhaltung der Datenschutzrichtlinien, sowie das verschlüsseln oder anonymisieren von Datenbeständen. In speziellen Fällen kann ein Datentreuhänder auch die Auswertung von Daten vornehmen. [**dt-bundesregierung2021datenstrategie**][**dt-richter2020ddvtalk**]

Da wie bereits angeführt viele Unternehmen die Weiterleitung ihrer Daten vermeiden, ist unter der Annahme eines etablierten Datentreuhänders ein deutliche größerer Datenbestand verfügbar. Bereits heute - vor einer großen Etablierung von Datentreuhänder - verspricht das Konzept einige gesellschaftliche Vorteile: [**dt-richter2020ddvtalk**]

1. Durch sie können Datenbestände besser vernetzt werden und Zusammenhänge hergestellt werden, welche zu Innovationen führen.
2. Der Wettbewerb unter Firmen wird gestärkt da mit besser zugänglichen Daten auch kleinere Unternehmen die kein Datenmonopol besitzen ihre Produkte aufwerten können.
3. Der individuelle Endnutzer erhält mehr Kontrolle und Transparenz über die Speicherung und Verwendung seiner Daten.

Allerdings ist das Verständnis eines Datentreuhänders nicht eindeutig. Jürgen Kühling beschreibt den Datentreuhänder als “ein schillerndes Wesen. Jeder kennt ihn, jeder setzt ganz eigene Hofnungen in ihn – und jeder stellt sich doch etwas anderes unter ihm vor“ [**dt-kuhling2021datentreuhander**]. Obwohl die Technologie eines Datentreuhänders bereits seit Jahren existiert und verwendet wird [**dt-hardinges2018data**] ist es nicht gelungen eine konkrete allumfassende Definition für die Technologie zu finden.

2.2.1 Definition

Die allgemeingültigste Definition stammt aus der Rechtswissenschaft und bezieht sich auf die Treuhandschaft im Allgemeinen. *[Treuhandschaften sind ein] Rechtsverhältnis, bei dem eine natürliche oder juristische Person (Treugeber) einer zweiten Person (Treuhänder) ein Recht unter der Bedingung überträgt, von diesem Recht nicht zum eigenen Vorteil Gebrauch zu machen. [...] Gemeinsames Charakteristikum ist die Uneigennützigkeit und Vertrauenswürdigkeit bei der Wahrnehmung fremder Interessen bzw. die uneigennützige Ausübung von amtlichen Befugnissen"*[**dt-beeck2013treuhandschaft**]. Diese Treuhandschaft wurde in der Vergangenheit verwendet, um bspw. Ländereien zu verwalten und im Namen einer lokalen Gemeinschaft Entscheidung zu treffen. [**dt-hardinges2018data**]. In dem Fall eines Datentreuhänders bedeutet dies konkret, dass eine Datengebende Person beim bereitstellen ihrer Daten den Datentreuhänder dazu ermächtigt, über die Daten zu verfügen. Darunter fällt unter anderen auch die Weitergabe der Daten, solange die im Sinne des Datengebers ist.

2.2.2 Verschieden Modelle

Insgesamt lassen sich alle bis zum heutigen Zeitpunkt in Betrieb genommenen oder geplanten Datentreuhandsysteme wie folgt kategorisieren. Zum einen besteht die Einteilung in Customer to Business oder Business to Business Systeme und zum anderen die risikobasierte Einteilung nach Zentralen/Dezentraler Datenspeicherung und Freiwilliger/Verpflichtender Nutzung. (siehe **??**)

Die Unterscheidung zwischen Customer to Business und Business to Business Datentreuhändern basiert ausschließlich auf den interagierenden Parteien. Im Falle einer B2B Interaktion kommunizieren zwei Unternehmen die vorhandenen Daten miteinander. In diesem Fall kommt es häufig vor, dass eines der Unternehmen durch eine staatliche Behörde dargestellt wird. Bei den gespeicherten Daten handelt es sich meist um personenbezogene Daten. Aufgrund dessen befassen sich B2B Datentreuhänder häufig mit der Pseudonymisierung und der Verwaltung der bereitgestellten Daten. Sie sind unter anderem im Gesundheitswesen viel vertreten. [**dt-blankertz2020datentreuhandmodelle**]

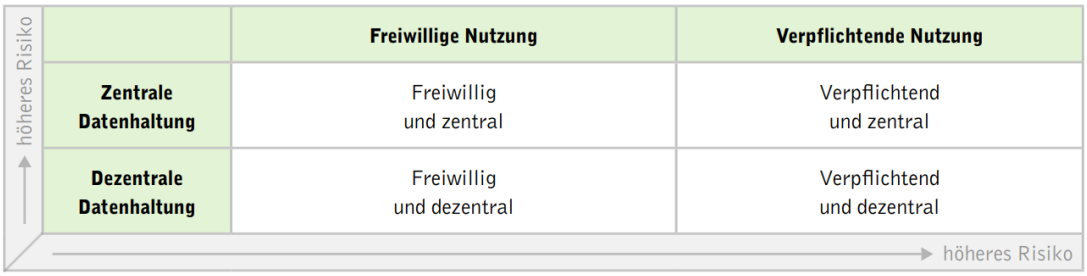

Abbildung 2.1: Risikobasierte Unterscheidung von Datentreuhandmodellen [**dt-blankertz2021neue**]

C2B Systeme umfassen solche, bei denen einen Endnutzer die Daten generiert und diese an ein Unternehmen zur weiteren Benutzung freigibt. Ihre Aufgabe ist hauptsächlich die Unterstützung des Nutzers bei der gerechten Weiterverarbeitung seiner Daten. [**dt-blankertz2020datentreuhandmodelle**] Hier sind dementsprechend die Pseudonymisierung der Daten sowie die Einhaltung von Datenschutzrechtlinien und einheitlicher Standarts die Hauptziele des Datentreuhänders.

Des weiteren lassen sich Datentreuhandsysteme anhand ihrer Datenspeicherung sowie Nutzung kategorisieren und Risikotechnisch bewerten. Die zentrale Speicherung der Daten bietet einige Vorteile für den Treuhänder. Sie ermöglicht es Daten vor zu verarbeiten, zu analysieren und Datenverarbeitende von dem direkten Zugang der Daten auszuschließen. Allerdings birgt die zentrale Datenspeicherung ein enormes Risiko der Datensicherheit, da hier ein single point of failure entsteht. Bei einem Angriff auf einen solchen Datentreuhänder fällt es einem Angreifer somit leichter eine große Menge an Daten zu stehlen. Bei einer dezentralen Speicherung werden die Daten durch den Treuhänder pseudonymisiert oder ganz anonymisiert und bei dem Datengeber gelagert. Aus diesem Grund bietet eine dezentrale Datenspeicherung mehr Sicherheit vor Diebstahl. Sie zieht allerdings auch eine eingeschränkte Verarbeitung und Analyse der Daten mit sich und erhöht die Komplexität der Verwaltung.

Es besteht eine weitere Unterteilung in Datentreuhandsysteme, dessen Benutzung freiwillig ist oder verpflichtend ist. Dabei fällt der größte Anteil an Systeme unter die freiwillige Benutzung. Es gibt jedoch auf Szenarien in denen die Verwendung einer Datentreuhand verpflichtend ist. Ein Beispiel hierfür wäre das Krebs- und Transplantationsregister aus dem medizinischen Bereich. Das Risiko steigt bei verpflichtenden Systemen, da sie meist einen wichtigen Bestandteil der Kommunikation ausmachen der nicht umgangen werden kann. Somit sind die potentiellen Schäden die bei einem Angriff entstehen können höher als in einem freiwilligen System.

