Money in the Big Data Age – Analyzing Blockchains

Seminar „Neueste Trends in Big Data Analytics“

WS 2017/18

Universität Hamburg

Referent:

Frederik The

(Doktoratsstudent, ETH Zürich)

Betreuer:

Dr. Julian Kunkel

Contents

Einleitung

Dies ist die begleitende schriftliche Ausarbeitung zu meiner Präsentation „Money in the Big Data Age“ im Seminar „Neueste Trends in Big Data Analytics.“ Die Analytics Beispiele stehen im Zusammenhang mit der Crypto Analytics API [2], die ich für das Projekt „Big Data“ verwende.

Neben zahlreichen Analytics Beispielen wird auch der Kryptomarkt als solcher sowie seine soziale- und geldtheoretische Einordnung beleuchtet, da Kryptowährungen sowohl einen sozialwissenschaftlichen, ökonomischen als auch kryptografisch-computerwissenschaftlichen Forschungsgegenstand darstellen, somit naturgemäß interdisziplinär sind und nicht isoliert betrachtet werden können ohne wesentliche Assoziationen zu vernachlässigen.

Diese Ausarbeitung ist wie folgt gegliedert: Ich beginne mit (1) der Motivation was Krypto-Währungen sind und was sie von zentralisierten Geldkonzepten unterscheidet. Es folgen unter (2) die technischen Hintergründe; in (3) werden Analytics Tools vorgestellt und in (4) entsprechende Analytics Beispiele. In (5) wird ein Blockchain Big Data Outlook auf dem Hintergrund versucht, dass es sich um eine noch extrem junge Technologie handelt und die meisten Anwendungen sich noch in der Test- und Entwicklunsphase befinden.

1. Motivation

Was ist Kryptogeld? Versuch einer Definition

Kryptogeld ist digitales "Vermögen", das Kryptografie verwendet, um (1) seine Transaktionen zu sichern, (2) die Erstellung zusätzlicher Einheiten zu kontrollieren und (3) die Übertragung von "Vermögen" zu überprüfen.

Blockchain-Tokens sind digitale „Nutzungsrechte", die Kryptografie verwenden, um (1) Transaktionen zu sichern, (2) die Erstellung zusätzlicher Einheiten zu kontrollieren und (3) die Übertragung von „Nutzungsrechten“ zu überprüfen.

Die Unterscheidung zwischen Kryptogeld und Tokens ist relevant, da der Begriff „Blockchain“-Technologie, genauer bezeichnet als ***Distributed-Ledger-Technologie*** (DLT), lediglich die technologischen Aspekte umfasst, wohingegen „Kryptogeld“ bzw. „Tokens“ Anwendungsbeispiele dieser Technologie darstellen.

Zentralisiertes vs. Dezentralisiertes Kryptogeld

Die Idee "kryptografischen Geldes" existiert bereits seit den 1980er Jahren (Chaum, 1983 [4]). Bitcoin war lediglich das erste *dezentralisierte* Konzept kryptografischen Geldes, was u. A. die enorme Popularität von Bitcoin bedingte und bis heute bedingt. Vorherige Konzepte kryptografischen Geldes beinhalteten stets eine zentrale Instanz, die die Rolle der Zentralbank lediglich auf die digitale Welt übertrugen. Bitcoin stellte das erste vollständig P2P-distribuierte System dar.

Heute gibt es 1000+ verschiedene "Krypto-Coins", wovon die meisten dezentrale Systeme verfolgen.[1]

Viele dieser Blockchains erheben keinen Anspruch darauf, ein Finanzinstrument zu sein, sondern sind eher mit Anteilen (sogenannten „Tokens“) an einer Firma oder einer speziellen Dienstleistung vergleichbar.

Zentralisiertes vs. Dezentralisiertes (Krypto)-Geld

"*[T]he main benefits are lost if a trusted third party is still required to prevent double-spending. We propose a solution to the double-spending problem using a peer-to-peer network...What is needed is an electronic payment system based on cryptographic proof instead of trust...*" (S. Nakamoto, 2009) [7]

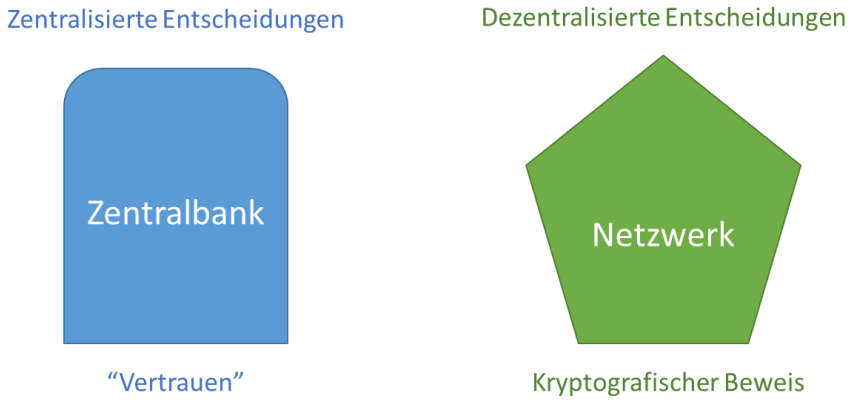
Satoshi Nakamotos Zitat macht deutlich, dass das entscheidende Element der Unterscheidung zwischen zentralisierten- und dezentralisierten Währungen von der Defintion des *Vertrauens* abgeleitet werden muss, da sich das Potential der Technologie nur entfalten kann, wenn Vertrauen durch Kryptografie geschaffen wird anstatt einer zentralen Instanz wie einer Zentralbank.

Da der Staat aus historischer Sicht die nationale monetäre Hoheit inne hat (oder im Falle der USA und Europa eine föderale Instanz diese Rolle einnimmt) lässt sich aus diesem zentralen Argument auch der „Krypto-anarchische“ Ruf herleiten, der Bitcoin in einen unlösbaren Konflikt mit den nationalen Institutionen stellt, der zur Abschaffung der Zentralbank wie wir sie kennen oder dem Scheitern von Bitcoin als (einzige) Währung führen muss.

In den letzten Jahren entstanden genau aus diesem theoretischen Konflikt heraus zahlreiche Hybrid-Konzepte und -Währungen (wie z.B. Ripple), die einen weniger drastischen Plan verfolgen und eine Integration mit dem bestehenden montären System und seinen Institutionen anstreben anstatt eine Ablösung zu verfolgen.

[1] Eine der größten Kryptowährungen nach Marktkapitalisierung, Ripple (XRP), bedient sich hindessen eines (semi-) zentralisierten Systems.

Fig 1: Zentralisiertes vs. Dezentralisiertes (Krypto)-Geld



Quelle: Autor

Drei Arten von Coins/Tokens

Prinzipiell muss man zwischen drei verschiedenen Arten von Kryptocoins/-Tokens unterscheiden:

Bitcoin lässt sich aufgrund der zuvor skizzierten Geschichte als "Genesis Blockchain" deklarieren, die den Markt für Blockchain-Technologie als solche kreierte. Es gab keinen formalen Prozess der Entwicklung oder Einführung im Falle von Bitcoin. Bitcoin existierten von heute auf morgen und Menschen fingen an es zu minen, zu transferieren und für verschiedene andere Zwecke zu nutzen. Seitdem wurde das Netzwerk selbst als auch die Technologie rundherum (2nd-Layer Technologien wie Lightning Network, etc.) ständig ausgebaut und weiterentwickelt. Bitcoin ist konzeptionell am ehesten mit "digitalem Gold" als natürlicher Rohstoff/Ressource vergleichbar.

Security Tokens stellen die zweite Kategorie dar und sind mit einer Kapitalbeteiligung an einer Organisation/Projekt/Firma zu vergleichen bzw. dem Anspruch auf zukünftig generierte Gewinne, die aus dieser Beteiligung resultieren und den Aktivitäten des jeweiligen Kryptonetzwerks zugeschrieben werden können. Nach amerikanischer Rechtsauffassung handelt es sich hierbei aller Wahrschinlichkeit nach[2] um Investitionsverträge, die dem sogenannten *Howie-Test* unterliegen, und laut dessen Kriterien als Securities deklariert und behandelt werden können.

Ether (Ethereum) könnte laut dieser Auffassung der größte Security Token nach Marktkapitalisierung sein, auch wenn solche Definitionen rückwirkend für Investoren kaum Bedeutung haben dürften. Das Thema der Definition als Security ist deshalb so brisant, da der enorme Initial Coin Offering (ICO) Boom im Jahr 2017 zu volumenstarken Kapitalinvestitionen durch Kleinanleger geführt hat, die laut US Recht nicht die Voraussetzungen für solche Investitionen in Securities besitzen und durch die unklare Rechtssprechung

[2] Da die SEC und das CFTC noch immer über die genaue Defintion und Auslegung verhandeln herrscht in diesem Fall noch keine endgültige Rechtssicherheit. Desweiteren unterliegen all diese Einstufungen nationaler Rechtssprechung und variieren dementsprechend gravierend von Land zu Land. Eine der „kryptofreundlichsten" Regulierungen der Welt bahnt sich in Malta an, weshalb der umsatzstärkste Exchange, Binance, inzwischen seinen Sitz nach Malta verlegt hat.

als auch Irreführung nicht ausreichend geschützt worden sind. Der Vorwurf der Irreführung von Investoren könnte theoretisch zu erheblichen Geld und sogar Freiheitsstrafen führen für jene DLT-Unternehmen, die in 2017 ohne Rechtsgrundlage Geld einsammelten.[3]

Utility Tokens bezeichnen den Zugang zu Dienstleistungen DLT-basierter Technologien und sind mit Nutzungsrechten für Produkte, Lizenzen für Software, In-Game Währungen oder pay-per-use SaaS Angeboten vergleichbar. Die meisten Utility Tokens basieren zurzeit noch auf Smart Contracts und bauen auf existierenden Blockchains wie Ethereum, Cardano oder Stellar auf. Viele Utility Tokens wurden allerdings auch nur während der ICO-Entwicklungsphase „on top" einer anderen Blockchain gebaut und entwickeln langfristig ihre eigene Blockchain (siehe TRON, VeChain, etc.).

