universität innsbruck

Prinzipien von Blockchain-Systemen

Datenschutz und Sicherheit

Rainer Böhme

Abgrenzung

Datensicherheit

(engl. “data / information security”)

schützt Daten durch
Schutz von IT-Systemen

Datenschutz

(engl. “data protection / privacy”)

schützt Menschen durch Schutz
ihrer personenbezogenen Daten

Wiederholung aus: VO Rechnernetze und Internettechnik, Kapitel „Sicherheit“, 6. Juni 2019, S. 7.

Agenda

1. **Datenschutz in Blockchain-Systemen**

2. Sicherheit in und von Blockchain-Systemen
   - Individuelle Sicherheit — *„Mein Geld bleibt meins.“*
   - Systemische Sicherheit — *Fortbestand des Systems*

3. Sicherheitsrisiken durch Blockchain-Systeme

Datenschutz in Blockchain-Systemen

Fundamentaler Konflikt zwischen:

- Öffentlichkeit der Blockchain-Daten, erforderlich für die dezentrale Verifikation,

und

- Persönlichkeitsrechten der Nutzer, denn Informationen in (Finanz-)Transaktionen lassen Rückschlüsse auf Gewohnheiten und Lebensumstände zu.

Erhoffte Milderung:

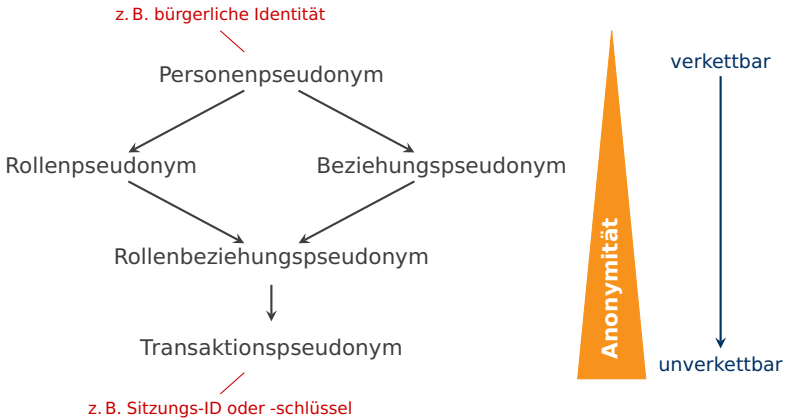
- Pseudonyme sind nicht unmittelbar natürlichen Personen zuordenbar.

Trotzdem sind Blockchain-Daten als **personenbeziehbar** anzusehen, denn:

- Personenbezug kann oft durch **Verknüpfung mit weiteren Daten** (z. B. in der Hand von Intermediären wie Wechselbörsen, P2P-Datenpakete) hergestellt werden.
- Die Unveränderbarkeit der Blockchain erhöht die Wahrscheinlichkeit, dass dies in der Zukunft geschieht.

Arten von Pseudonymen

Einteilung nach der Wiederverwendung von Identifikatoren



Pfitzmann, A., Köhntopp, M. Anonymity, Unobservability, and Pseudinymity — A Proposal for Terminology. In H. Federrath, ed., *Proceedings of Anonymity 2000*. LNCS 2009, Springer, Berlin, 2001, 1–9.

Ansätze zur Deanonymisierung bei Bitcoin

Pseudonyme	Know your customer (KYC)
Blockchain	Clustering-Heuristiken
Overlay-Netz	Echtzeitüberwachung Aktive Angriffe auf Clients

Verkettung von Bitcoin-Adressen zu Entitäten

Bekannte Heuristiken um Adressen in Entitäten (d.h. vermuteten Parteien) zu clustern:

1. Multiple-Input-Heuristik

- **Annahme** Wenn Inputs aus mehreren Adressen stammen, werden diese von der gleichen Partei kontrolliert.

$$T_3 = \big((T_{1,1}, T_{2,1}), ((1.0, \boldsymbol{t}_{\text{Casino}})), (\boldsymbol{s}_{\text{Pseudonym A}}(T_3'), \boldsymbol{s}_{\text{Pseudonym B}}(T_3'))\big)$$

2. Wechselgeld-Heuristik

- **Annahme** Die Wechselgeldadresse gehört zur Wallet des Senders.