2.3 Anwendungsfälle

Das mögliche Spektrum an Anwendungsfällen ist denkbar breit. In beinahe jedem Bereich der eine große Menge an Daten benötigt oder verwaltet ist die Verwendung eines Datentreuhänders angedacht. Beispiele dafür sind:

- **Patientendaten** Der Datentreuhänder sorgt für eine pseudonymisierung von Patientendaten zur Bereitstellung an Forschungseinrichtungen. Hierbei behält der Patient die Hoheheit über seine Daten und kannst selbst entscheiden mit wem seine Daten geteilt werden.
- **Autonomes Fahren** Beim autonomen Fahren werden enorm viele Daten generiert, welche seit 2017 per Gesetz gespeichert werden müssen [**dt-bundesdruckereiDatentreuhÃd'nder**]. Leider ist die Zugehörigkeit der Daten rechtlich weder dem Autohersteller noch dem Autoinhaber zuzuschreiben [**dt-richter2020ddvtalk**]. An dieser Stelle kann ein Datentreuhänder die Kommunikation erleichtern und exklusiven Zugang von bspw. Versicherungen oder Automobilkonzernen ausschließen.

- **E-Government** Durch die Verwendung eines Datentreuhänders können bei der behördlichen Verwaltung von Bürgerdaten große Fortschritte erzielt werden. Unter anderem müssen so notwendige Informationen aus anderen Registern für Verwaltungsvorgänge nicht mehr vom Bürger bereitgestellt werden. Der Bürger gibt lediglich seine Einwilligung zum Abruf der Daten zu Beginn des Verwaltungsvorgangs. Auf diese Weise können Verwaltungsvorgänge erheblich effizienter Ablaufen.
- **KI-Datenpools** Eines der größten Probleme bei Entwicklung von KI-Software ist der Zugang zu einer ausreichend großen Menge an Trainingsdaten. Ein Datentreuhänder kann solche Daten die zur Verwendung für KI Training freigegeben wurden sammeln und pseudonymisiert an mehrere Interessierte verteilen. Dadurch entsteht ein einheitlicher Zugang zu Trainingsdaten der gleichzeitig nur freigegebene Daten beinhaltet und so rechtlichen Streit über das Copyright aus dem Weg geht.
- **Industrie** In der Industrie besteht eine hohe Abhängigkeit von Warenbewegungen, seien es in Lieferketten, Logistik oder Handel. Durch die Verwendung eines Datentreuhänders können diese Informationen pseudonymisiert an Warenempfänger weitergegeben werden. Vor allem in diesem Bereich besteht durch die Zusammentragung an Lieferinformationen in Kombination mit Algorithmen der Graphentheorie ein großes Innovationspotential.
- **PIMS** Personal Information Management Systeme befassen sich grundsätzlich mit der Wahrung von personenbezogenen Daten und bietet ihren Nutzern mehr Kontrolle über diese. Datentreuhänder sind für solche Systeme vor allem von Vorteil da sie im Umgang mit personenbezogenen Daten dem Nutzer wieder die Kontrollen über seine Daten zurückgeben.

[**dt-blankertz2021regulierung**][**dt-blankertz2021neue**][**dt-bundesdruckereiDatentreuhÃd'nder**]

Der erste Datentreuhänder entstand einer bereits im Jahre 2006 in England. Seit dem das Thema Datentreuhänder in den letzten Jahren immer mehr zum Trend wurde [**dt-richter2020ddvtalk**]. Die "UK Biobank" ist eine biomedizinische Datenbank die sowohl medizinische Daten als auch biologische Proben von einer halben Millionen Teilnehmer aus Großbritanien speichert[**dt-hardinges2018data**]. Die gespeicherten Daten sind pseudonymisiert und werden ausschließlich an Forscher im Feld der Medizin weitergegeben, was die UK Biobank zu einem Paradebeispiel für einen Patientendatentreuhänder macht.

2.4 Bestehende Anreize

Generell sei gesagt, dass die Verwendung eines Datentreuhänders keine direkten Nachteile mit sich bringt. Wenn Daten ohnehin geteilt werden oder werden müssen, so behält der Nutzer bei der Speicherung der Daten über einen Datentreuhänder mehr Kontrolle über die Verwendung seiner geteilten Daten verglichen mit dem direkten Teilen mit Unternehmen. Durch den Datentreuhänder wird ihm ermöglicht zu einem beliebigen Zeitpunkt das weitere Teilen seiner Informationen einzustellen. Vermutlich teilen deswegen Privatpersonen ihre Daten lieber mit Datentreuhändern als auf direktem Weg. [**dt-tresor24study**]

Allerdings gibt es nur wenige Vorteile die die freiwillige Benutzung eines Datentreuhänders reizvoll machen. Bei freiwilligen C2B Datentreuhändern, hängt die Entscheidung für oder gegen die Nutzung des Treuhänders beim Nutzer. Es ist also an ihm abzuwägen ob die Verwendung ausreichend Vorteile liefert. Im Fall von PIMS Treuhänder wird die Verwendung von manchen als erstrebenswert angesehen, da der Nutzer mehr Kontrolle über die Verbreitung von persönlichen Daten erhält und selbst entscheiden kann, mit wem diese geteilt werden. Die Erkenntnissen von Jai et al. [**dt-jai2016privacy**] zeigen hingegen, dass vor allem jüngere Erwachsene weniger Wert auf den Schutz ihrer persönlichen Daten legen.

Allerdings gibt es keinen direkten Anreiz zur Verwendung einer solchen Software oder der Weitergabe der persönlichen Daten, da nur das Business von den gesammelten Daten einen Mehrwert hat. Der

Nutzer der seine Daten freigibt erhält keine Kompensation in irgendeiner Form. Folglich kann es für einen Datentreuhänder dessen Verwendung freiwillig ist mühsam sein neue Nutzer zu gewinnen und die Technologie als solche auszubauen.

3 Anforderungen und verwandte Arbeiten

3.1 Anforderungen

Da eine Lösung zur Verwendung durch einen Datentreuhänder konzeptioniert ist, decken sich die Anforderungen an eine mögliche Lösung stark mit denen eines üblichen Datentreuhänders. Im Grunde kann zwischen funktionalen und nichtfunktionalen Anforderungen unterschieden werden.

3.1.1 Funktionale Anforderungen

1. Ein Bezahlsystem ermöglicht es einem Datengebenden Geld für seine Daten zu erhalten.
2. Die Präsenz eines Reputationswertes gibt den Datennutzenden vor Erwerb der Daten eine Einschätzung über deren Qualität.
3. Nach Abschluss der Transaktion kann ein Datennutzender den entsprechenden Datengebenden aufgrund der Qualität der übermittelten Daten bewerten.
4. Ein Datengebender muss eine Bewertung seiner Daten ermöglichen.
5. Ein Datennutzender kann pro Austausch nur genau eine Bewertung für einen Datengebenden abgeben. Mehrfache Bewertungen ist nur im Fall von mehrfachem Erwerb möglich.