Je nachdem ob Utility Tokens durch einen Airdrop verschenkt wurden oder von Nutzern der Dienstleistungen für Fiatgeld erworben werden mussten ist die logische Trennung zu Security Tokens schwierig zu ziehen und stellt eine Grauzone dar. Viele der heute am Markt existierenden Utility Tokens könnten wahrscheinlich auch als Securities definiert werden trotz ihres eindeutigen Dienstleistungscharakters wenn die Nutzer die Tokens mit Fiatgeld erwerben mussten.

2. Technische Hintergründe

Elemente dezentralisierten Kryptogeldes

Im Grunde genommen basieren alle dezentralisierten Kryptowährungen auf drei Elementen:

Das erste ist ein **Zustandsübergangssystem**, das Block an Block reiht, in die alle Transaktionen geschrieben werden und somit unveränderbar für alle Zeiten festgehalten werden. Dies bedingt die ***„Immutability"-Eigenschaft*** von DLT Technologien.

Das zweite Element ist ein **Konsessystem**. Die meistverwendeten Konsessysteme heutzutage sind Proof-of-Work (PoW) und Proof-of-Stake (PoS). Sowohl Bitcoin als auch Ethereum nutzen bis heute ein PoW System, auch wenn Ethereum und viele andere Blockchains aus Energie-Effizienzgründen vermehrt auf PoS umstellen. Da es sich bei PoS um ein spieltheoretisches Konsessystem handelt, das keinerlei ASIC-Mining-Hardware benötigt, ist die Umstellung ein Politikum, da beträchtliche Hardware Investitionen der Miner-Community durch Staking wertlos werden.

Proof-of-Work setzt sich zusammen aus zwei logischen Schritten:

1. Finde eine kryptografische Nonce, die einen Hash (double SHA-256 für Bitcoin) unterhalb eines Zielwertes Z liefert, dem Schwierigkeitsgrad des Systems zum Zeitpunkt t.
2. Der Schwierigkeitsgrad selbst wird anhand eines gleitenden Mittelwertes basierend auf der Anzahl der durchschnittlichen Blöcke pro Stunde bestimmt (siehe Fig 2 Appendix).

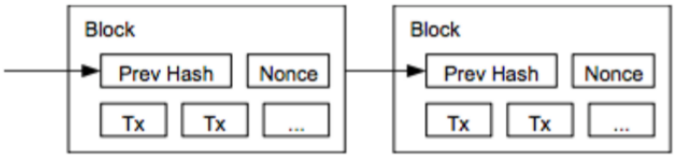
Das Konsessystem ist wohl das bedeutsamste Element, da es die ***„Trust"-Eigenschaft*** von DLT Technologie garantiert ohne eine übergeordente Instanz wie eine Zentralbank oder einen Market-Maker zu benötigen.

[3] In diesem Sinne hat der ICO Boom des Jahres 2017 auch eine Diskussion über neue Formen der Kapitalbeschaffung für Firmen ausgelöst.

Konsenssysteme wie PoW und PoS stellen die entscheidende Neuerung dar, die vollständig dezentralisierte P2P Systeme zurück in das Licht der Öffentlichkeit gerückt haben nachdem P2P file sharing Anfang der 2000er an rechtlichen Hürden scheiterte.

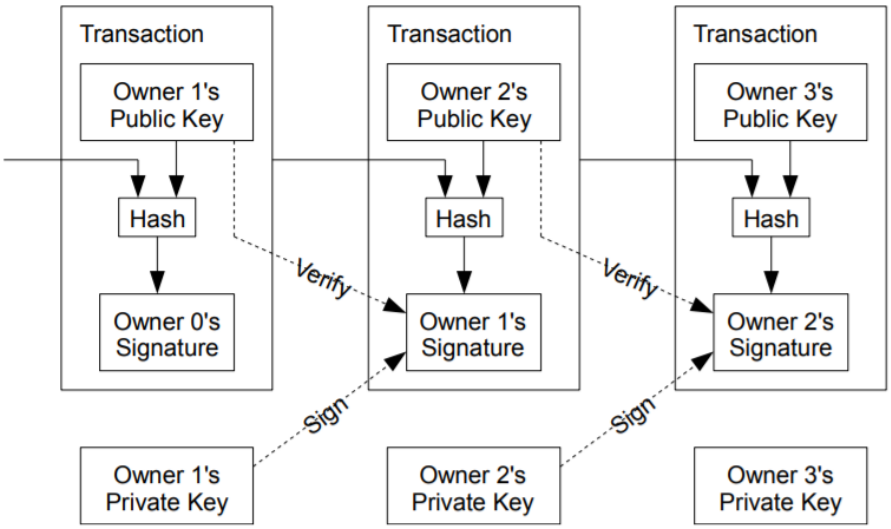
Das Letzte Element ist ein **Timestampserver**, der für nicht anderes verantworltlich ist als einen linearen verlauf der Zeit zu garantieren, so dass die gesamte Historie unveränderlich auf der Blockchain festgehalten werden kann. Das zeitlich lineare Aneinanderreihen von Blöcken ermöglicht eine einheitliche „Geschichtsschreibung".[4] Dies lässt sich als ***„Memory"-Eigenschaft*** von DLT Technologien beschreiben (The, 2014 [11]).

Fig 2: Blocks



Quelle: "Bitcoin: Peer-to-Peer Electronic Cash System", S. Nakamoto (2009) [7]

Fig 3: Transaktionen (Zustandsübergangssystem)



Quelle: "Bitcoin: Peer-to-Peer Electronic Cash System", S. Nakamoto (2009) [7]

[4] Genau diese *Memory*- und *Immutability* Eigenschaften von DLT-Technologie stehen nun in Europa ab Mai 2018 in Konflikt mit der General Data Protection Regulation (GDPR, www.eugdpr.org). Diese lässt zwar eine Menge Interpretationsspielraum offen, betrachtet jedoch selbst gehashte Adressen und Keys als PII, was das Validieren von Blockchains sowie das Betreiben einer Full Node in einen neuen rechtlichen Kontext setzen und für zahlreichen Blockchain Anwendungen in Europa eine enorme rechtliche Hürde bedeuten könnte.

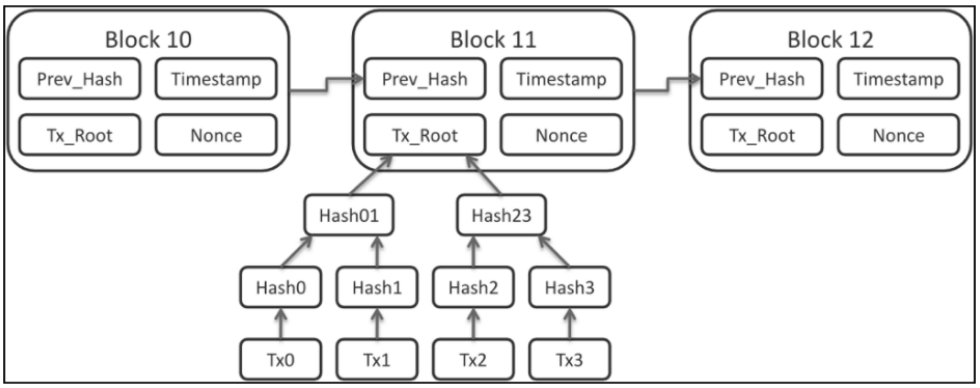
Blockchain Algorithmus (Archetyp)

Zusammengenommen lässt sich aus dem zuvor erklärten folgender generischer Blockchain Algorithmus herleiten:

1. Prüfe ob der vorherige Block existiert und valide ist
2. Prüfe, dass Timestamp des Blocks zum Zeitpunkt *t > t-1*
3. Prüfe, dass Proof-of-Work für den Block valide ist
4. *S[0]* sei der Zustand am Ende des vorherigen Blocks
5. Angenommen *TX* ist die Liste aller *n* Transaktionen des Blocks. Für alle *i* in *0 ... n-1*, prüfe dass *S[i+1] = APPLY (S[i], TX[i])*. Im Falle eines Fehlers EXIT und return FALSE
6. Ansonsten TRUE und registriere *S[n]* als den Status am Ende des jetzigen Blocks

Nachdem der grundlegende Algorithmus skizziert ist noch einen Kommentar zur Skalierung und Praktikabilität von Blcokchains: Da jeder Block einer Blockchain theoretisch Tausende und Abertausende von Transaktionen enthält wird es sehr zeitaufwendig sein, alle Daten in jedem Block als eine Reihe zu speichern. Dies würde das Finden einer bestimmten Transaktion extrem umständlich und zeitaufwendig gestalten. Merkle-Bäume bieten hier eine Lösung. Sie reduzieren die Zeit, die benötigt wird, um herauszufinden, ob eine bestimmte Transaktion in einen Block gehört oder nicht. Anstatt den mühsamen Prozess zu durchlaufen, um jeden einzelnen Hash zu betrachten und zu sehen, ob er zu den Daten einer bestimmten Transaktion gehört oder nicht, lässt er sich verfolgen indem man der Spur der Hashes im Merkle Baum folgt, die zu den Daten der jeweiligen Transaktion führt.[5]

Fig 4: Vom Block zur Blockchain



Quelle: Blockgeeks, https://blockgeeks.com/guides/what-is-hashing/

[5] Siehe Appendix Fig 1a und 1b um die genaue Blockstruktur einzusehen.

3. Analytics Tools

Im Appendix Fig 3 lässt sich erkennen, dass auch wenn die auf Blockchains gespeicherte Datenmenge rasant wächst man streng genommen noch nicht von „Big Data“ sprechen kann. Nichtsdestotrotz muss man sagen, dass Blockchain Technologie noch in den Kinderschuhen steckt und die meisten kommerziellen, sozialen und privaten Anwendungen basierend auf DLT-Technologie erst 2018 ihre Mainnets lancieren und sich noch in der Testphase befinden (dies gilt für Bitcoin Lighnting, Ethereum Metropolis, TRON Exodus, VeChain Thor, Cardano, etc.)

