blockninjas

Endbericht | Call 12 | Projekt ID 2200

Inhalt

1 Einleitung

Blockchains haben die Eigenschaften einer persistenten, manipulationssicheren und verteilten Datenbank. Trotz dieser Eigenschaften ist es schwierig, bestimmte Transaktionsflüsse, Einträge und Ereignisse aus einer Blockchain zu extrahieren und nachzuvollziehen. Dieser Umstand erschwert nicht nur die Arbeit von Ermittlern, die sich mit der Aufklärung von Straftaten im Bereich von Crypto-Währungen beschäftigen, sondern beschäftigt auch Unternehmen wie Banken und Trading-Plattformen, die aufgrund spezieller Regulierungen (“Know Your Customer") darauf angewiesen sind, dass Transaktionen und Vorgänge in Blockchains transparent und nachvollziehbar sind.

Das Projekt blockninjas verfolgt daher das Ziel eine Plattform zu entwickeln, um die Analyse & Visualisierung der Bitcoin Blockchain zu ermöglichen. Basierend auf statistischen Methoden bzw. Machine Learning und Graphen-Theorie sollen Verfahren entwickelt werden, die eine Erkennung bestimmter Transaktion-Muster erlauben sollen.

2 Projektbeschreibung

Das Projekt blockninjas beschäftigt sich mit der Analyse von Überweisungs-Strukturen in Blockchains. Für die Umsetzung des Projektes im Rahmen von netidee haben wir uns auf die Blockchain von Bitcoin fokussiert um einen Analyse-Software Prototypen zu entwickeln.

Die zentrale Zielgruppe des entwickelten Systems wird von Behörden und Institutionen gebildet, die sich mit Ermittlungen zu Bitcoin bzw. Crypto-Währungen befassen. Weitere mögliche Anwender sind Fintech Unternehmen wie Trading-Plattformen bzw. Exchanger, Banken und Finanz-Institutionen, für die Transparenz bzw. Nachverfolgbarkeit von Crypto-Währungs-Transaktionen von grundlegendem Interesse ist, um einen sicheren als auch rechtskonformen Betrieb ihrer Services zu ermöglichen.

In den folgenden Abschnitten gehen wir auf die Kernelemente unseres Projektes ein und beschreiben die entwickelten Analyse-Verfahren sowie das entstandene System.

2.1 Datenbank-Import & Clustering

Eine Kernkomponente unseres Backend-Systems stellt der “Blockchain Analyzer” dar. Sie umfasst die folgenden Funktionalitäten:

- Deserialisierung der rohen Blockchain Daten
- Überführung der Blockchain in ein relationales Modell und Import in eine Postgres Datenbank
- Analyse der Transaktionen, um zusammenhängende Adress-Cluster zu identifizieren

2.2 Structure Detection

Als Teil unseres Projektes beschäftigen wir uns damit, wie man das Überweisungsverhalten von Teilnehmern in der Bitcoin Blockchain analysieren kann, um wiederkehrende Transaktions-Muster nachweisen zu können. Solche wiederkehrenden Strukturen weisen häufig auf automatisierte Überweisungs-Mechanismen hin, wie sie zum Beispiel von

Malware angestoßen werden können. In diesem Abschnitt möchten wir kurz zusammenfassen, wie solche Muster mittels Graphentheorie erkannt werden können.

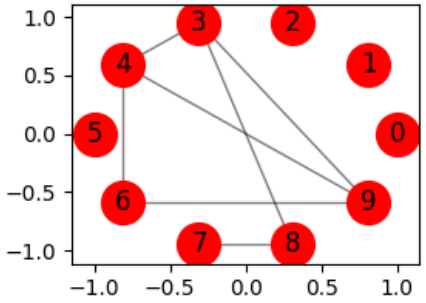
Die Bitcoin Blockchain bildet ein öffentliches Verzeichnis aller jemals durchgeführter Transaktionen. Diese Transaktions-Historie kann als Graph illustriert werden:

- die Knoten des Graphen entsprechen den Adressen der Transaktions-Teilnehmer
- je zwei Knoten im Graphen sind durch eine Kante verbunden, wenn eine Transaktion zwischen den entsprechenden Adressen durchgeführt wurde

Da Transaktionen durch Kanten repräsentiert werden, können Teilgraphen als zeitlich lokale Transaktions-Abfolgen, oder Transaktions-"Muster", verstanden werden. Diese Darstellung der Bitcoin Blockchain ermöglicht die Analyse von Transaktionen durch die Anwendung graphentheoretischer Algorithmen.