3.1.2 Nicht funktionale Anforderungen

1. **Anonymität** Die Identität des Datengebenden darf durch den Austausch von Zahlungsmitteln oder durch dessen Reputation nicht nachverfolgbar sein.
2. **Zeitsensitivität** Der Bezahlvorgang muss in vernachlässigbarer Zeit geschehen.
3. **Skalierbarkeit** Die Rechenzeit des Systems soll bei linear steigender Menge an Datengebenden und Datennutzenden auch mit linearem Zeitaufwand zunehmen.
4. **Vertraulichkeit** Die kommunizierten Daten dürfen nicht durch unbefugte Dritte ausgelesen werden können.
5. **Integrität** Die kommunizierten Daten dürfen nicht unbemerkt durch unbefugte Dritte verändert werden.

3.2 GNU Taler

Das im Paper “Enabling Secure Web Payments with GNU Taler“ von J. Burdges et al. eingeführte Zahlungssystem GNU Taler, ist ein elektronisches Online-Zahlungssystem, das Datenschutz für Customer und Mechanismen zur steuerlichen Nachverfolgung für Merchants bietet [**gnu-burdges2016enabling**]. Es verwendet einen Exchangeservice, um Münzen mithilfe von blinden Signaturen zwischen Nutzern und Händlern zu transferieren. Im Folgenden werden diese Münzen als Taler bezeichnet, um Verwirrung zu vermeiden. Das System basiert auf 4 übergeordneten Rollen, dessen Interaktion grob in Abbildung **??** skizziert ist. Der Customer möchte ein Gut oder eine Dienstleistung bei dem Merchant erwerben und

bezahlt diesen dafür mit Talern, welcher er beim Exchange erworben hat. Der Merchant kann die erhaltenen Taler wieder beim Exchange für herkömmliche Währungen eintauschen. Ein Auditor überprüft währenddessen die Liquidität des Exchange, um sicherzustellen, dass dieser auch bei Datenverlust von Taler noch in der Lage ist allen Beteiligten, Auszahlungen zu ermöglichen.

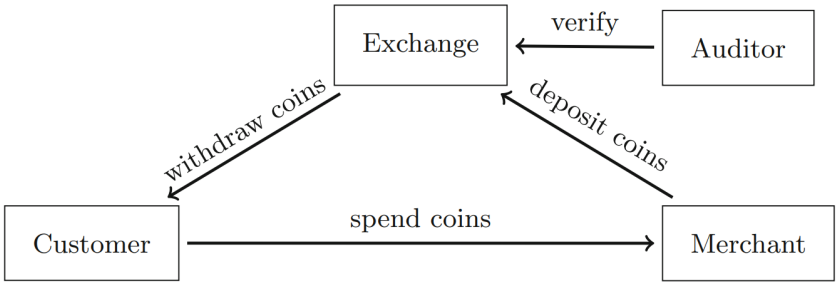

Abbildung 3.1: Grundlegender Ablauf des GNU Taler Systems [**gnu-burdges2016enabling**]

Taler abheben. Damit ein Customer Geld auf sein Wallet laden kann, muss er sich zuerst bei seiner Bank anmelden. Sollte die Bank GNU Taler native unterstützen, so kann der Customer in einem Formular eine Summe auswählen, die er in Taler übertragen möchte und einen Exchangeservice, über welchen der Tausch abgewickelt wird. Nachdem der Customer die Transaktion bestätigt hat, wird die ausgewählte Summe transferiert, der Exchangeservice signiert die equivalente Summe an Talern und überträgt diese in das Wallet des Customers.

Taler ausgeben. Gehen wir von der Situation aus, dass eine Customer bei einem Merchant (hier ein Onlineshop), etwas erwerben möchte. Nach der Auswahl des Produktes und GNU Taler als Zahlungsmittel, erstellt der Merchant einen Zahlungsvertrag, der Details wie den Gesamtpreis, mögliche offene Umwandlungsgebühren und akzeptierte Exchangeservices beinhaltet und sendet diesen an das Wallet des Customers. Wenn der Customer daraufhin die Zahlung übermittelt, so leitet der Merchant die erhaltenen Taler direkt an den Exchange weiter. Wenn der Exchange den Eingang bestätigt, so kann der Merchant dem Customer die Transaktion bestätigen und der Kauf ist somit abgeschlossen.

In dem Treuhändermodell eignet sich GNU Taler aber nur teilweise als Zahlungssystem. Der Datengebende stellt dabei den Merchant da, der seine Daten als digitales Gut anbietet. Der Datennutzende stellt hier die Rolle des Customers dar und möchte diese Daten erwerben. Sollte der Datengebende der Anfrage des Datennutzende zustimmen, so würde er einen Zahlungsvertrag formulieren, um den Anspruch auf seine Vergütung zu formalisieren. Allerdings soll der Datengebende im Datentreuhändermodell so gut wie möglich vor Informationsgewinnung geschützt werden. Bei dem von GNU Taler vorgeschlagenen Prozess wird jedoch die Identität des Datengebenden nachverfolgbar, während der Datennutzenden anonym bleibt. Zusätzlich besteht eine direkte Kommunikation zwischen Datengebenden und Datennutzenden, was weitere schützenswert Information über die Identität des Datengebenden preis gibt.

3.3 Privacy Pass

Ein VPN bietet einige Vorteile in der regulären Kommunikation über das Internet. Bspw. erhöht er die Anonymität des Nutzers da dessen private IP Adresse so geschützt bleibt. Allerdings gibt es auch Nachteile die häufig übersehen werden. Einer davon ergibt sich daraus, dass wenn der Internetverkehr von hunderten Benutzern des VPNs über die gleiche IP Adresse verteilt wird, dann genügt ein bösartiger Nutzer der die IP Adresse für einen Angriff oder o.ä. nutzt, um der IP einen schlechten Ruf bei populären Content Delivery Networks wie Cloudflare zu verschaffen. Ein Content Delivery Network (kurz CDN) ist dafür zuständig häufig angefragt Websiten wie Google.com oder Netflix.com in seinem Cache aufzubewahren und so die Zugriffszeit welche durch physikalisch große Distanzen zwischen Nutzer und

Server entsteht zu verkürzen [**pp-cdn**]. Zusätzlich liefert ein CDN ein Menge an Sicherheitsfunktionen, wie unter anderem IP Adressen mit schlechtem Ruf ein CAPTCHA präsentieren, um Botzugriffe zu verhindern. Daraus resultiert dass ein regulärer Nutzers eines VPNs erheblich mehr CAPTCHAs lösen muss, als ein Nutzer der keinen VPN verwendet.

3.3.1 Umgehen vom CAPTCHAs

Privacy Pass ist eine Browsererweiterung die es ermöglicht, vorab ein CAPTCHA zu lösen und damit eine Menge an Token zu erhalten. Solange ein Nutzer über mindestens einen Token verfügt, so kann er das nächste Mal, wenn ein CDN ihn aufgrund eines schlechten IP Rufwertes zum lösen eines CAPTCHAs auffordert, anstatt einen Token einlösen und kann die Aufgabe so überspringen. Dadurch erhöht sich die Nutzerfreundlichkeit unter der Verwendung eines VPNs da die Anzahl an zu lösenden CAPTCHAs rapide sinkt. [**pp-davidson2018privacy**]

3.3.2 Funktionsweise

Das System ist aufgeteilt in eine sogenannte Signierphase und Einlösephase. Die Signierphase startet nachdem der Nutzer erfolgreich ein CAPTCHA gelöst hat. Sie ist dafür zuständig dem Nutzer eine Anzahl an Coins auszustellen die durch den Server signiert sind. Die Einlösephase beginnt wenn ein CDN ein CAPTCHA für den Zugang zu einem Webinhalt fordert. Bei ihr wird einer der gespeicherten signierten Token eingetauscht um die Lösung des CAPTCHAs zu überspringen.