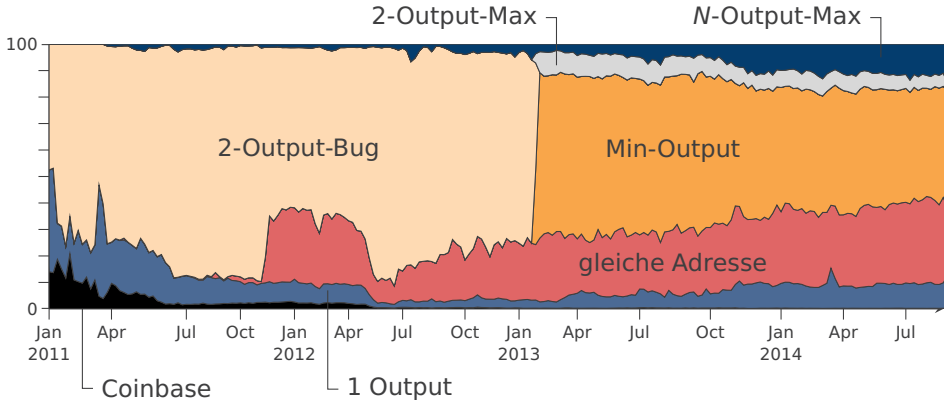
$$T_4 = \big((T_{3,1}), ((0.9, \boldsymbol{t}_{\text{Gewinner}}), (0.05, \boldsymbol{t}_{\text{Pseudonym C}})), (\boldsymbol{s}_{\text{Casino}}(T_4'))\big)$$

- Bitcoin-Clients verschleiern das Wechselgeld durch Randomisierung der Outputs.

Meiklejohn, S. et al. A Fistful of Bitcoins: Characterizing Payments among Men with No Names. *CACM*, 59 (4), 2016, 86–93.

Bestimmung der Wechselgeldadresse

Die Grafik zeigt den Erfolgsanteil (in %) von sieben Heuristiken im Zeitverlauf

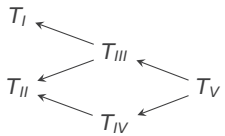


Möser, M., Böhme, R. Unveröffentlichte Analyse zur Schätzung des Netto-Transaktionsvolumens, 2014.

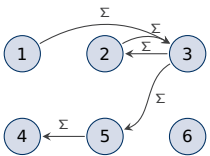
Graphische Blockchain-Analyse

Darstellungsvarianten zur Verfolgung von Zahlungsströmen in Bitcoin

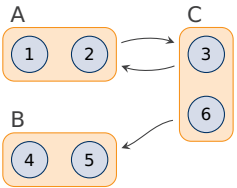
Transaktionsgraph



Adressgraph



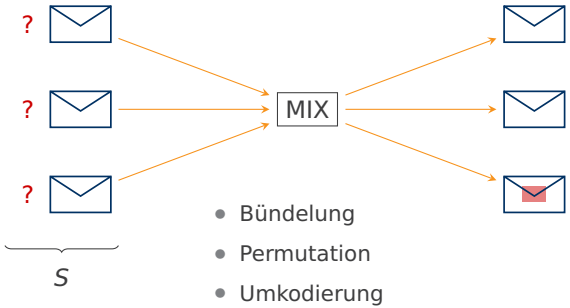
Entitätengraph



Angelehnt an: Abb. 6 aus Tschorsch, F. and Scheuermann, B. Bitcoin and Beyond: A Technical Survey on Decentralized Digital Currencies. *IEEE Communications Surveys and Tutorials*, 18, (3), 2016, 2084–2123.

Das Mix-Prinzip

Herstellung von **Unverkettbarkeit** in Kommunikationssystemen

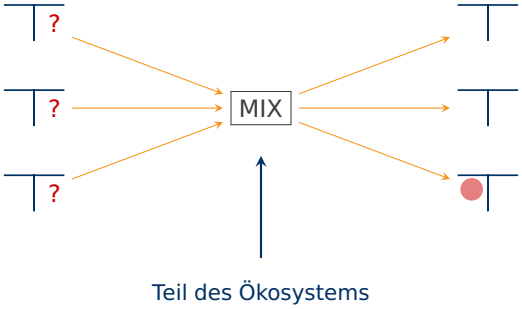


Die Kardinalität der **Anonymitätsmenge** $|S|$ ist eine einfache Datenschutzmetrik.