Folgende Analytics Tools existieren heutzutage u. A.:

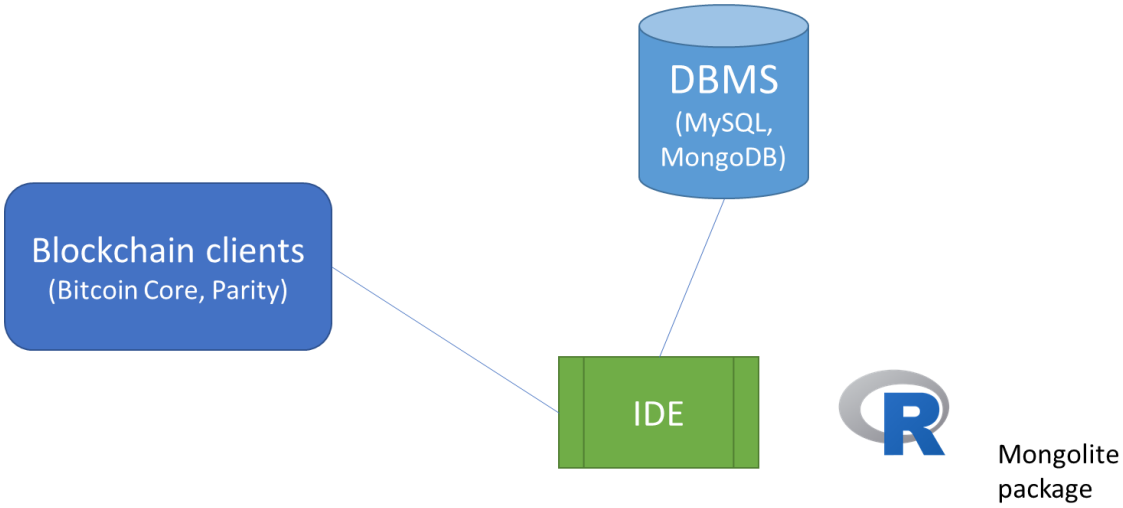
Tabelle 1: Überblick Analytics Tools/APIs

Blockchain Explorer APIs	**Trading Platform APIs**	**Custom APIs**
Blockexplorer	Kraken	“Crypto Analytics API” (Project Scala-MongoDB)
Etherscan	Coinbase	Andere Github repos
blockchain.info		
Coindesk		

Die von mir im Projekt eingesetzte „Crypto Analytics“ API [1] hat die Besonderheit, dass sich Daten aus der Blockchain mit externen Daten verbinden lassen. Die API setzt sich zusammen aus dem jeweiligen Blockchain client (Bitcoin Core in meinem Fall[6]), einem DBMS wie MongoDB oder MySQL (MongoDB in meinem Fall). Ein IDE wie IntelliJ erleichtert mir die Arbeit und das Mongolite package in R erlaubt es mir, R’s Data Analytics packages zu verwenden. Die schematische Darstellung folgt unten.

[6] Eine Idee, die ich dieses Jahr beabsichtige in meinen Doktoratsforschungsplan miteinzubringen, ist die Anzahl der Clients zu erweitern und die TRON Blockchain als Studienobjekt zu verwenden. Dies ist besonders interessant, da es sich bei TRON, um eine *dezentrale* Social Media und Content Sharing Plattform handelt. Mit Blick auf die Datenskandale aus der jüngsten Vergangenheit bei *zentralisierten* Social Media Plattformen wie Facebook erhoffe ich mir hiervon einen vielversprechenden Datensatz für mein Promotionsthema „Die Monetarisierung persönlicher Daten“.

Fig 5: Schematische Darstellung Project Scala-MongoDB



Quelle: Autor

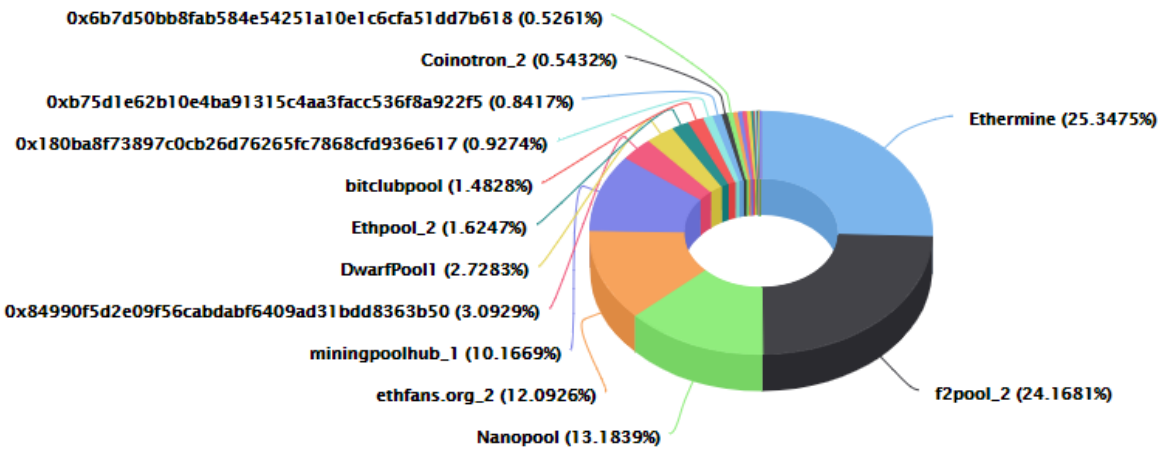
4. Analytics Beispiele

Ethereum Top 25 Miners by Blocks

Aus der Grafik wird klar, dass PoW Mining ein zentralisiertes Geschäft darstellt. Große Miningpools, in denen Miner ihrer Ressourcen bündeln und Gewinne untereinander aufteilen, dominieren das Geschehen (*Staking* mittels PoS wird dies nicht ändern). In einem gewissen Sinne führt diese Konzentration, die bei Bitcoin übrigens identisch ist, zu einer entdemokratisierung des eigentlich vollständig dezentralen Konzeptes, welches das Konsessystem Proof-of-Work darstellen soll. Satoshi Nakamoto's Motto „*One CPU, one vote*" wird hierdurch allerdings nicht zwangsläufig relativiert. Andere Crypto-Währungen wie Decred, dessen Community glaubt, das Bitcoin an seinen eigenen Maßstäben gescheitert sei, haben alternative Konzepte entwickelt.[7]

[7] Siehe https://www.decred.org/

Fig 6: Ethereum Top 25 Miners by Blocks (letzten 7 Tage)

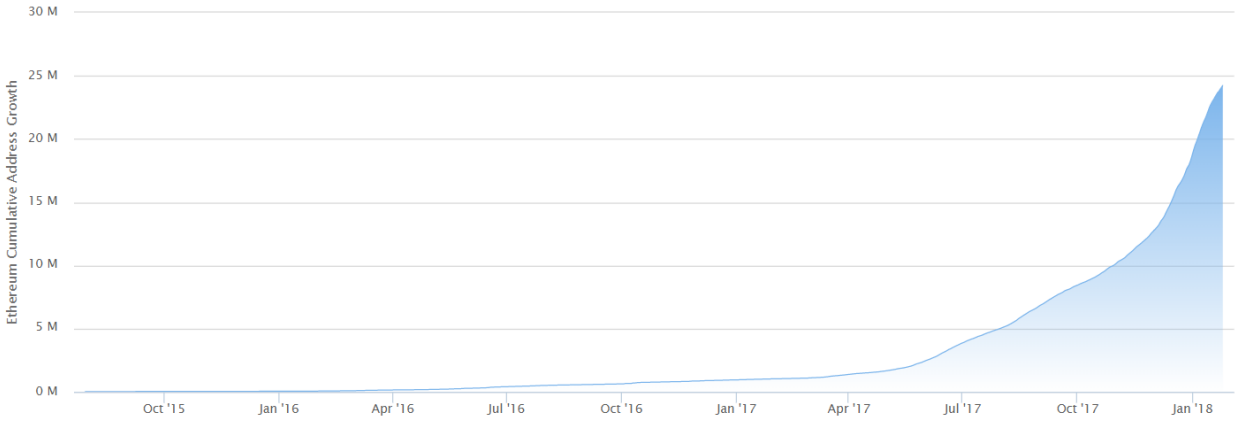


Quelle: EtherScan.io, https://etherscan.io/stat/miner?range=7&blocktype=blocks

Ethereum Adresswachstum

Die Grafik unten zeigt den enormen Adresszuwachs von Ethereum, die größte Blockchain für Smart Contracts, ohne die der Initial Coin Offering Boom des Jahres 2017 nicht möglich gewesen wäre. Etherum bot die Infrastruktur, um die Vielzahl an Geschäftsmodellen und Startup Unternehmen zu beherbergen, ohne dass diese gezwungen waren, eine eigene Blockchain-Technologie zu entwickeln.

Fig 7: Ethereum Adresswachstum



Quelle: EtherScan.io, https://etherscan.io/chart/address

5 Größten Bitcoin Adressen in BTC

Die „reichesten" Bitcoin Adressen lassen sich mit Hilfe der oben genannten Explorer sehr leicht identifizieren.[8] Die respektiven Adressen besitzen ca. 0.5 – 1.2 Prozent aller Bitcoins. Somit handelt es sich um Werte die im Jahr 2018 ca. eine Milliarde USD betragen. Interessanterweise ist die viert größte Adresse mit BTC im Wert von 994 Millionen USD eine reguläre Binance Wallet, ohne einen physisch getrennten kryptografischen Kern (*cold storage*). Mit Blick auf die Häufigkeit von Hacker Angriffen und Binance eigenen technischen Debakel in der Vergangenheit[9] ist dies hervorzuheben.