Das wiederkehrende Auftreten eines Transaktions-Musters in der Blockchain, entspricht damit dem wiederkehrenden Auftreten gleichartiger Teilgraphen in diesem Transaktionsgraph. Unser Prototyp bietet die Möglichkeit solche Teilgraphen auf Ähnlichkeit zu untersuchen, was erforderlich ist um deren Wiederauftreten nachweisen zu können. Dabei wird jedem Knoten eines Graphen ein Wert zugewiesen, der durch seinen Knotengrad bestimmt wird. Um festzustellen, ob zwei Teilgraphen eine ähnliche Struktur aufweisen, können die Werte der jeweiligen Knoten verglichen werden.

Hier eine Illustration eines simplen Graphen:



Den Knoten des Graphen werden durch den Algorithmus folgende Werte zugewiesen, die anschließend einen Vergleich der Knoten ermöglichen:

0 vs. 1 = 0.0	3 vs. 6 = 0.0044434564116	8 vs. 9 = 0.230183507595
0 vs. 2 = 0.0	4 vs. 6 = 0.0044434564116	6 vs. 8 = 0.364181448502
0 vs. 5 = 0.0	6 vs. 9 = 0.0044434564116	6 vs. 7 = 0.832449584601
1 vs. 2 = 0.0	3 vs. 4 = 0.115852559292	7 vs. 8 = 0.832449584601
1 vs. 5 = 0.0	3 vs. 9 = 0.132517373294	3 vs. 7 = 1.75819368327
2 vs. 5 = 0.0	3 vs. 8 = 0.230183507595	4 vs. 7 = 3.26297247635
4 vs. 9 = 0.0	4 vs. 8 = 0.230183507595	

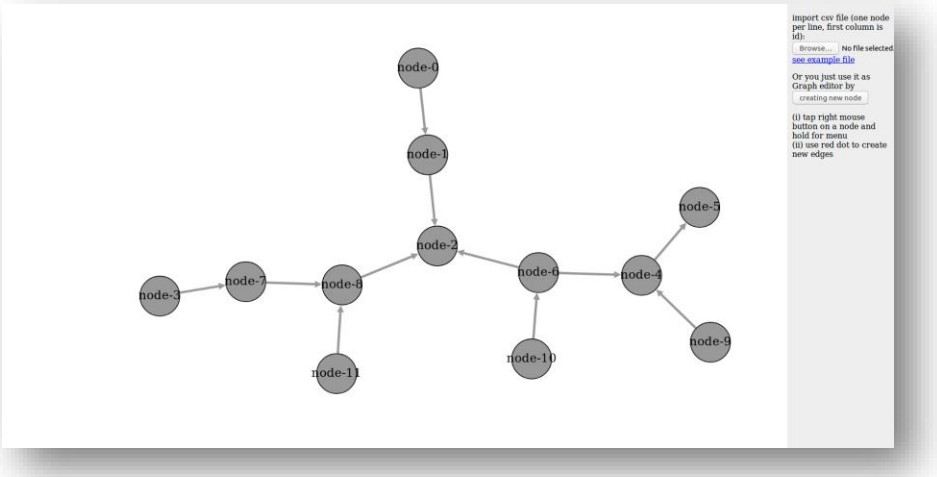
Diese Werte drücken eine Differenz aus. Betrachtet man die erste Zeile, werden die Knoten 0 und 1 als gleichwertig erkannt, was im illustrierten Graphen leicht erkennbar ist, denn

beide Knoten besitzen keine Nachbarn. Dasselbe gilt auch für die Knoten 4 und 9, die jeweils mit drei weiteren Knoten verbunden sind.