Chaum, D. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. *CACM*, 24 (2), 1981, pp. 84–88.

Anwendung des Mix-Prinzips

Herstellung von Unverkettbarkeit in **Transaktionssystemen**



Erhebliches **Vertrauen** in Mix-Betreiber erforderlich.

Ausgewählte Datenschutztechniken für Bitcoin

Mixer, Tumbler

- Zentral kontrollierte Intermediäre im Ökosystem
- Erfordern Vertrauen im Bezug auf die Sicherheit der Rückzahlung und die Wahrung der Anonymität

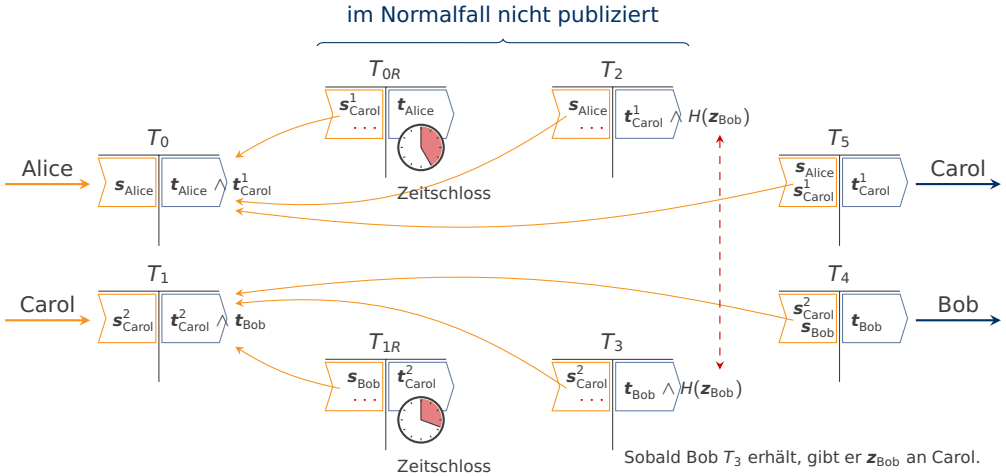
CoinSwap

- Protokoll für atomaren Tausch über dritte Partei ohne direkte Verkettbarkeit in der Blockchain: z. B. Alice $\rightarrow$ Carol$_1$, Carol$_2$ $\rightarrow$ Bob
- Wahrung der Anonymität erfordert immer noch Vertrauen

CoinJoin

- Gemeinsam signierte Transaktion führt zum Versagen der Multiple-Input-Heuristik
- Wahrung der Anonymität abhängig vom (off-chain) Kommunikationskanal und der Zusammensetzung der CoinJoin-Partner

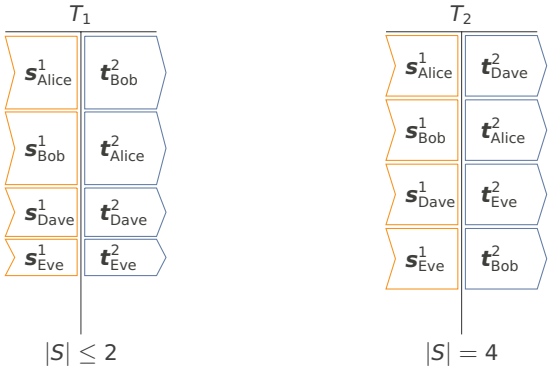
CoinSwap



Maxwell, G. *CoinSwap: Transaction Graph Disjoint Trustless Trading*, 2013.

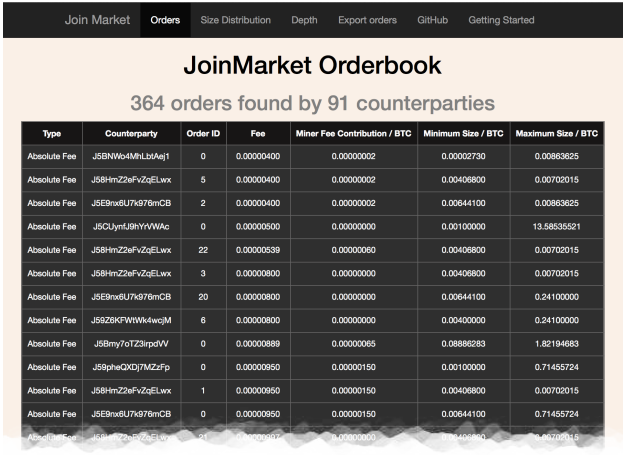
CoinJoin

Prinzip (unter Vernachlässigung von Transaktionsgebühren)



Maxwell, G. *CoinJoin: Bitcoin Privacy for the Real World*, 2013.