Tabelle 2: Die 5 reichsten Bitcoin Adressen

Address	Balance ▵1w/▵1m	% of coins	First In	Last In	Number Of Ins
3D2oetdNuZUqQHPJmcMDDHYoqkyNVsFk9r wallet: Bitfinex-coldwallet	193,349 BTC ($1,923,024,641 USD) +1589 BTC / +2500 BTC	1.14%	2017-01-05 13:34:15	2018-05-05 05:24:47	4893
16rCmCmbuWDhPjWTrpQGaU3EPdZF7MTdUk wallet: Bittrex-coldwallet	132,203 BTC ($1,314,873,357 USD) -5000 BTC / -10000 BTC	0.7775%	2016-02-27 19:00:09	2018-04-30 17:55:36	143
3Nxwenay9Z8Lc9JBiywExpnEFiLp6Afp8v wallet: Bitstamp-coldwallet	103,848 BTC ($1,032,860,611 USD) +0.01 BTC / +4000 BTC	0.6107%	2015-10-16 16:43:06	2018-05-04 20:23:20	181
16ftSEQ4ctQFDtVZiUBusQUjRrGhM3JYwe wallet: Binance-wallet	99,947 BTC ($994,059,926 USD)	0.5878%	2017-12-08 08:51:10	2018-04-22 00:29:53	101
3Cbq7aT1tY8kMxWLbitaG7yT6bPbKChq64 wallet: Huobi-wallet	98,041 BTC ($975,106,969 USD) / +5694 BTC	0.5766%	2017-09-08 18:41:05	2018-04-26 04:53:53	174

Quelle: Bitinfocharts.com, https://bitinfocharts.com/top-100-richest-bitcoin-addresses.html

Extrahieren von Bitcoin Metadata

Mit Hilfe der Crypto Analytics API [2] lassen sich zum Beispiel die Anzahl der Transaktionen pro Bitcoin Protokoll analysieren. Der Query sowie die Grafik sind hier abgebildet.

Die Query startet bei Block 290000, da Metadaten in der *OP_Return* Variable erst ab Bitcoin Core Release 0.9.0 verfügbar waren. Sie iteriert über jeden Block und fügt eine neues Dokument in die Mongo Collection, wenn es sich um eine OP_Return Transaktion handelte. Die *OpReturn.getApplication* Methode der API ermöglicht es den Klarnamen des Protokolls hinzuzufügen.

Colu ist ein Protokoll für ein Community Payment System. Offensichtlich weist allerdings selbst das größte Bitcoin Protokoll („Smart Contract") eine sehr geringe Anzahl an Transaktionen auf mit einem totalen Volumen von weniger als 250 tausend Transaktionen. Dies ist nicht verwunderlich, da Bitcoins Smart

[8] Siehe: https://bitinfocharts.com/top-100-richest-bitcoin-addresses.html
[9] Siehe Wood, A. (07-03-2018) „Possible Hack Of Third-Party Tools Affects Binance Exchange Users ", https://cointelegraph.com/news/possible-hack-of-third-party-tools-affects-binance-exchange-users.

Contract Möglichkeiten sehr beschränkt sind, was u. A. den enormen Erfolg von Ethereum bedingt hat, welches als erstes diese Möglichkeit erkannt und addressiert hat.

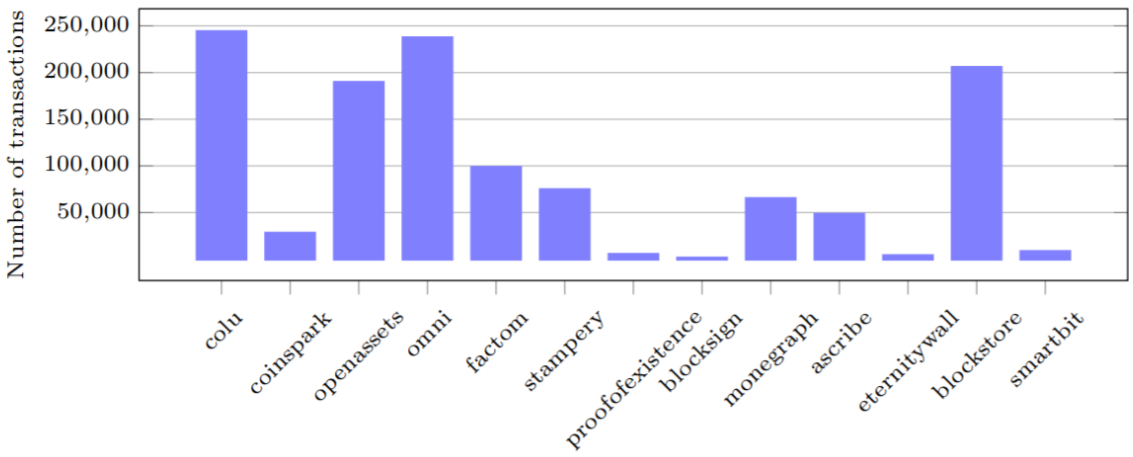
Query 1: Transaktionen pro Bitcoin Protokoll

```
val opReturnOutputs = new Collection("opReturn", mongo)

blockchain.start(290000).end(473100).foreach(block => {
    block.bitcoinTxs.foreach(tx => {
        tx.outputs.foreach(out => {
        if(out.isOpreturn()) {
          opReturnOutputs.append(List(
          ("txHash", tx.hash),
          ("date", block.date),
          ("protocol", OpReturn.getApplication(out.outScript.toString)),
          ("metadata", out.getMetadata())
          ))
        }
        })
    })
})
```

Quelle: Bartoletti et al. (2017) [2]

Fig 8: Transaktionen pro Bitcoin Protokoll



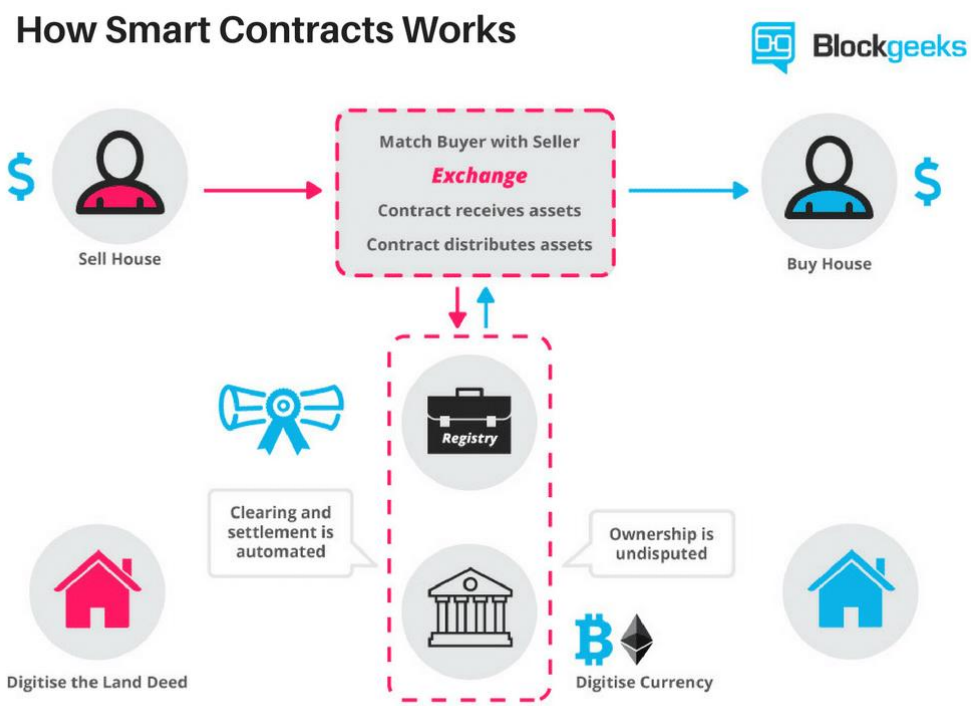
Quelle: Bartoletti et al. (2017) [2]

Smart Contracts (Decentralized Applications)

Smart Contracts, auch als ***Decentraliced Applications*** oder kurz "DAPPs" bezeichnet, stellten im DLT-Bereich die erste große Innovation seit Proof-of-Work als Konsenssystem dar. Wie bereits beschrieben führte Vitalik Buterin der Ethereum Foundation massgeblich diesen Trend an.

Smart Contracts können in beliebiegen Programmiersprachen (Solidity, Java, C#, etc.) programmiert werden (dies hängt von der jeweiligen Blockchain selbst ab) und sind im Grunde nichts anderes als ein simples Regelwerk, das grundlegende Geschäftsoperationen, Entscheidungen und Konsequenzen formal hard-coded und somit in der Lage ist, eine existierende Dienstleistung wie einen Notar-Service (siehe Fig. 9) in die DLT-Welt zu überführen. Das Elegante an der Lösung ist, dass der Contract selbst den Mittelsmann/Market-Maker und somit jegliche zentrale Instanz überflüssig macht. Ob Social Media Platform, Notar oder Zentralbank. DLT-Technologie besitzt theoretisch das Potential, Facebooks Dienstleistungen, den Notar sowie den Zentralbänker überflüssig zu machen.

Fig 9: Beispiel Smart Contract - Notarservice



Quelle: https://blockgeeks.com/guides/smart-contracts/

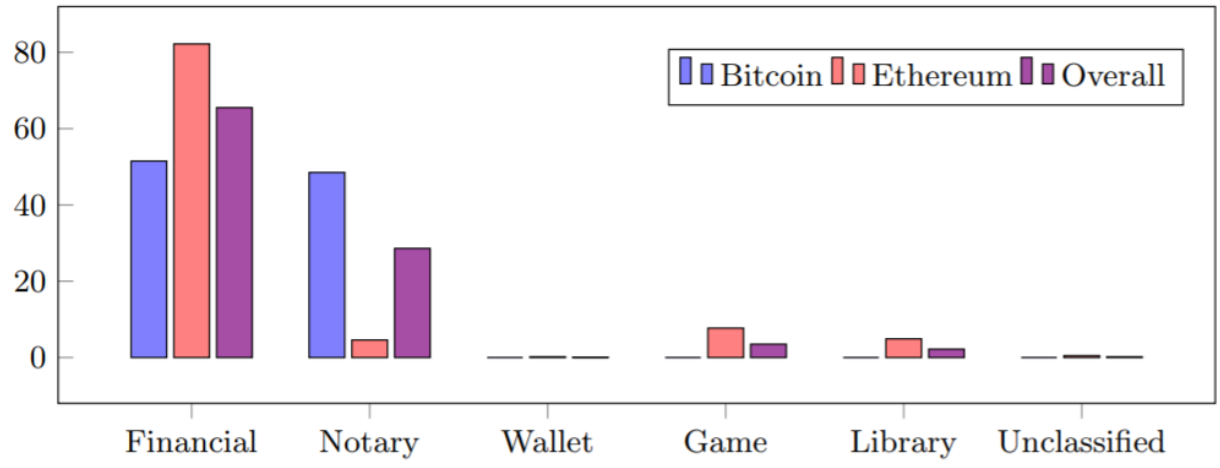
Smart Contracts lassen sich in verschiedene Kategorien klassifizieren, wobei Finanzdienstleistungen den Löwenanteil ausmachen. Die Statistiken in Tabelle 3 und Fig 10 zeigen abermals, dass Ethereum im Bereich Smart Contracts der Marktführer ist verglichen mit Bitcoin.