Signierphase Erstellen und signieren von Token

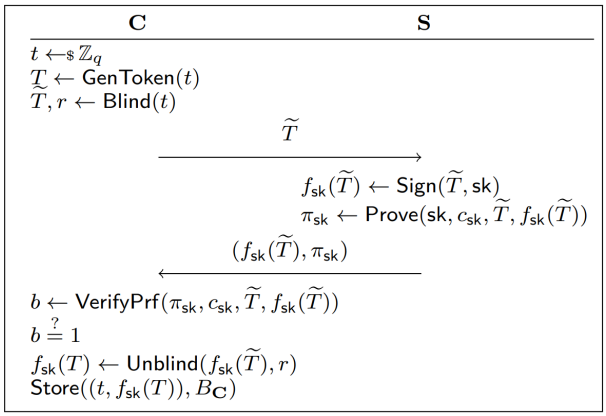

Abbildung 3.2: Signierphase von Privacy Pass [**pp-davidson2018privacy**]

Zuerst erstellt der Nutzer (C - Client) einen Tokenseed t mit $t \in_R \mathbb{Z}_q$. Daraus generiert er Token T bspw. mit einer Hashfunktion und blindet diesen mir r wie in **??** beschrieben, um $\widetilde{T}$ zu erhalten. Dieser geblindete Token wird nun an der Server gesendet, damit dieser in signieren kann. Beim Server angekommen beginnt dieser damit, den Token mit seinem privaten Schlüsselt sk zu signieren. Anschließend erstellt er einen sogenannten Batch Discrete Log Equivalence Proof, um dem Nutzer zu beweisen, dass er nicht für jeden Nutzer einen eigenen privaten Schlüssel verwendet um ihn über längere Zeit zu deanonymisieren. Der signierte Token und der BDLEQ werden wieder an den Nutzer gesendet, dieser prüft die Korrektheit des Beweises, unblindet den signierten Token und speichert ihn für spätere Verwendung.

Einlösephase Signierten Token einlösen

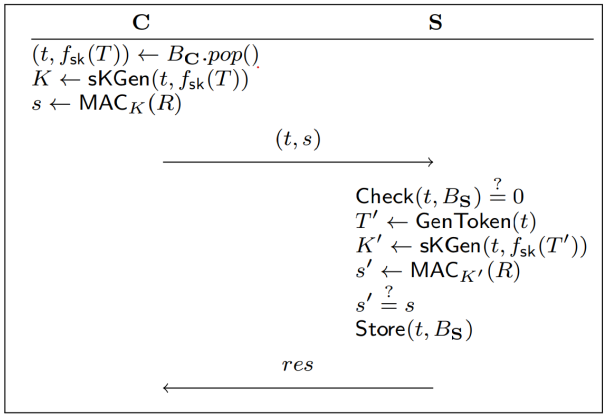

Abbildung 3.3: Einlösephase von Privacy Pass [**pp-davidson2018privacy**]

Der Nutzer beginnt damit einen gespeicherten signierten Token auszulesen und einen anfragenabhängigen Wert R zu berechnen. Dieser könnte einfach die Domain der Anfrage sein. Er generiert einen shared key K aus dem Tokenseed und dem signierten Token und verschlüsselt R mit K als key. Daraufhin sendet er das Tupel aus Tokenseed t und $s \leftarrow MAC_K(R)$ mit shared key verschlüsseltem R an den Server. Der Server prüft ob t bereits für eine vorherige Anfrage verwendet wurde. Falls dies nicht der Fall ist, berechnet er auf Grundlage von t alle Schritte des Nutzers und erhält so ein s'. Sollte $s' = s$ gelten, dann ist der Token valide, das CAPTCHA kann übersprungen werden und der Nutzer erhält Zugriff auf die angefragt Webressource.

3.4 Privacy-Preserving Reputation Management

In ihrem Paper beschreiben R Petrlic et al. ein Reputationssystem, um Nutzern den Dienst eines Dienstleisters anonym bewerten zu lassen und so anderen Nutzern eine Einschätzung über die Qualität der Dienstleistung zu geben [**petrlic2014privacy**]. Es werden 3 Rollen charakterisiert. Ein Reputation Provider RP, der sich um das Verwalten der verschiedenen Reputationswerte kümmert. Eine Menge an Service Providern SP, welche einen Dienst anbieten, der durch Nutzer bewertet werden soll. Und zuletzt eine Menge an Nutzern U, die die Dienstleistung der Service Provider bewerten, wobei keiner der Nutzer gleichzeitig ein Service Provider SP sein kann, $U \cap SP = \varnothing$.

Der im Paper beschriebene Ablauf des Bewertens, eines $sp \in SP$ durch einen Nutzer $u \in U$, umfasst grob die Erstellung eines Schemas für einen Bewertungsvektors durch sp, welcher von u später verwendet wird, um den sp zu bewerten. Nach dem sp dieser Vektor mit dem RP kommuniziert wurde, kann u eine Dienstleistung von sp in Anspruch nehmen. Dieser antwortet, zusätzlich zur Erbringung der Dienstleistung, mit einer Beschreibung des Bewertungsvektors, einem Schlüsselpaar und einem Token. Dank der Beschreibung kann der Nutzer selbst einen Bewertungsvektor erstellen, in welchen er seine Bewertung der Dienstleistung einbaut. Im Anschluss wird dieser mit dem Schlüsselpaar verschlüsselt, signiert und zusammen mit dem signierten Token, an den RP gesendet. Der Token dient hier zur Erkennung von doppelten Bewertungen.

Die Berechnung der Reputation ist in Zeitslots eingeteilt, damit ein sp nicht in der Lage ist einen alten Reputationswert anzugeben. Wenn nun ein Nutzer u den Reputationswert eines sp's einsehen möchte und der Wert für den gewünschten Zeitraum noch nicht bestimmt wurde, so muss dieser zwischen RP und sp berechnet werden. Hierfür bestimmt RP die Summe aller verschlüsselten Bewertungsvektoren für den Zeitraum und sendet diese, zusammen mit weiteren Prüfwerten an sp. Dieser kann die übermittelten Werte prüfen und auf eine Blacklist hinzufügen, um Replay Attacken von RP auszuschließen. Wenn die Überprüfung gelingt, signiert sp die Summe der Vektoren zusammen mit einer ID und verifiziert somit den neuen Reputationswert an RP.

Bei einer Anfrage des Reputationswertes von sp durch u, schickt der sp eine Reihe an Werte zu u. Diese erlauben es u, den Bewertungsvektor zu interpretieren und zu prüfen, dass der übermittelte

Werte sowohl aktuell ist, als auch nicht beeinflusst wurde. Das Konzept von Petrlic et al. bietet unter anderem Schutz vor:

1. **Whitewashing**, bei welchem sich ein sp mit schlechter Reputation als neuer sp ausgeben kann, um somit den Wert zurücksetzen kann.

2. **Transaction-independent Ratings**, bei welchen ein Nutzer die Dienstleistung bewerten kann, obwohl er besagt Dienstleistung nicht in Anspruch genommen hat.