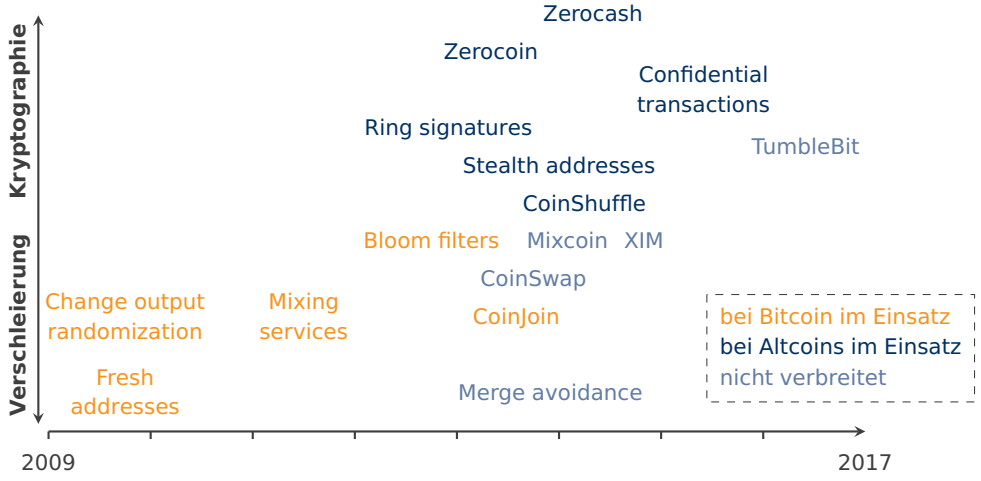
Partnersuche für CoinJoins



Type	Counterparty	Order ID	Fee	Miner Fee Contribution / BTC	Minimum Size / BTC	Maximum Size / BTC
Absolute Fee	J5BNWb4MhLbtAej1	0	0.00000400	0.00000002	0.00002730	0.00863625
Absolute Fee	J58Hm22eFvZqELwx	5	0.00000400	0.00000002	0.00406800	0.00702015
Absolute Fee	J5E9nx6U7k976mCB	2	0.00000400	0.00000002	0.00644100	0.00863625
Absolute Fee	J5CUymfJ9hYrVWAc	0	0.00000500	0.00000000	0.00100000	13.58535521
Absolute Fee	J58Hm22eFvZqELwx	22	0.00000539	0.00000060	0.00406800	0.00702015
Absolute Fee	J58Hm22eFvZqELwx	3	0.00000800	0.00000000	0.00406800	0.00702015
Absolute Fee	J5E9nx6U7k976mCB	20	0.00000800	0.00000000	0.00644100	0.24100000
Absolute Fee	J59Z6KFWtWk4wcjM	6	0.00000800	0.00000000	0.00400000	0.24100000
Absolute Fee	J5Bmy7oTZ3irpdVV	0	0.00000889	0.00000065	0.08886283	1.82194683
Absolute Fee	J59pheQXDj7MZzFp	0	0.00000950	0.00000150	0.00100000	0.71455724
Absolute Fee	J58Hm22eFvZqELwx	1	0.00000950	0.00000150	0.00406800	0.00702015
Absolute Fee	J5E9nx6U7k976mCB	0	0.00000950	0.00000150	0.00644100	0.71455724

Quelle: `http://joinmarket.io`, Abruf: 4. Januar 2017

Datenschutztechniken in Blockchain-Systemen

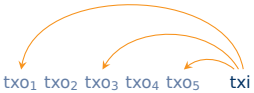


Narayanan, A., Möser, M., *Obfuscation in Bitcoin: Techniques and Politics*, 2017.