Tabelle 3: Smart Contract Typen – Bitcoin vs. Ethereum

Category	Platform	Contracts	Transactions
Financial	Bitcoin	6	470,391
	Ethereum	373	624,046
Notary	Bitcoin	17	443,269
	Ethereum	79	35,253
Game	Bitcoin	0	0
	Ethereum	158	58,257
Wallet	Bitcoin	0	0
	Ethereum	17	1,342
Library	Bitcoin	0	0
	Ethereum	29	37,034
Unclassified	Bitcoin	0	0
	Ethereum	155	3,679
Total	Bitcoin	23	913,660
	Ethereum	811	759,611
	Overall	834	1,673,271

Quelle: Bartoletti et al. (2017), Table 2 [1]

Fig 10: Smart Contract Typen – Bitcoin vs. Ethereum



Quelle: Bartoletti et al. (2017), Fig. 2 [1]

Transaktionsgebühren

Die Query 2 unten iteriert über alle Blocks und kreiert eine neue Mongo Collection die die Transaktionsgebühren als Differenz aus Input und Output Summe errechnet. Die Boolean Variable in Line 1 ist true, da hier ein Deep Scan der Blockchain vorgenommen wird.[10]

Query 2: „*Whale*"-Transaktionen

```
1  val blockchain = BlockchainLib.getBitcoinBlockchain(new BitcoinSettings("user","
       password","8332",MainNet,true))
2  val mongo = new DatabaseSettings("myDatabase", MongoDB, "user", "password")
3  val txWithFees = new Collection("txWithFees", mongo)
4
5  blockchain.end(473100).foreach(block => {
6      block.bitcoinTxs.foreach(tx => {
7          txWithFees.append(List(
8              ("blockHash", block.hash),
9              ("txHash", tx.hash),
10             ("fee", tx.getInputsSum() - tx.getOutputsSum()),
11             ("date", block.date),
12             ("rate", Exchange.getRate(block.date))
13         ))
14     })
15 })
```

Quelle: Bartoletti et al. (2017), Fig. 9 [2]

Trotz eines arithemtischen Mittelwertes von lediglich 0.41 USD und einer recht hohen Standardabweichung von 12.09 USD lassen sich eindeutig sogenannte „*Whale Transactions*" ausmachen, in denen die Miner bis zu 136 tausend USD verdienen konnten, wenn sie die entsprechende Transaktion validierten und in den nächsten Block einfügten.

Tabelle 4: „*Whale*"-Transaktionen

Fee (USD)	Date	Transaction hash
136243.37	2016-04-26 14:15:22	cc455ae816e6cdafdb58d54e35d4f46d860047458eacf1c7405dc634631c570d
56493.50	2017-01-04 20:01:28	d38bd67153d774a7dab80a055cb52571aa85f6cac8f35f936c4349ca308e6380
39502.15	2017-05-31 14:28:51	cb95ab3aef378c14bc59d0db682d96202b981c1f8fad7d66e23e0be06f2a00c4
25095.71	2017-05-31 14:28:51	8e12a1aba87e4657f5fabec1121ed57f706805ad6d4ffe88c6fce78596bd9b75
23518.00	2013-08-28 10:45:17	4ed20e0768124bc67dc684d57941be1482ccdaa45dadb64be12afba8c8554537

Quelle: Bartoletti et al. (2017), Fig. 9 [2]

[10] Während die Werte von Outputs explizit in Transaktionen gespeichert werden, ist dies im Flle von Inputs nicht der Fall: Um sie zu erhalten muss man von einem vergangenen Block die Transaktion abrufen, die durch den Input eingelöst wurde. Dies kann durch einen "deep" Scan der Blockchain erreicht werden. Die Crypto Analytics API [2] beinhaltet diese nützliche Funktion.

Bitcoin Preisvorhersagen

Preisvorhersagen (Forecasting) ist ein sehr populäres Thema an Finanzmärkten generell und insbesondere im Falle von Crypto-Währungen, da diese die Volatilität regulärer Finanzmärkte um ein Vielfaches übersteigen.

Madan, I., Saluja, S., & Zhao, A. (2015 [6]) haben in einer Studie die unten aufgeführten 16 Features benutzt, um den Preis von Bitcoin vorherzusagen. Anhand der Features wurde ein binomialer Klassifikationsalgorithmus entwickelt, um die Vorzeichenänderung im Bitcoin-Preis basierend auf täglichen Datenpunkten vorherzusagen. Der Datensatz bestand aus 16 Variablen, die über einen Zeitraum von 5 Jahren täglich beobachtet wurden. Der Train-Test Split lag bei 70-30 Prozent.

Die Forscher fanden, dass 10-Minuten-Intervalle besser als 10-Sekunden-Intervalle für Presivorhersagen funktionieren, was für jeden Crypto-Trader nicht verwunderlich sein dürfte, da die enorme Volatilität auf Krypto-Börsen vor allem durch einen Mangel an Liquidität und das Fehlen institutioneller Investoren zuruckzuführen ist, was Krypto-Märkte und insbesondere einzelne Börsen und Marktplätze zum Spielball von FUD-, Hype-Phasen und sogar Marktmanipulationen macht.

Tabelle 5: Features für Preisvorhersagen

FEATURE	DEFINITION
Average Confirmation Time	Ave. time to accept transaction in block
Block Size	Average block size in MB
Cost per transaction percent	Miners revenue divided by the number of transactions
Difficulty	How difficult it is to find a new block
Estimated Transaction Volume	Total output volume without change from value
Hash Rate	Bitcoin network giga hashes per second
Market Capitalization	Number of Bitcoins in circulation * the market price
Miners Revenue	(number of BTC mined/day * market price) + transaction fees
Number of Orphaned Blocks	Number of blocks mined / day not off blockchain
Number of TXN per block	Average number of transactions per block
Number of TXN	Total number of unique Bitcoin transactions per day
Number of unique addresses	Number of unique Bitcoin addresses used per day
Total Bitcoins	Historical total Number of Bitcoins mined
TXN Fees Total	BTC value of transaction fees miners earn/day
Trade Volume	USD trade volume from the top exchanges
Transaction to trade ratio	Relationship of BTC transaction volume and USD volume

Quelle: Madan, I., Saluja, S., & Zhao, A. (2015) [6]

Die Resultate im Detail zeichnen allerdings ein noch deutlicheres Bild dieser Situation: Während ein (simpleres) GLM und ein (komplexerer) Random Forest in Tabelle 6 ähnlich gute Resulate bei einem 1-tägigen Lookback-Window liefern ist die Vorhersage auf Basis von 10-minütigen Intervallen nur minimal besser als ein Münzwurf. Dies zeigt eindeutig wie volatil Crypto-Währungen in diesen kurzen Zeitfenstern

sind und dass eine Preisvorhersage aller Wahrscheinlichkeit nach in diesen Zeitintervallen auch nicht sinnvoll ist (zumindest nicht zum jetzigen Zeitpunkt).

Table 6: Vorhersagen auf Basis von 1-tägigen Intervallen (oben) vs. 10-minütigen Intervallen (unten)

STATISTIC	BINOMIAL GLM	SVM	RANDOM FOREST
Sensitivity (TPR)	0.9790	0.0348	1.0
Specificity (TNR)	0.9939	0.5514	0.9392
Precision (PPV)	0.9790	0.0839	0.7762
Accuracy (ACC)	0.9879	0.2716	0.9498

STATISTIC	10 SECOND GLM	10 MINUTE GLM	10 MINUTE RANDOM FOREST
Sensitivity (TPR)	0.5429	0.524	0.540
Specificity (TNR)	0.577	0.576	0.619
Precision (PPV)	0.574	0.551	0.581
Accuracy (ACC)	0.085	0.539	0.574

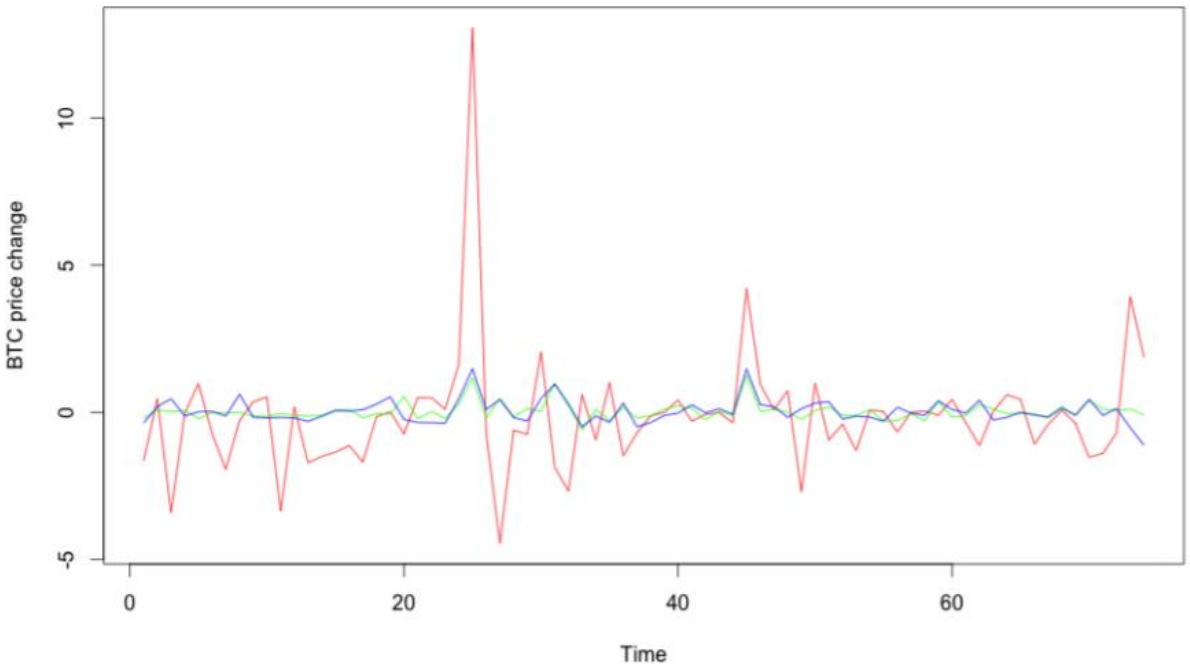
Quelle: Madan, I., Saluja, S., & Zhao, A. (2015) [6]