3. **Sybil Attacks**. Ein Nutzer bewertet eine Dienstleistung unter mehreren Identitäten und täuscht so seine Meinung als Gruppenmeinung vor.

4. **Delta Analysis**. Eine teilweise Deanonymisierung des Nutzers durch vergleichen von gesammelten Bewertungen in unterschiedlichen Zeitabständen.

In der vorliegenden Situation kann das Konzept zu großen Teilen verwendet werden. Die Rollen können unter anderen Namen genauso verwendet werden. In diesem Fall wird der Service Provider zum Datengebenden, der Nutzer wird zum Datennutzenden und die zu bewertende Dienstleistung wird zu den übermittelten Daten. Allerdings ist das Konzept darauf ausgelegt, dass zu Beginn eine Kommunikation zwischen dem Datengebenden und Datennutzenden besteht, so dass der Reputationswert direkt vom Datennutzenden abgefragt wird. An dieser Stelle muss es also etwas abgewandelt werden, da zu Beginn keine direkte Kommunikation zwischen einem Datennutzender und einem Datengebenden besteht. Erst nachdem der Treuhänder dem Datennutzenden eine List an Datengebenden mitgeteilt hat, kann eine direkte Kommunikation entstehen. Stattdessen soll der Austausch des Reputationswertes direkt über den Treuhänder geschehen.

3.5 Blinde Signaturen

Neue rl müssen hier noch eingefügt werden. Texte nochmal überarbeiten -> besser formulieren und technischer beschreiben.

Das von David Chaum im Jahre 1983 veröffentlichtes Paper “Blind signatures for untraceable payments“ beschreibt den theoretischen Ansatz, dass ein Nutzer eine verifizierbare Signatur für eine Nachricht erhält, ohne dass der Unterzeichner den Inhalt der Nachricht kennt. [**chaum1983blind**]. Ein Beispiel hierfür wäre ein Zahlungsdienstleister, der eine neuen Coin in dem Umlauf bringen möchte. Dafür sendet er seinen Reputationswert und den unkenntlich gemachten Coin an eine Zentrale Institution. Wenn diese den Reputationswert als hoch genug ansieht und somit den Zahlungsdienstleister als vertrauenswürdig erkennt, so kann sie den Coin blind signieren und an den Dienstleister zurückschicken. Wenn nun ein späterer Inhaber des Coins, prüfen möchte, ob sein Coin von einem vertrauenswürdigen Dienstleister erstellt wurde, so kann er die blinde Signatur mit dem öffentlichen Schlüssel der Institution entschlüsseln. Das Ergebnis dieser Berechnung zeigt, dass die Institution den Coin tatsächlich signiert hat und sie dem Coinaussteller vertraut. Der hierbei essentielle Punkt ist, dass diese Institution nicht weiß, was sie unterschreibt. Somit kann die Institution die Beziehung des Dienstleisters und des Coininhabers nicht nachverfolgen.

Die für das Verfahren nötigen Rechenschritte sind im Folgenden beschrieben. Angenommen der Unterzeichnende verfügt über eine private Signierfunktion s' und eine öffentliche Funktion s, sodass $s(s'(x)) = x$. Der Nutzer verfügt über die privaten Funktionen c und dessen invers c', sodass $c'(s'(c(x))) = s'(x)$.

1. Der Nutzer beginnt nun damit, sich ein x auszusuchen. Dieses wird durch zufällige Redundanz vor Kollisionen geschützt und mit $c(x)$ unkenntlich gemacht.

2. Anschließend erhält der Unterzeichnende $c(x)$, berechnet $s'(c(x))$ und schickt den entsprechenden Wert an den Nutzer zurück.

3. Wenn der Nutzer nun $c'(s'(c(x))) = s'(x)$ berechnet, so erhält er den signierten Ursprungswert x ohne, dass der Unterzeichnende x je kannte.

Daraufhin kann jede weitere Person die Unterschrift überprüfen, indem diese die öffentliche Funktion s verwenden um $s(s'(x))$ zu berechnen und das Ergebnis mit x abgleichen.

Blinde Signaturen sind bei dem Ansatz von privatsphäreschützenden Zahlungs- und Reputationssystemen einer der Kernbausteine der verwendet wird, um einen vertrauenswürdigen Austausch zwischen den Parteien zu ermöglichen. Sie liefern die Grundlage der Kommunikation+0

3.6 Partiell Blinde Signaturen

Partiell blinde Signaturen ähneln sich vom Effekt stark zu blinden Signaturen von Chaum aus dem vorherigen Absatz **??**. Der jedoch entscheidende Unterschied ist, dass es bei partiell blinden Signaturen eine Infowert gibt der sowohl Nutzer als auch Unterzeichnende bekannt ist. Dieser kann genutzt werden, damit der Unterzeichnende möglicherweise Prüfung ausführen kann und zu entscheiden ob er die Nachricht wirklich unterzeichnen möchte. Diese Eigenschaft wird im späteren Verlauf bspw. dazu verwendet einen Coin partiell blind zu signieren. Hierbei ist es essentiell, dass der Unterzeichnende den monetäre Wert des Coins prüfen kann, da ein Nutzer sonst freie Kontrolle über den Wert seines eigenen Coins hat. Mit partiell blinden Signaturen, kann genau dies gewährleistet werden. Dabei ist die monetäre Wert des Coins die Info und der kryptographische Wert des Coins die Nachricht. Nun kann der Unterzeichnende prüfen, dass der monetäre Wert der vereinbarte Wert ist und kann den kryptographischen Wert des Coins blind signieren. Der springende Punkte ist der Zusammenhang zwischen Info und Nachricht, da sonst der monetäre Wert des Coins nach dem Unterzeichnen durch den Nutzer verändert werden könnte. Um dies zu verhindern, verwendet das Verfahren den Hash der Info als Teil der Signatur, so dass diese nur gültig ist, solange die Info unverändert bleibt.

3.6.1 Funktionweise

3.7 Elliptische Kurven Kryptographie

Die Kryptographie über elliptische Kurven (ECC - elliptic curve cryptography) ist ein Zweig der Kryptographie welcher bereits seit 1985 besteht [**ecc-miller1985use**] und trotz dessen nur wenig Aufmerksamkeit genießt . Hierbei geht es um das Ver- und Entschlüsseln von Nachrichten anhand von Punkten auf einem elliptischer Graph, was verglichen mit der weit verbreiteten Faktorisierungskryptographie große Performanzsteigerung liefert.

3.7.1 Trapdoor functions

Das asymmetrische RSA Verschlüsselungsverfahren ist heute weltbekannt und wird in über 90% der Onlinekommunikation verwendet [**ecc-rsa_amount**]. Dessen Sicherheit basiert genauso wie die von ECC auf sog. *Trapdoor functions*. Ein Trapdoor function ist ein mathematisches Problem, welches in eine Richtung leicht zu berechnen ist, jedoch das Inverse enorm schwer. Im Falle von RSA ist die Trapdoor function das Faktorisierungsproblem. Es ist leicht 2 sehr große Zahlen miteinander zu multiplizieren. Allerdings ist es enorm schwer bei einem gegebenen Produkt dessen Faktoren zu bestimmen, insbesondere wenn die Faktoren jeweils Primzahlen sind. Dies ist der Grund weshalb RSA sicher ist und zuverlässig die Kommunikation schützt. Bei ECC ist die Trapdoor function, die die Sicherheit der Verschlüsselung ausmacht, eine andere. Jedoch kann anhand dieser, ein public und private key generiert werden wie es für asymmetrische Kryptographie nötig ist.