Kryptowährungen mit Datenschutz als Entwurfsziel

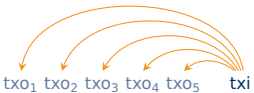
Monero (XMR)

- Bekannteste Instanz des Cryptonote-Protokolls
- Verwendet Ringsignaturen um Mehrdeutigkeiten im Transaktionsgraph zu erzeugen



ZCash (ZEC)

- Verwendet Zero-Knowledge-Technik um ausgewählte Transaktionen völlig unverkettbar zu machen.
- Sehr rechenaufwändig
- Vertrauen in die sichere Wahl globaler kryptographischer Parameter erforderlich



Die Transaktionslogik dieser und ähnlicher Systeme ist eng an die kryptographischen Algorithmen gebunden und beschränkt sich bislang auf Zahlungen.

Agenda

1. Datenschutz in Blockchain-Systemen

2. **Sicherheit in und von Blockchain-Systemen**
   - Individuelle Sicherheit — *„Mein Geld bleibt meins.“*
   - Systemische Sicherheit — *Fortbestand des Systems*

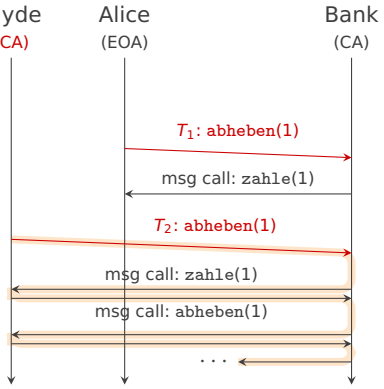
3. Sicherheitsrisiken durch Blockchain-Systeme

Sicherheitslücken durch Programmierfehler

Beispiel: Ethereum-Bankraub durch Wiedereintritt

Funktion `zahle`(a, b)

5 // Saldi aktualisiert
6 Wenn $a.y \geq b$ dann
7 a.call(`abheben`, b)



Zustand

Menge der Kunden K (Adressen)
Einlagen $E : K \mapsto \mathbb{N}_0$
Saldo y

Funktion `abheben`(Sender a, Betrag b)

1 Wenn $b > 0$ und $a \in K$ dann
2 Wenn $E[a] \geq b$ dann
3 a.call(`zahle`, b)
4 $E[a] \leftarrow E[a] - b$

Atzei, N., Bartoletti, M., and Cimoli, T. A Survey of Attacks on Ethereum Smart Contracts. In M. Maffei and M. Ryan, eds., *Principles of Security and Trust*. Lecture Notes in Computer Science 10204, Springer, Berlin, Heidelberg, 2017, 164–186.

Zufall als Fehlerquelle

Kein Zufall in der Ausführungsumgebung

Die Ausführungsumgebung ist deterministisch. Wie realisieren wir eine Lotterie?

- Lösung: Pseudozufall durch Hashwerte über Blockchain-Daten
- Aber: Wenn der Miner ein Interesse hat (selbst oder gegen „Bestechung“), kann er ungünstige Ausgänge zurückhalten.
- **Merke:** Zufall bedeutet, durch den Miner gewählt!

Schwacher Zufall in Wallets

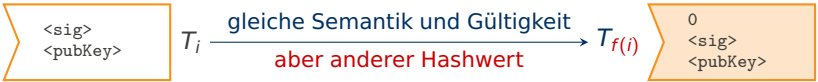
- Offensichtlich: Private Schlüssel dürfen nicht erratbar sein.
- Subtiler: Die in Blockchain-Systemen verwendeten ECDSA-Signaturen erlauben die Berechnung des privaten Schlüssels $\boldsymbol{s}$, wenn verschiedene Nachrichten $\boldsymbol{m}_1, \boldsymbol{m}_2, \ldots$ mit dem selben Schlüsselpaar und der gleichen Zufallszahl $\boldsymbol{z}'$ signiert wurden.
- Bsp.: Schwäche im Zufallszahlengenerator von Android $\rightarrow$ Bitcoin-Wallets leer.

Quelle: `https://bitcoin.org/en/alert/2013-08-11-android`

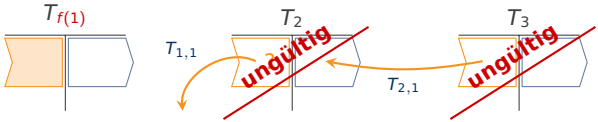
Bitcoins Kinderkrankheit: Transaction Malleability

engl. *malleable*: **verformbar**, in der Kryptographie die Möglichkeit, ohne Kenntnis des Schlüssels gültige Schlüsseltexte in andere gültige Schlüsseltexte umzuwandeln.