Auch in Fig 11 lässt sich erkennen, dass beiden Algorithmen - GLM und RF - gut abschneiden, jedoch nicht Noise und Schwankungen in den tatsächlichen Bitcoin-Preisänderungen erfassen, was der Grund für die meisten Testfehler sein dürfte. Die Forscher finden signifikante Hinweise auf Preis-Clustering bei runden Zahlen, ohne dass ein signifikantes Muster bei den Renditen nach runden Zahlen existiert. Aus der Sicht eines Traders ist auch dies zu erklären, da wiederum keine professionellen Trader sondern Kleinanleger einen Großteil der Nutzer auf Plattformen wie Coinbase und Binance, etc., ausmachen. Erst im Jahr 2018 haben Institutionen der Wall Street offiziel damit begonnen, sich mit Kryptowährungen zu beschäftigen, so dass signifikantes institutionelles Geld in naher Zukunft in Kryptomärkte fließen könnte, was die beschriebenen Dynamiken verändern könnte.[11]

Die Forscher zeigen auch, dass Preis und Volumen eine signifikante positive Beziehung zu Preisclustern haben. Desweiteren schlagen die Autoren eine Verbesserungen ihres Modells vor indem sie nicht auf tatsächlichen Daten, sondern auf einer Abstraktionsebene, die durch K-means-Clustering von jeweils 100 Datenpunkten erreicht werden soll, Vorhersagen treffen. Diese "Muster/Cluster" sollen dann als Eingabe für das ursprüngliche RF- und GLM-Prognosemodell dienen.

[11] Hugh Son , Dakin Campbell , and Sonali Basak (December 21, 2017), “Goldman Is Setting Up a Cryptocurrency Trading Desk” https://www.bloomberg.com/news/articles/2017-12-21/goldman-is-said-to-be-building-a-cryptocurrency-trading-desk

Fig 11: Tatsächliche Preise vs. Vorhersage



Quelle: Madan, I., Saluja, S., & Zhao, A. (2015) [6]

Clustering Addressen/Transaktionen

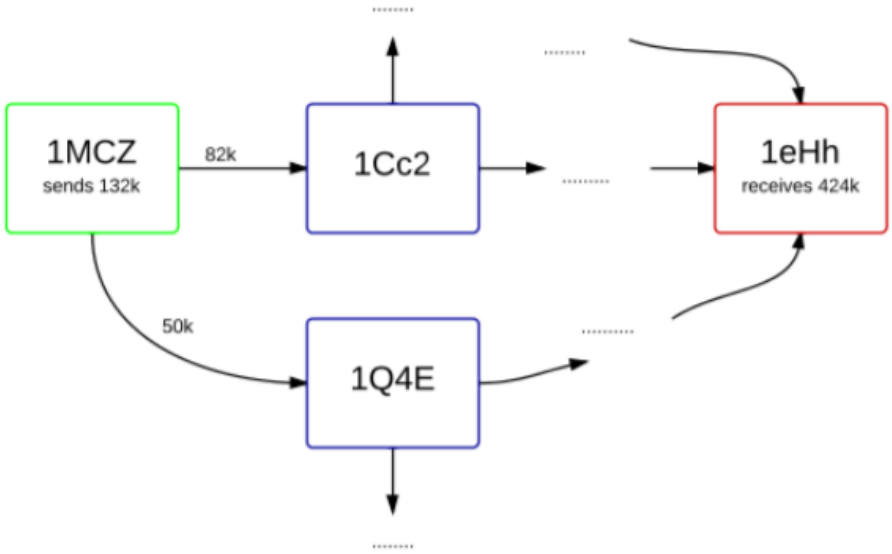
Hirshman et al. (2013 [5]) haben sich mit dem Clustering von Adressen und Transaktionen im Bitcoin Netzwerk beschäftigt. Zu diesem Zweck verwenden sie K-means and RolX Algorithmen. Die Forscher finden „Anomalien", die auf Geldwäsche hindeuten, da viele kleine Bitcoin Transaktionen am Ende auf eine einzige Receiving-Adresse gebündelt werden, was auf eine Verschleierungs-Taktik hinweisen könnte. Die Forscher können ihre Hypothese allerdings nicht weiter belegen und eine Alternative Erklärung könnten Sicherheitsaspekte sein im Falle, dass eine einzelne Transaktion fehlschlägt.

Die Anonymität, die eine weitere wichtige Eigenschaft des Hashings von Adressen im Bitcoin Netzwerk darstellt, ist je nach Rechtsgrundlage im jeweiligen Land ein Problem oder nicht. Der NSA hat laut Berichten[12] an einem System zur Deanonymisierung von Bitcoin Transaktionen und Adressen gearbeitet sich dabei aber auf das physische Layer-1 wie Glasfaser konzentriert und Layer-2 Daten wie MAC Adressen direkt entnommen, so dass Bitcoin Adressen deanonymisiert werden konnten. Interessanterweise ist Hashing laut General Data Protection Regulation (GDPR, www.eugdpr.org) in Europa ab Mai nicht

[12] siehe: Alexandre, A. (Mar-21, 2018), https://cointelegraph.com/news/us-national-security-agency-develops-system-to-identify-bitcoin-users-say-leaked-docs

„anonym“ genug, da GDPR das „Right to be forgotten“ für jeden EU Bürger einführt, was aus technischer Sicht schwierig umzusetzen scheint und mit DLT-Technologie im Widerspruch steht.

Fig 12: Indirekte Transaktionen als „Anomalien“



Quelle: Hirshman et al. (2013), Fig. 3 und Fig. 6. [7]

5. Blockhain’s Big Data Outlook

Wenn man „Big Data“ mit dem von Gartner[13] geprägten „Triple-VVV“ bemisst, dann haben die heute großen Blockchains wie Bitcoin, Ethereum, TRON, etc., noch nicht das *Volumen*, die *Velocity* oder die *Variety* erreicht, die heute für Big Data Anwendungen üblich sind.[14] Ein viertes „V“, das manchaml in diesem Zusammenhang erwähnt wird – die sogenannte *Veracity* – wird von Blockchains allerdings besser erfüllt als von irgendeiner anderen Big Data Datenquelle heutzutage, da Veracity durch das Konsessystem ein integraler Bestandteil der DLT-Technologie selbst ist. Veracity ist allerdings auch ein normatives Konzept, wohingegen Volume, Velocity und Variety deskriptive Kriterien darstellen.

Auch wenn die TRON Blockchain laut White Paper bereits 1500 Transaktionen pro Sekunde verarbeiten kann[15] stehen nahezu alle Blockchains heute noch vor einem Skalierungsproblem, das einem Big Data

[13] Siehe: https://www.gartner.com/it-glossary/big-data
[14] Es ist Stand heute immer noch möglich mehrere Blockchains vollständig auf einen handelsüblichen Laptop mit 500GB Festplatte zu speichern.
[15] TRON White Paper, https://o836fhe91.qnssl.com/tron/whitebook/TronWhitepaper_en.pdf

Datenvolumen, wie es bei indsutriellen Anwendungen üblich ist, im Wege steht, selbst wenn die Nutzerbasis bestehen sollte. Da DLT-Technologie allerdings erst langsam ins Licht einer „Main Stream" Anwendung rückt und noch in den Kinderschuhen steckt ist dies auch nicht verwunderlich. Im Folgenden wird ein Überblick über gängige Skalierungs-Lösungsansätze gegeben, die, falls erfolgreich, ein zukünftiges Transaktionsvolumen von 1000+ Tx/sec ermöglichen würden und die Größe einzelner Blockchains deutlich erhöhen könnten (100+ TB).

Skalierbarkeit

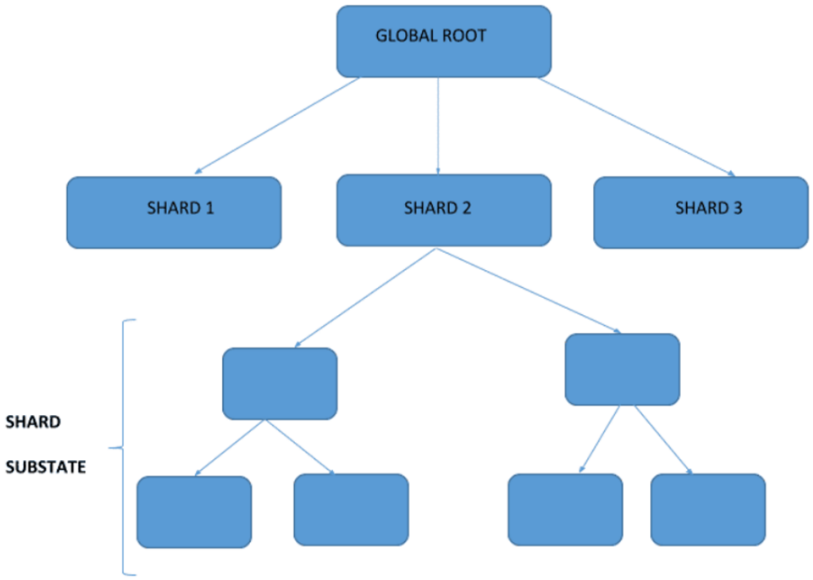
Der erste Ansatz zum Skalieren von DLT- Technologie ist das klassische ***Scale-up***. Hier wird die Blockgröße einfach erhöht, was dazu führt, dass mehr Transaktionen in den einzelnen Block geschrieben werden können. Bitcoin Cash (BCC) hat diesen Ansatz verfolgt und hat sich somit durch *hardforking* von Bitcoin (BTC) abgespalten. Das Problem beim Scale-up Ansatz ist, dass die grundlegende Problematik lediglich aufgeschoben wird, da auch ein 8 MB Block mit ca. 60 Tx/sec wie es der Fall bei Bitcoin Cash ist irgendwann zu klein wird und den Nachteil mit sich bringt, dass 8 MB große Blöcke die Chain auch acht mal so schnell wachsen lassen, was mehr Speicherplatz für jede einzelne Node bedeutet.[16]

Der zweite Ansatz zur Skalierung ist das sogennante ***Scale-out*** – auch Parallelisierung genannt. Scale-out ist kein neues Konzept und exitsiert seit den 90er Jahren unter dem Begriff *Sharding*. Beim Sharding werden die Validatoren fragmentiert, so dass nicht das gesamte Netzwerk alle Transaktionen validiert sondern lediglich jene, die ihrem Shard zugeordnet werden. Scale-out bedeutet somit eine vertikale Partitionierung der Blockchain und ist daher ein ***On-Chain*** Ansatz, da die Daten auf der Blockchain in fragmentierter Form zu jedem Zeitpunkt verbleiben (Ethereum und Telegram[17] verfolgen diesen Ansatz massgeblich).