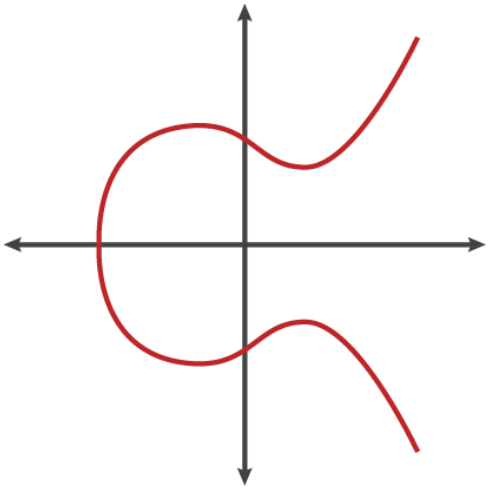

Abbildung 3.4: Mögliche Form einer elliptischen Kurve [**ecc-cloud2013elliptic**]

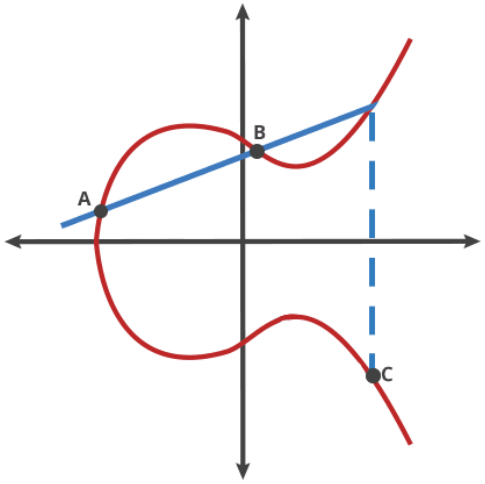

Abbildung 3.5: Operation $\circ$ auf Punkten $A, B \in M$ [**ecc-cloud2013elliptic**]

3.7.2 Funktionsweise

Um die Trapdoor function hinter ECC zu verstehen müssen wir uns zuerst elliptische Kurven anschauen. Die algebraische Struktur von Elliptische Kurven ist eine Gruppe, welche aus einer Menge von Punkten M und einer binären Operationen $\circ$ auf 2 Punkten der Menge besteht. Die Definition einer Gruppe fordert, dass die Element die Eigenschaften der Assoziativität, Identität, Existenz eines inversen Elements und je nach Quelle auch Abgeschlossenheit erfüllen. [**ecc-aradi2016einfuhrung**][**ecc-bogopolÊźskij2008introduction**] Zudem müssen die Koordinaten aller Punkte $(x, y) \in M$ die in der Menge liegen aus dem endlichen Feld stammen, sowie die Gleichung:

$$y^2 = x^3 + ax + b \tag{3.1}$$

erfüllen. Zusammen bildet die Menge an Punkten M einen Graph der je nach Wahl der Parameter a, b einer Form aus Abbildung **??** ähnelt. Aufgrund der Gleichung **??** entsteht der Zusammenhang, dass ein Gerade durch 2 zufällig gewählte Punkte $P, Q \in M$, den Graph immer an genau einer dritten Stelle schneidet. Zusätzlich ist der Graph an der X-Achse durch das y^2 gespiegelt. Mit diesen Eigenschaften kann nun die Operation $\circ$ definiert werden.
Diese arbeitet wie folgt: Bei Eingabe von $A, B \in M$ finde den invertierten dritten Schnittpunkt mit dem Graph. Dieser sei hier mit C beschrieben. Ein Ausführung von $A \circ B = C$ ist in Abbildung **??** verdeutlicht.

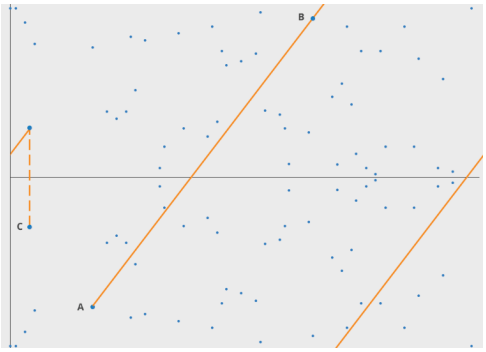

Abbildung 3.6: Graph nach Anwendung von $\mathbb{Z}_p$ [**ecc-cloud2013elliptic**]

Die Operation $\circ$ kann beliebig oft hintereinander mit dem jeweils neu entstehenden Punkt ausgeführt werden. Genau hier liegt die Trapdoor function von ECC. Mit Kenntnissen über der Startpunkt A und die Anzahl der Ausführungen n, ist es einfach den Endpunkt E zu berechnen. Wenn allerdings nur der Endpunkt E und Startpunkt A gegeben sind, ist es sehr rechenaufwändig die Anzahl an Ausführungen zu bestimmen.

3.7.3 Anwendung in der Kryptographie

Allgemein ist zu beachten, dass elliptische Kurven die für ein kryptographischen Verwendung in Frage kommen, eine andere Form haben, als die eben aufgeführt Abbildungen (**??**,**??**). Entscheidend ist hierfür die Wahl des endlichen Feldes. In der Kryptographie wird hier meist $\mathbb{Z}_p$ verwendet. Diese Wahl bringt 2 Eigenschaften mit sich:

1. Das p in $\mathbb{Z}_p$ besagt, dass alle Werte $x \in \mathbb{Z}_p$, $0 \leq x \leq p-1$ erfüllen müssen. Durch die Wahl dieses Wertebereichs, können die X und Y-Achse nie den Wert p überschreiten sondern beginnt anstatt wieder bei 0.

2. Die zweite Eigenschaft ist die ausschließliche Betrachtung ganzer Zahlen als X und Y-Koordinate. Hierdurch wird aus dem Graphen eine Menge an zufällig aussehend gewählten Punkten. Der Graph für die kryptographische Verwendung einer elliptische Kurve ist ein Abbildung **??** visualisiert.

Nun muss ein passendes p sowie a, b bestimmt werden, um die genaue Menge an Punkten zu spezifizieren. Hierfür gibt es bereits ein große Auswahl und Werten in der Literatur [**ecc-lochter2010elliptic**][**ecc-merkle2013ellipt**
Eine der prominentesten ist die “Curve25519“ von Daniel J.Bernstein [**ecc-bernstein2006curve25519**]. Sie hat den Werte $p = 2^{255} - 19$ (daher der Name) und die elliptische Funktion $y^2 = x^3 + 486662x^2 + x$. Dies ist die Kurve, welche im folgenden verwendet wird.

3.7.4 Vor und Nachteile

Allgemein betrachtet ist die Tatsache, dass ECC eine schneller Berechnungszeit liefert als RSA, nicht von der Hand zu weisen [**ecc-cloud2013elliptic**]. Da die Laufzeiten von Verschlüsselung und Entschlüsselung durch RSA und ECC asymptotisch nicht gleich verhalten ist es schwer eine endgültige Antwort zu nennen. Allerdings schlägt ECC RSA bei der Gesamtzeit für Ver- und Entschlüsselung je nach Nachrichtengröße um einen Faktor von ~$3-20$ [**ecc-mahto2018performance**][**ecc-bao2022research**]. Zudem ist eignet sich ECC vorallem als Ver- und Entschlüsselungsmethode auf rechenschwachen Geräte wie Mobiltelefonen aufgrund von kleineren Schlüsselgrößen[1]. So können Rechenaufwand, Energieverbrauch und RAM-Auslastung gesenkt werden [**ecc-gupta2011ecc**].