Ursache Input-Skripte werden nicht mit signiert ($T_i \leftrightarrow T_i'$). (leider nicht die einzige)



Problem Referenzen auf schwebende Transaktionen sind unsicher.



Transaction Malleability schränkt die Möglichkeiten für Off-Chain-Anwendungen ein.

Seit August 2017: Transaktionen, die nur SegWit-Outputs ausgeben, sind nicht malleable.

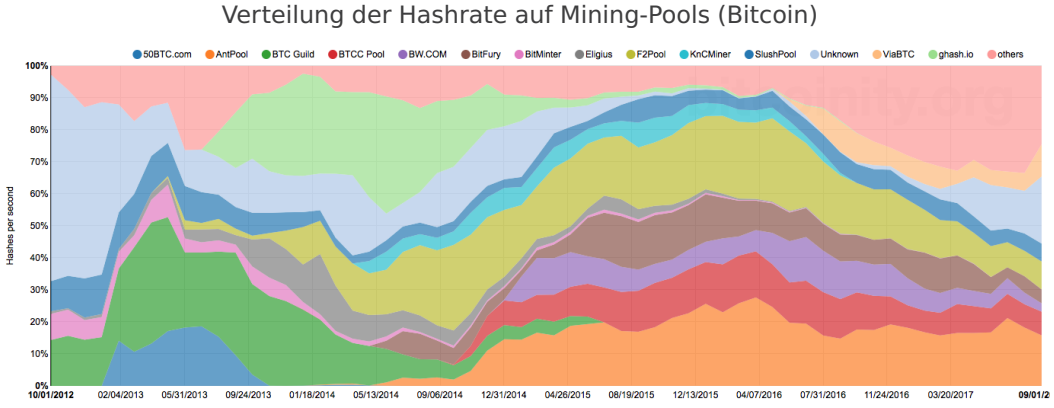
51 %-Angriff

Miner, die nachhaltig mehr als 50 % der Rechenkapazität aufwenden (können), kontrollieren den Schreibzugriff auf den Ledger:

- Sie können Transaktionen bevorzugen oder ausschließen.
- Sie können einen Teil der Geschichte umschreiben.
- Sie können keine Signaturen fälschen.

Abschätzung des RENT-Angriffs: Bonneau, J. *Hostile Blockchain Takeovers*, Financial Cryptography Workshop, 2. März 2018

Historische Daten



Quelle: `https://data.bitcoinity.org/`, September 2017. Im Stream schauen wir uns aktuelle Daten an und besprechen eine neue Methode, die nicht gekennzeichnete „Mystery“-Blöcke Mining-Pools zuordnen kann (Romiti et al., WEIS 2019).

Zurückhalten von Informationen

Blockchain-Protokolle basieren auf Informationsaustausch.

Manipulation ist schwer, denn Kryptographie sichert die Integrität der Nachrichten.

Jedoch verhindert sie nicht, dass Knoten Informationen **strategisch** zurückhalten.

Zurückhaltung von Transaktionen

- Bei niedrigen Transaktionsgebühren: Leere Blöcke verwaisen seltener.
- Bei hohen Transaktionsgebühren: Kein Anreiz zur Weiterleitung von Transaktionen an konkurrierende Miner

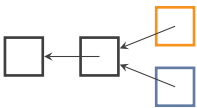
Zurückhaltung von Blöcken

- Miner’s Dilemma: Sabotage konkurrierender Mining-Pools
- Selfish Mining: Kontrolle des Konsens-Mechanismus mit weniger als 50 % Leistung

Beispiel: Selfish Mining bei Bitcoin

Angenommen, wir finden einen Block. Was passiert, wenn wir ihn nicht veröffentlichen?

Mit Wahrscheinlichkeit p veröffentlichen andere Miner einen konkurrierenden Block. Wir ziehen sofort nach.



Wer das Rennen gewinnt, hängt von Glück, Netzanbindung und Blockgröße ab.

ODER

Zurück zum ersten Schritt. Mit Wahrscheinlichkeit $1 - p$ finden wir vor allen anderen einen zweiten Block.