[16] *Merge Mining* is ein weiterer Ansatz, der Throughput durch einen Faktor *N* erhöhen könnte, allerdings auch das Problem hat, dass sich Speicher- und Rechenleistung jeder einzelenen Node um den Faktor *N* erhöhen. Merge Mining ist zudem nur auf zwei Blockchains anzuwenden, die denselben Mining-Algorithmus verwenden (z. B. Bitcoin und Namecoin)

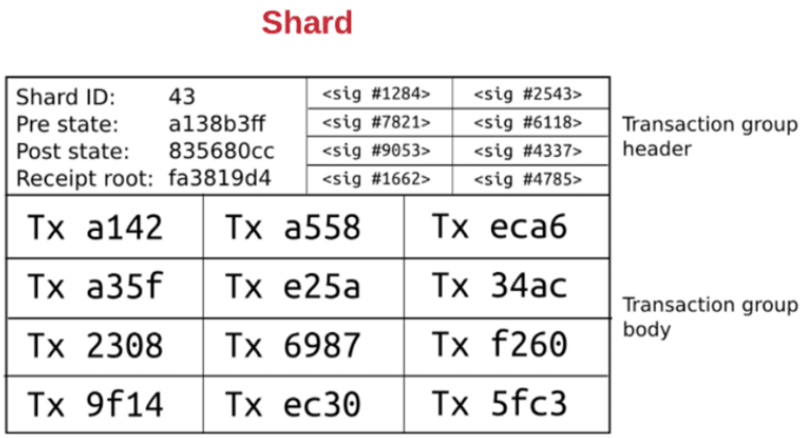
[17] Telegram Open Network Technical Summary. Available: https://icorating.com/upload/whitepaper/gNQ7e9z3lCGi519Wz8mmC0Kg8aA0goeZKAQ802vo.pdf

Fig 13: Sharding



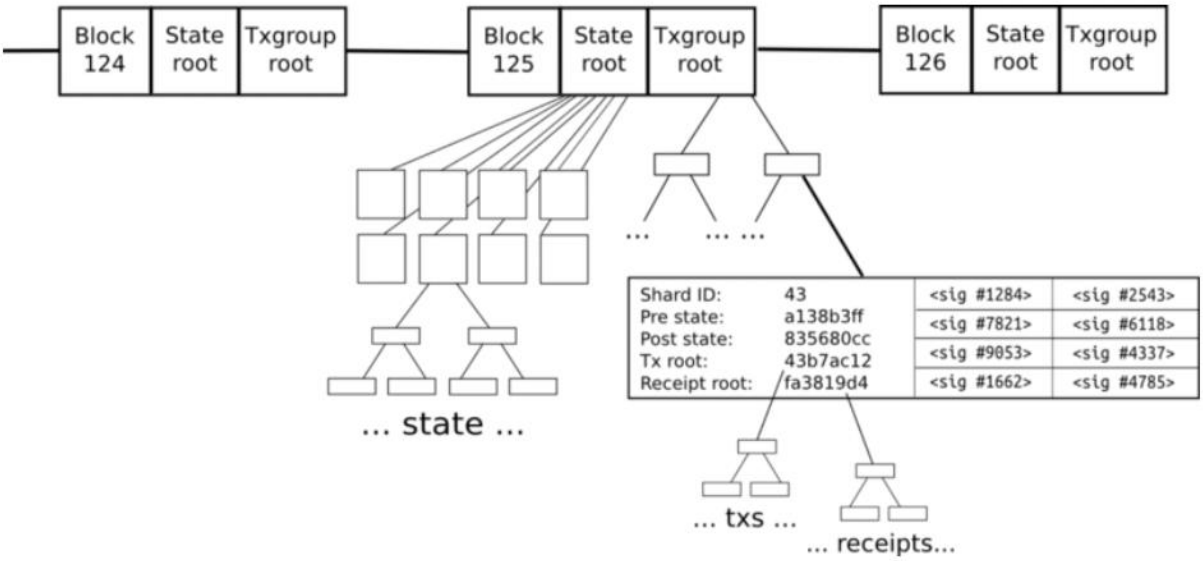
Quelle: Rosic, A. "What are Ethereum Nodes And Sharding." Blockgeeks, Oct. 2017 [9]

Fig 14: Shard Struktur



Quelle: Rosic, A. "What are Ethereum Nodes And Sharding." Blockgeeks, Oct. 2017 [9]

Fig 15: Sharding als On-Chain Technik



Quelle: Rosic, A. "What are Ethereum Nodes And Sharding." Blockgeeks, Oct. 2017 [9]

Im Gegensatz zu diesen On-chain Ansätzen verfolgen Raiden Network[18] für Ethereum und Lightning Network[19] für Bitcoin einen ***Off-Chain*** Ansatz. Hier werden über sogenannte Tunnel-Technologien, die Transaktionen „ausgelagert". Es handelt sich also um eine 2nd-Layer Technologie, die über bilaterale Zahlungskanäle einen Routing-Graphen erzeugt (z.B. basierend auf dem TOR-Onion Protokoll), der es ermöglicht, zwischen beliebeigen Parteien im Graphen BTC zu versenden, wobei abgesehen von der initialen Transaktion keine weiteren Transaktionen in dem Block der übergeordneten „Mutter-Blockchain" gespeichert werden, wodurch diese entlastet und die Skalierbarkeit deutlich erhöht wird.

Ein weiterer Off-Chain Ansatz sind sogenannte ***Subchains*** wie es das Plasma Projekt[20] oder die Cardano Blockchain verfolgen. Die technische Komplexität der einzelnen Ansätze übersteigt den Umfang dieser Seminararbeit, die grundlegende Idee ist jedoch die, dass Child-Blockchains in 2nd+-Layern einer Root-Blockchain zugeordnet werden, was theoretisch inifintes Skalieren und zero Transaction Fees ermöglichen würde. Die Sicherheitsaspekte solcher Child-Blockchains sind jedoch kontrovers, da die einzelnen Child-Chains nicht im selben Maß validiert werden wie die Root-Chain. Wie bereits erwähnt sind die technischen Details der einzlenen Lösungen komplex, die Plasma Lösung basiert z.B. auf zwei grundlegenden Designelementen: (1) der Reformulierung aller Blockchain Berechnungnen als MapReduce Funktionen und dem Proof-of-Stake Token-Bonding von Child zu Root-Chains.

[18] Raiden Network Technical Summary. Available: https://raiden.network/101.html
[19] Lightning Network Technical Summary. Available: https://lightning.network/lightning-network-technical-summary.pdf
[20] Poon, J. and Buterin, V. (2917). Plasma: Scalable Autonomous Smart Contracts. Available: https://plasma.io/plasma.pdf

Zusammenfassend lässt sich sagen, dass das Skalierungsproblem, vor dem Blockchains heutzutage stehen, ein altes Problem der Computerwissenschaften als solche darstellt. Das CAP Thereom besagt, dass in der gegenwart von Netzwerk Partitionen, eine Entscheidung zwischen Konsistenz und Verfügbarkeit getroffen werden muss wenn es zu einem Netzwerkversagen kommen sollte, was für jede Blockchain angenommen werden muss. Leslie Lamperts Arbeiten aus den 80er Jahren beschreiben bereits sehr ähnliche Probleme.[21] Man könnte sogar argumentieren, dass das Web 1.0 und 2.0 genau deshalb erfolgreich waren weil das 5-Layer OSI Modell beduetet, dass nicht jeder Netzwerknutzer alle Informationen braucht. Diese abgeschwächte Informationsbedingung, die das große Datenvolumen in das Application-Layer auslagert, steht daher in einem gewissen Widerspruch zu den normativen Zielen der DLT-Technologie. Eine der größten „Vaporwares" der Geschichte der Computerwissenschaften, Projekt Xanadu, wurde durch Tim Berner-Lees Protokolle HTML und XML überholt und abgelöst, da das umfassende informationstheoretische Konzept, das Xanadu zugrunde lag, nicht umsetzbar war. Blockchains müssen erst noch beweisen, dass ihnen nicht dasselbe Schicksal droht.

Zusammenfassung

"The whole human memory can be, and probably in a short time will be, made accessible to every individual. [T]his new all-human cerebrum…can have at once the concentration of a craniate animal and the diffused vitality of an amoeba…" (H. G. Wells, World Brain, 1938)

Well's Zitat aus dem Jahr 1938 zeigt, dass die Faszination von universal verfügbaren, validen Informationen in einer Form von Peer-to-Peer System nichts Neues ist sondern vielmehr einer sehr alten Menschheitsfantasie entspringt.[22] Bitcoin und DLT-/Blockchain-Technologie haben diesen Traum wiederbelebt und stehen nun vor dem Problem theoretische Konzepte in reale Produkte, Dienstleistungen und Unternehemen zu integrieren.