1. Um eine security bit level von 256bits mit RSA zu erreichen, ist eine Schlüssellänge von 15360bits nötig während es bei ECC lediglich 512bits sind.[**ecc-mahto2018performance**]

Allerdings wurde 2007 bekannt, dass der Pseudo Random Number Generator (PRNG) Dual_EC_DRBG eine potentielle Backdoor hatte, die es Nutzern mit Informationen über einen Wert die Zufälligkeit problemlos knacken lies [**ecc-green2013backdoor**]. Dual_EC_DRBG wurde 2005-2006 zusammen von NIST (National Institute of Standarts and Technology) und der NSA veröffentlicht und basiert die Auswahl der zurück gegebenen pseudozufälligen Zahlen auf Berechnung über elliptischen Kurven. Dieser Vorfall schwächt bis heute das generelle Vertrauen in ECC.

[**ecc-cloud2013elliptic**]

4 Zielsetzung

Im folgenden soll ein Systeme vorgestellt werden welche einen Anreiz für Datengeber stellt, seine persönlichen Daten durch einen Datentreuhänder weiter zugeben. Um die bösartige Ausnutzung dieses Anreizes zu verhindern wird das System Mechanismen enthalten welche einem Datengebenden einen Ruf zuteilt und verwaltet. Das System wurden in dem Tresor-Projekt der Universität Hamburg implementiert um anhand der Implementation Auswertungen zu erstellen.

Payment System. Hier ist vorgesehen, dass ein Datennutzender bezahlen muss bevor er Zugang zu den Daten erhält. Dafür sendet der Datennutzende einen oder mehrere interne Coins an den Datengebenden, welcher diese durch den Datentreuhänder prüfen lässt. Wenn die Prüfung erfolgreich ist, gibt der Datengebende den Zugang zu seinen Daten frei und erstellt erstellt einen Bewertungstoken. Der Datennutzenden kann nun die Daten auswerten und später mit dem Bewertungstoken eine Bewertung für den Datengebenden ausstellen. Anhand dieser Bewertung wird ein Ruf für jeden Datengebenden bestimmt, der eine Einschätzung über die Qualität der Daten liefert. Sollte also ein Datengebender ein großen Menge an nutzlosen Daten anbieten und so das System auszunutzen, so kann dies durch einen schlechten Ruf von Datennutzenden erkannt werden bevor die Daten erworben werden.

4.1 Begründungen

4.1.1 Geld als Anreiz

Der hier gewählte Anreiz für den Datengeber ist eine kleine monetäre Summe als Kompensation. Dies hat vor allem 5 Gründe:

1. Kein Handelsgut oder Dienstleistung ist so universelles begehrt wie Geld. Zwar ist es möglich das ein Business welches hier als Datennutzer auftritt, für die Preisgabe der Daten eine Dienstleistung anbietet. Bspw. können Social Media Plattformen wie X konstenlose zeitlichbedingte Premiumabonnenments oder vergleichbar Angebote in Aussicht stellen. Jedoch sind ist das Interesse an solchen Angeboten subjektiv zu bewerten, was sie als mögliche Anreize zwar denkbar aber trotzdessen suboptimal macht.

2. Wenn der Anreiz durch den Datennutzende gestellt wird, wird dieser in der Vielzahl der Fälle ein unterschiedlicher sein. Aus der Kommunukationsstruktur folgt das alle unterschiedlichen Kompensationen auf irgendeinem Weg über der Datentreuhänder laufen. Dies bedeutet einen enormen Aufwand an Entwicklung und Implementation.

3. Manche Datennutzende wie bspw. Forschungsinstitutionen haben keine Dienstleistungen oder Produkte die sie Datengebenden anbieten können. Sie sind ausschließlich an der Forschung interessiert und können den Datengebenden nichts bieten was sie zum preisgeben ihrer Daten veranlassen würde.

4. Viele C2B Treuhänder legen Wert auf die Anonymität oder Pseudonymität der Datengeber, welche durch die Beanspruchung von Angeboten der Datennutzer verloren gehen können.

5. Der Datentreuhänder kann bei anderen Anreizen nicht garantieren dass der Datengebende die von ihm angeboten Dienstleistungen auch einhält.

Durch eine die Verwendung von Geld als Anreiz werden diese Punkte umgangen. Das Interesse der Datengebenden ist nun ausschließlich abhängig von dem Ruf und dem persönlichen Interesse an einem Datengebenden und nicht von dessen Anreizangebot. Die in Anspruchnahme des versprochenen Anreizes ist für jeden Datennutzenden gleich Implementiert und hängt maximal von der Wahl anderer Parameter ab. Es kann davon ausgegangen werden dass alle Datennutzenden über Geld verfügen und dieses als Kompensation verwenden können. Aufgrund der (trivialität?) des Wertes von Geld sind bereits viele Wegen bekannt wie diesen Wert anonym oder pseudony an andere überträgt. Der Datentreuhänder hat die volle Kontrolle, dass ein zu zahlender Geldbetrag vom Datennutzenden abgegeben wird und auch bei Datengebenden ankommt.

5 Entwurf der Systeme

5.1 Generelles

In diesem Abschnitt wird genau beschrieben wie die Pay-First und Pay-Later Systeme aufgebaut sind und funktionieren.
Um den Zahlungsablauf unabhängig von anderen Technologien wie Bankanbindungen o.ä. zu gestalten wird in den folgenden Systemen eine eigene interne Währung etabliert. Da diese ausschließlich innerhalb des Systems nutzbar ist trägt sie keinen bestimmten Namen und wird im folgenden mit Coin oder Coins referenziert. Zwar ist ein Tausch von Coins für Geld wie im Falle von Kryptowährungen möglich, allerdings ist der monetäre Wert eines Coins in diesem gespeichert, was den Tausch zu Spekulationszwecken sinnlos macht.
Weiter sei angemerkt, dass für die Systeme ein weiterer Akteur in die Kommunikation mit dem Datentreuhänder eingeführt wird. Im Standartmodell eines Datentreuhänder existieren die Rollen des Datengebenden der seine Daten anbietet, des Datennutzenden der an den Daten des Datengebenden interessiert ist und die des Datentreuhänders der Datengebende und Datennutzende zusammenführt und deren Interaktion verwaltet. Im folgenden wird eine Datentreuhänder (Payment) eingeführt. Er ist ausschließlich für die Verwaltung von Zahlungsgeschäften zwischen Datennutzenden und Datengebenden zuständig. In der Regel ist angedacht diesen Akteur zusammen mit dem Datentreuhänder innerhalb eines Netzwerks oder Rechner laufen zu lassen, allerdings kann dieser auch separat an einer anderen Stelle laufen.
Alle Kommunikationen zwischen 2 Akteueren sind mit ECC verschlüsselt sowie signiert, um Vertraulichkeit und Integrität zu gewährleisten.

5.2 Coin Generierungsphase

Hier werden die Coins die im späteren Verlauf beider Systeme verwendet werden vom Datennutzenden erstellt und vom Datentreuhänder signiert. Die einzelnen Schritte sind in Abbildung **??** verdeutlich und werden im folgenden beschrieben.