Jetzt gewinnen wir praktisch jedes Rennen gegen protokollkonforme Miner.

Eyal und Sirer (FC 2014) haben gezeigt, dass Selfish Mining mehr Rechenleistung verschwendet als sofortige Veröffentlichung, aber die Selfish Miner haben einen **höheren Erwartungswert** für Blockgewinne.

Ein **Selfish-Mining-Pool** kann mehr auszahlen, gewinnt so neue Miner und baut damit seine Macht aus.

Agenda

1. Datenschutz in Blockchain-Systemen

2. Sicherheit in und von Blockchain-Systemen
   - Individuelle Sicherheit — *„Mein Geld bleibt meins.“*
   - Systemische Sicherheit — *Fortbestand des Systems*

3. **Sicherheitsrisiken durch Blockchain-Systeme**

Ransomware

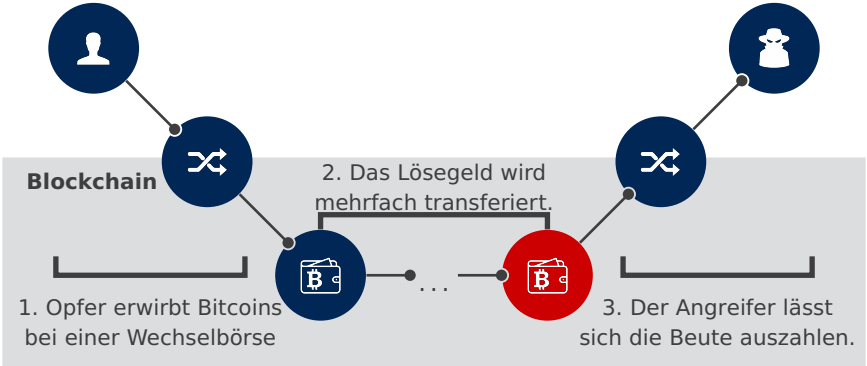
Schadsoftware, die ein Lösegeld (engl. *ransom*) erpresst indem sie persönliche Dateien verschlüsselt oder einem Opfer den Zugang zum System verwehrt.

Digitale Werbetafel in der Kranebitter Allee, Innsbruck (Oktober 2019)

Ablauf einer Ransomware-Infektion

Ablauf einer Lösegeldzahlung

- Deposit-Adresse: vom Angreifer erzeugt um Zahlungen des Opfers zu empfangen
- Sammeladresse: von Angreifer zur Zusammenführung von Zahlungen verwendet



Messung der Ransomware-Aktivität

Problem: keine verlässlichen Daten zu Infektionen und Deposit-Adressen

Ansatz: Schätzung einer unteren Schranke für Lösegeldzahlungen durch Blockchain-Analyse und Identifikation von Deposit- und Sammeladressen

Journal of Cybersecurity, 2019, 1–11
doi: 10.1093/cybsec/tyz003
Research paper

Research paper

Ransomware payments in the Bitcoin ecosystem

Masarah Paquet-Clouston[1,*], Bernhard Haslhofer[2] and Benoit Dupont[3]

[1]GoSecure, 630 René-Lévesque Blvd W #2630, Montreal, QC H3B 1S6, Canada, [2]AIT Austrian Institute of Technology, Giefinggasse 4, 1210 Vienna, Austria and [3]Université de Montréal, 2900 Edouard Montpetit Blvd, Montreal, QC H3T 1J4, Canada

*Corresponding address: Research and Development Department, GoSecure, 630 Boul. René-Lévesque West, #630, Montréal, QC, H3B 1S6. E-mail: mcpc@gosecure.ca

Received 9 August 2018; revised 18 December 2018; accepted 0 Month 0000

Abstract

Ransomware can prevent a user from accessing a device and its files until a ransom is paid to the attacker, most frequently in Bitcoin. With over 500 known ransomware families, it has become one of the dominant cybercrime threats for law enforcement, security professionals, and the public. However, a more comprehensive, evidence-based picture on the global direct financial impact of ransomware attacks is still missing. In this article, we present a data-driven method for identifying and gathering information on Bitcoin transactions related to illicit activity based on footprints left on the public Bitcoin blockchain. We implement this method on-top-of the GraphSense open-source platform and apply it to empirically analyze transactions related to 35 ransomware families. We estimate the lower bound direct financial impact of each ransomware family and find that, from 2013 to mid-2017, the market for ransomware payments has a minimum worth of USD 12 768 536 (22 967.54 BTC). We also find that the market is highly skewed with only a few number of players responsible for the majority of the payments. Based on these research findings, policy-makers and law enforcement agencies can use the statistics provided to understand the size of the illicit market and make informed decisions on how best to address the threat.