Blockchains haben sich beträchtlich weiterentwickelt seit Satoshi Nakamoto im Jahr 2009 das Bitcoin Projekt vorstellte. Wenn Bitcoin in seiner ursprünglichen Form die **erste Generation** von Blockchains darstellt, repräsentiert Ethereum mit seinem Fokus auf Smart Contracts und Blockchains als Infrastruktur die **zweite Generation**. Jene Blockchains, die nun entwickelt werden wie Cardano und Telegram Open Network, etc., und speziell die Skalierbarkeit der Blockchain Technologie addressieren, stellen die **dritte Generation** von Blockchains dar.

Smart Contracts (Metadaten und Tokens) führen zu einem rasanten Anstieg von (1) Transaktionen pro Sekunde, (2) der benötigten Netzwerkbandbreite und (3) dem Datenvolumen, das letztendlich – sollte es gelingen, Blockchains zu skalieren und in Main-Stream Anwendungen zu integrieren, zu einem Big Data Paradigma führen wird.

[21] Lamport, L. (1978). Time, clocks, and the ordering of events in a distributed system. Communications of the ACM, 21(7), 558-565.
[22] Ted Nelsons Projekt Xanadu (1960) und sogar Googles Projekt Projekt "BackRub" (1997), aus dem der heutige Google Search Engine entstanden ist, lassen sich thematisch dieser Vision zuordnen.

Diese Seminararbeit ist ergänzend zu dem Big Data Projekt und fügt sich thematisch in mein Promotionsthema „Die Monetarisierung persönlicher Daten“ ein. Ich beabsichtige die im Projekt verwendete API dieses Jahr in meinen Doktoratsforschungsplan an der ETH Zürich miteinzubringen. Der Plan wäre die Anzahl der Clients zu erweitern und die TRON Blockchain als Studienobjekt zu verwenden. Dies ist besonders interessant, da es sich bei TRON, um eine *dezentrale* Social Media und Content Sharing Plattform handelt. Mit Blick auf die Datenskandale aus der jüngsten Vergangenheit bei *zentralisierten* Social Media Plattformen wie Facebook erhoffe ich mir hiervon einen vielversprechenden Datensatz für mein Promotionsthema.

Literatur

- [1] Bartoletti, M., Pompianu, L. “An empirical analysis of smart contracts: platforms, applications, and design patterns.” In International Conference on Financial Cryptography and Data Security, pp. 494-509, April, 2017. Springer, Cham. [Online]. Available: https://arxiv.org/abs/1703.06322v1
- [2] Bartoletti, M., Bracciali, A., Lande, S., & Pompianu, L. “A general framework for blockchain analytics.” 2017. arXiv preprint arXiv:1707.01021. [Online]. Available: https://arxiv.org/abs/1707.01021v2.
    - Github Repo [Online]. Available: https://github.com/bitbart/blockchain-analytics-api
- [4] Chaum, D. “Blind signatures for untraceable payments.” Advances in Cryptology, pp. 199-203, 1983.
- [5] Hirshman, J., Huang, Y., and Macke, S. "Unsupervised approaches to detecting anomalous behavior in the bitcoin transaction network.” Technical report, Stanford University, 2013 [Online]. Available: http://cs229.stanford.edu/proj2013/HirshmanHuangMacke-UnsupervisedApproachesToDetectingAnomalousBehaviorInTheBitcoinTransactionNetwork.pdf
- [6] Madan, I., Saluja, S., and Zhao, A. “Automated Bitcoin Trading via Machine Learning Algorithms.” Technical report, Stanford University, 2015 [Online]. Available: http://cs229.stanford.edu/proj2014/Isaac%20Madan,%20Shaurya%20Saluja,%20Aojia%20Zhao,Automated%20Bitcoin%20Trading%20via%20Machine%20Learning%20Algorithms.pdf
- [7] Nakamoto, S. “Bitcoin: A Peer-to-Peer Electronic Cash System.” 2008 [Online]. Available: https://bitcoin.org/bitcoin.pdf
- [8] Rosic, A. “What Is Hashing? Under The Hood Of Blockchain.” Blockgeeks, Sep. 2017. Retrieved Jan 2018 from: https://blockgeeks.com/guides/what-is-hashing/
- [9] Rosic, A. “What are Ethereum Nodes And Sharding.” Blockgeeks, Oct. 2017. Retrieved Jan 2018 from: https://blockgeeks.com/guides/what-are-ethereum-nodes-and-sharding/
- [10] Rosic, A. “Smart Contracts: The Blockchain Technology That Will Replace Lawyers.” Blockgeeks, Feb. 2017. Retrieved Jan 2018 from: https://blockgeeks.com/guides/smart-contracts/
- [11] The, F., “eCash in a Social Theory of Money: Bitcoin and Other Cryptocurrencies” 2014 [Online]. Available: http://dx.doi.org/10.2139/ssrn.2491743
- [12] “Block.” In Bitcoin Wiki, n.d. Retrieved Jan 2018 from: https://de.bitcoin.it/wiki/Block
- [13] “Block hashing algorithm.” In Bitcoin Wiki, n.d. Retrieved Jan 2018 from: https://en.bitcoin.it/wiki/Block_hashing_algorithm
- [14] “Controlled supply.” In Bitcoin Wiki, n.d. Retrieved Jan 2018 from: https://en.bitcoin.it/wiki/File:Controlled_supply-supply_over_block_height.png

Appendix

Fig. 1a Appendix: Blockstruktur (Bitcoin)

Feld	Beschreibung	Größe
Fixer Wert	immer 0xD9B4BEF9	4 Bytes
Blockgröße	Zahl der Bytes bis zum Ende des Blocks	4 Bytes
Blockheader	besteht aus sechs Teilen	80 Bytes
Transaktionszähler	positive Ganzzahl	1 bis 9 Bytes
Transaktionen	die (nichtleere) Liste von Transaktionen	So viele Transaktionen, wie im Transaktionszähler genannt

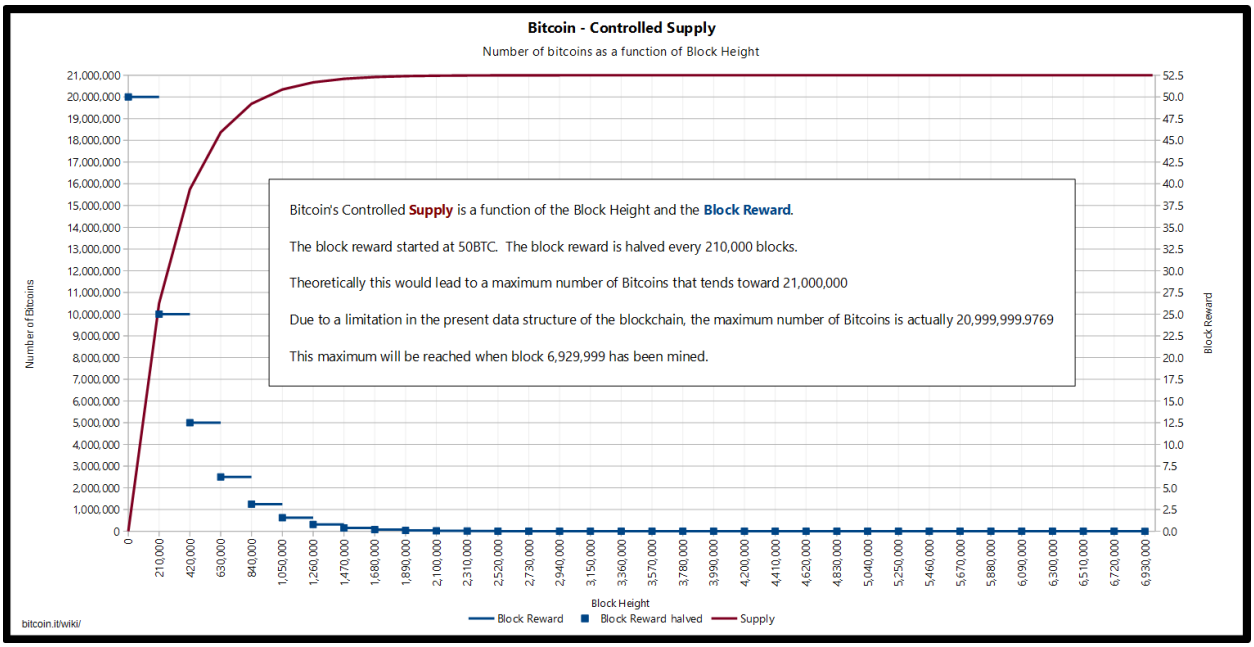
Quelle: Bitcoin Wiki, https://de.bitcoin.it/wiki/Block

Fig. 1b Appendix: Blockstruktur (Bitcoin)

Field	Purpose	Updated when...	Size (Bytes)
Version	Block version number	You upgrade the software and it specifies a new version	4
hashPrevBlock	256-bit hash of the previous block header	A new block comes in	32
hashMerkleRoot	256-bit hash based on all of the transactions in the block	A transaction is accepted	32
Time	Current timestamp as seconds since 1970-01-01T00:00 UTC	Every few seconds	4
Bits	Current target in compact format	The difficulty is adjusted	4
Nonce	32-bit number (starts at 0)	A hash is tried (increments)	4

Quelle: Bitcoin Wiki, https://en.bitcoin.it/wiki/Block_hashing_algorithm

Fig. 2 Appendix: Finites Angebot (Bitcoin)



Quelle: Bitcoin Wiki, https://en.bitcoin.it/wiki/File:Controlled_supply-supply_over_block_height.png

Fig. 3 Appendix: Daten- & Transaktionsvolumen (Bitcoin vs. Ethereum)

	Bitcoin	**Ethereum**
Größe in GB (Jan 2018)	~ 140	~ 45
Transaktionsvolumen (Total) in Millionen	~ 300 (seit 2009)	~ 135 (seit 2015)
Transaktionen / sec[23]	3-7	7-15 (25 max)
Kapitalisierung in Milliarden USD (Jan 2018)	~ 230	~ 120
Umlauf	~ 16 million BTC	~ 97 million ETH

Quelle: Coinmarketcap, https://coinmarketcap.com/currencies/

[23] VISA kommt auf ca. 2000+ Transaktionen/sec zum Vergleich, neue 3rd Generation Blockchains wie TRON und Cardano kommen auf ca. 10.000 Transaktionen/sec.