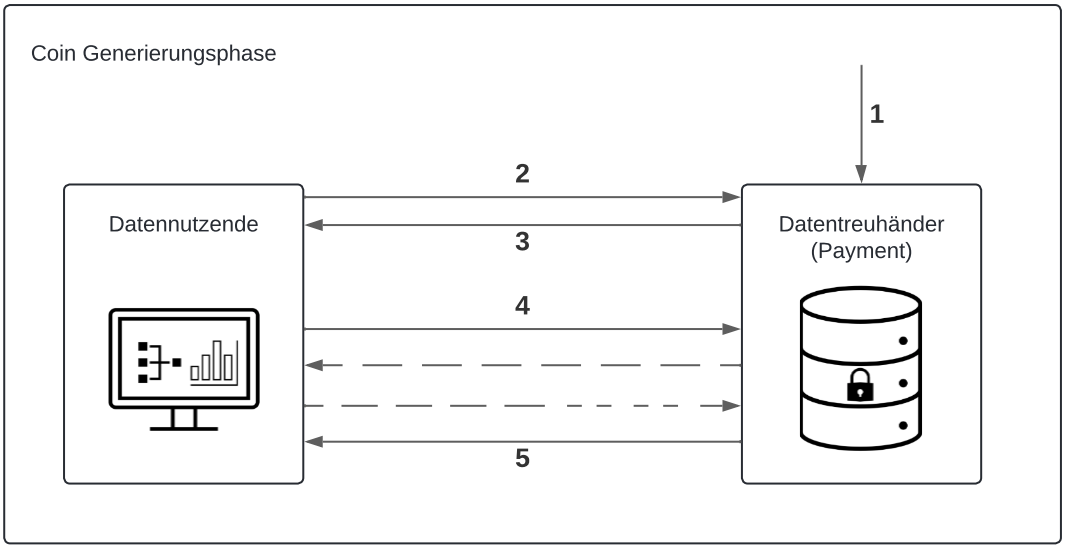

Abbildung 5.1: Coin Generierungsphase Ablauf

1. Zahlungseingang $(DN\text{-}ID, ES \leftarrow ErhalteneSumme)$
Der Datentreuhänder (Payment) erhält über einen beliebigen Weg eine Summe und eine Datennutzende ID. Die genaue Umsetzung des Zahlungseingang liegt bei dem Datentreuhänder (Payment). Er kann alles zwischen einer Banküberweisung mit der ID als Verwendungszweck, bis hin zu einem Briefumschlag mit Bargeld und einer leserlich aufgeschrieben ID sein. Nach Eingang einer Zahlung erstellt der Datentreuhänder (Payment) einen sogenannten CoinGenerationToken bestehend aus $(nonce, DN\text{-}ID, ES, spent \leftarrow false)$ und speichert diesen intern.

2. CoinGenerationToken abrufen $(DN\text{-}ID)$
Da je nach Zahlmethode große zeitliche Abstände zwischen Zahlungsabsendung und -eingang entstehen kann und nicht gewährleistet werden kann, dass der Datennutzende zum Zeitpunkt des Zahlungseingangs erreichbar ist, kann der Datentreuhänder (Payment) ihn nicht zuverlässig benachrichtigen. Aufgrund dessen kann der Datennutzende beim Datentreuhänder (Payment) nachfragen ob es offene CoinGenerationToken für ihn gibt. Dies wird in diesem Schritt getan.

3. CoinGenerationToken übertragen $[(nonce, DN\text{-}ID, ES)]$
Auf eine CoinGenerationToken Anfrage des Datennutzenden antwortet der Datentreuhänder (Payment) mit allen noch nicht eingelösten Token. Nachdem der Datennutzende die Token erhalten hat, kann er damit beginnen selbst Coins zu erstellen. Ein Coin besteht aus $(nonce, value)$. Bei der Erstellung sind 2 Sachen zu beachten. Zuerst muss der summierte Wert aller erstellten Coins gleichgroß sein wie die ES des Tokens. Des weiteren gibt es eine Menge an möglichen Werten PV, sodass $\forall value \in PV$ gilt. Dies ist vor allem wichtig, da $value$ bei der Signierung und bei der Einlösung des Coins für den Datentreuhänder (Payment) sichtbar sind und dieser bei einer Wahl von selten vorkommenden $value$ ein Verbindung zwischen den Phasen herstellen kann.

4. Signierung Anfragen $(nonce, DN\text{-}ID, ES), [(\widehat{nonce}, value)]$
Eine Signieranfrage besteht aus einem CoinGenerationToken und einer Menge an Coins die partiell geblindet wurden, sodass $value$ einsehbar ist aber $nonce$ geblindet wurde. Zuerst kann der Datentreuhänder (Payment) prüfen ob der gesendete Token noch nicht eingelöst wurde. Anschließend summiert er alle $value$ auf und prüft ob die Summe gleich dem ES des Tokens ist. Wenn beide Überprüfungen akzeptieren, kann der Datentreuhänder (Payment) die Coins partiell blind signieren und bei seinem gespeichert Token $spent \leftarrow true$ setzen, um eine doppelte Einlösung zu verhindern.

Gestrichelte Pfeile
Die Pfeile zwischen 4. und 5. sind hier gestrichelt eingetragen, da sie notwendige Kommunikationsschritte von partiell blinden Signaturen abbilden. Sie sind essentiell für die Funktionsweise und wurden bereits in **??** erklärt, weshalb sie hier nur zur Vollständigkeit aufgelistet aber nicht weiter benannt werden.

5. Signieranfrage Antwort $[(sign(\widehat{nonce}), value)]$
Vorrausgesetzt alle Überprüfungen aus Schritt 4 akzeptieren, so erhält der Datennutzende nun eine Menge an partiell blinden signierten Coins, kann diese wieder unblinden und für spätere Verwendung lokal speichern. Sollten die Überprüfungen nicht akzeptieren, so kann der Datennutzende entweder bei Schritt 4 mit anderen Coins oder bei Schritt 2 wieder ansetzen.

Auf diese Weise kann ein Datennutzender Geld bei einem Datentreuhänder (Payment) einzahlen und eine dem Geldbetrag entsprechende Menge an Coins erhalten, ohne dass der Datentreuhänder (Payment) die von ihm ausgehändigten Coins nachverfolgen. Gleichzeitig ist es für einen Datennutzende nicht möglich Coins zu erhalten für die er keinen monetären Gegenwert bereitgestellt hat.

5.3 Bezahlvorgang

5.4 Reputationsvergabe

6 Implementation

Wird kurz gehalten. Implementierung beschreiben

6.1 TRESOR Projekt

7 Auswertung

Evaluation ist wichtiger als guter Ansatz
Sicherheit gegen Angreifer Modelle: - Honest but curios DT - bösartige DN
Graphs and Stats to make: ECC vs RSA weighted shifting average vs linear reputation calculation security bit level performance

7.1 Setup / Implementation

7.2 Metrics

8 Bewertung

8.1 Erfüllen der Anforderungen

8.2 Research question 1

8.3 Research question 2

8.4 (Discussion)

9 Limitations

10 Conclusion

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit im Bachelorstudiengang Software-System-Entwicklung selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel — insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen — benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Arbeit vorher nicht in einem anderen Prüfungsverfahren eingereicht habe.

Hamburg, den 9. September 2024

Vorname Nachname