Key words: ransomware; economics; Bitcoin; graph analysis

© The Author(s) 2019. Published by Oxford University Press.

Tracking Ransomware End-to-end

Danny Yuxing Huang[1], Maxwell Matthaios Aliapoulios[2], Vector Guo Li[3]
Luca Invernizzi[4], Kylie McRoberts[4], Elie Bursztein[4], Jonathan Levin[5]
Kirill Levchenko[3], Alex C. Snoeren[3], Damon McCoy[2]

[1] Princeton University [2] New York University [3] University of California, San Diego [4] Google Inc [5] Chainalysis

≈ USD 12,8 Mio.
(2013 bis Mitte 2017)

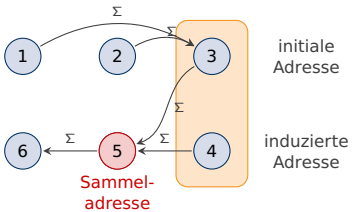
≈ USD 16 Mio.
(2016 bis Mitte 2017)

Verfolgung von Lösegeldzahlungen

öffentliche Foren | Datenbanken | Honey-„Opfer“

initiale Deposit-Adressen

konstant über mehrere Infektionen
(z. B. WannaCry attack)



eindeutig für jede Infektion
(z. B. Locky & Cerber)

↓

Erweiterung der Adressmenge
→ Multi-Input-Heuristik

↓

Ausschluß von False Positives
(z. B. durch CoinJoins)

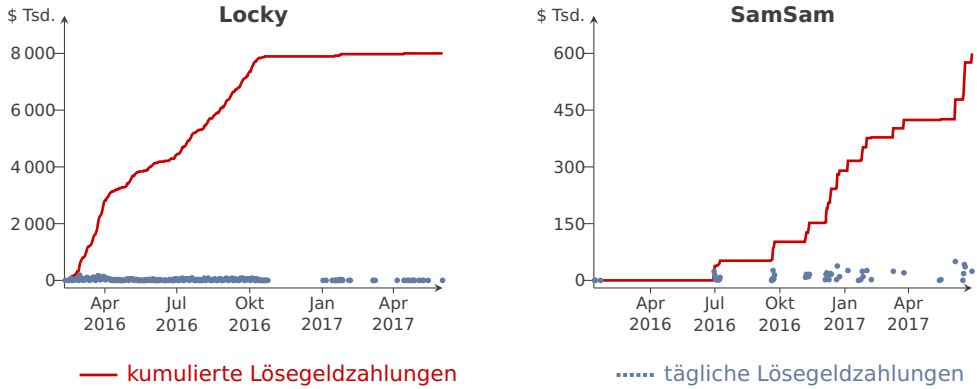
↓

Erkennung von Sammeladressen

Lösegeldzahlungen im Zeitverlauf

Locky automatische Verbreitung führt zu epidemischer Infektion und Zahlungen

SamSam gezielte Angriffe mit manueller Verbreitung



Paquet-Clouston, M. et al. Ransomware Payments in the Bitcoin Ecosystem. *Journal of Cybersecurity*, 5(1), 2019, 1–11.

Syllabus

05.03.20	1. Einführung und Grundlagen	
23.04.20	2. Infrastruktur für Blockchain-Systeme	
30.04.20	3. Transaktionslogik in Bitcoin und Ethereum	
07.05.20	Übung: Blockchain-Analyse mit BlockSci	(Martin Plattner)
14.05.20	4. Datenschutz und Sicherheit	
28.05.20	Übung: Ethereum-Programmierung mit Solidity	(Michael Fröwis)
04.06.20	5. Skalierbarkeit, Off-Chain-Transaktionen, Governance	
18.06.20	6. Wiederholung, Fragestunde	
25.06.20	Klausur	

Stand: 14. Mai 2020. Änderungen vorbehalten.