Der entwickelte Algorithmus wurde als Teil einer Publikation bei der *IEEE International Conference on Blockchain* eingereicht.

2.3 Graph Visualization Helper

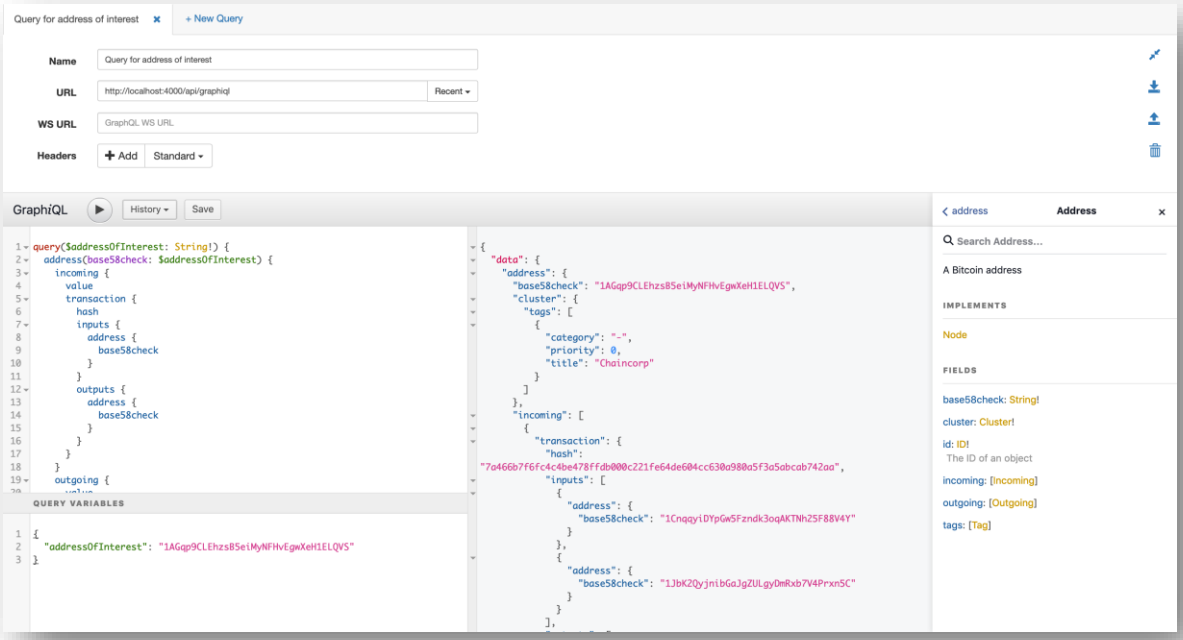
Ein Beiprodukt der Evaluierung des Strukturerkennungs-Algorithmus ist das *Graphtool*. Es ermöglicht die einfache Erstellung und Visualisierung allgemeiner Graphen. Folgend sieht man einen Screenshot von einer Beispielvisualisierung eines Graphen in unserem Tool:



2.4 GraphQL API

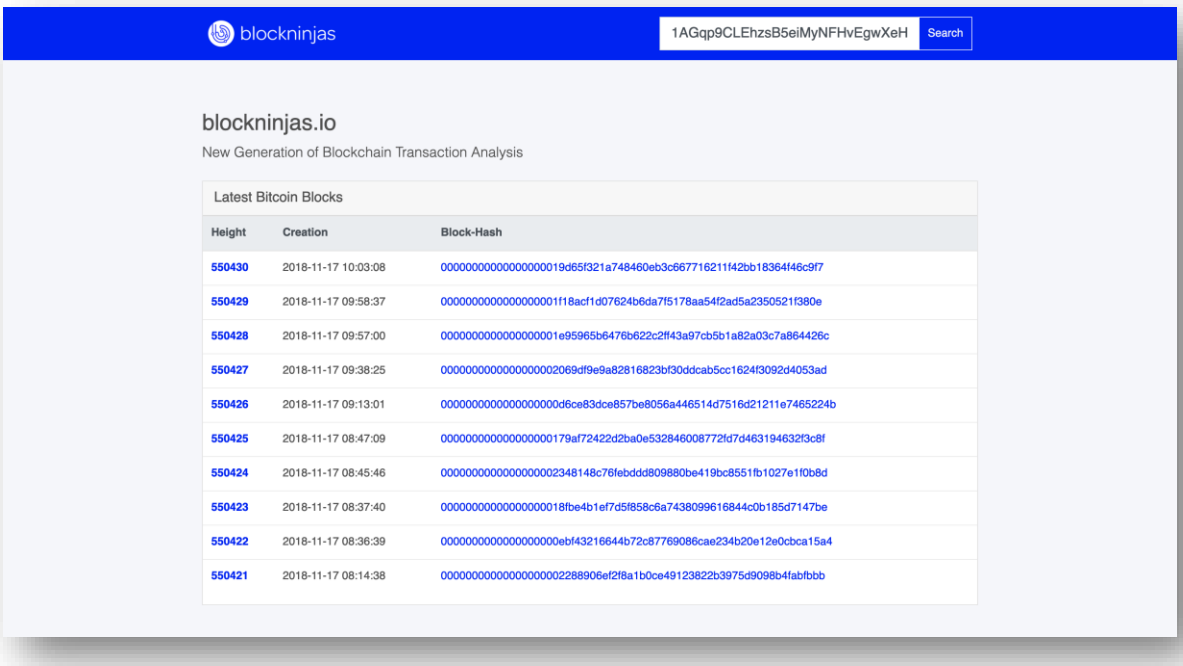
Die importierten Blockchain-Daten sowie die Ergebnisse der Analysen werden über eine typisierte API zur Verfügung gestellt, die auf der Open-Source-Datenabfragespache *GraphQL* basiert.

Zur Illustration und Entwicklung von Queries kann die integrierte graphische Oberfläche *GraphiQL* genutzt werden. Die folgende Graphik zeigt einen beispielhaften Query zur Abfrage der Informationen zu einer einzelnen Bitcoin Adresse:

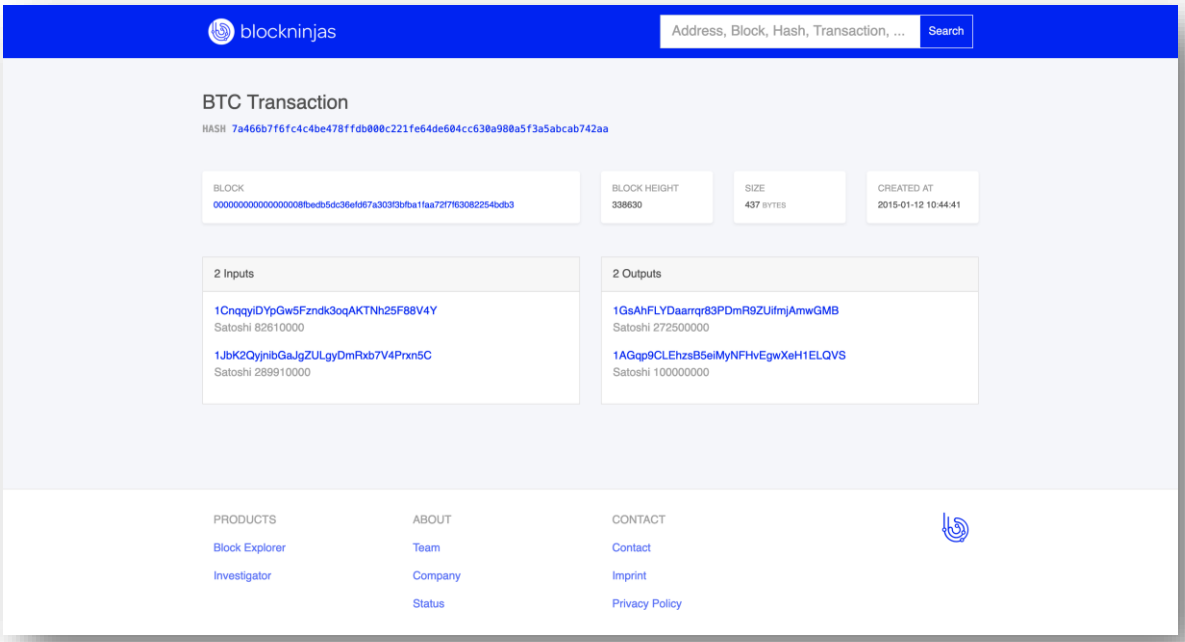


2.5 Block Explorer

Als *Bock Explorer* wird im Bitcoin-Umfeld ein System bezeichnet, das es ermöglicht mittels graphischer Oberfläche die Blöcke und Transaktionen der Bitcoin Blockchain zu durchsuchen. Eine solche Komponente wurde ursprünglich nicht als Teil des Projektes eingeplant, wurde jedoch als Entwicklungshilfe und schließlich auch als Feature in unser System integriert. Wie in folgendem Screenshot ersichtlich, ermöglicht der Block Explorer eine Einsicht in den Live-Zustand der Bitcoin Blockchain:



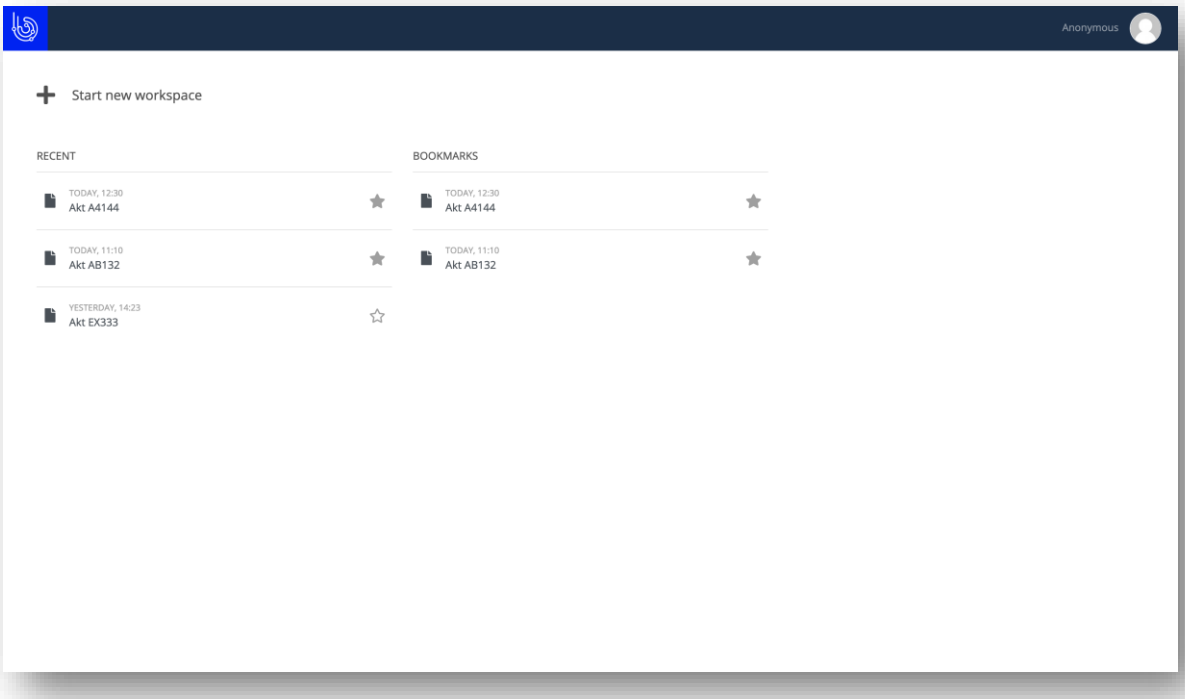
Eine Suche nach Adressen, Block-Hashes und Transaktions-Hashes ermöglicht es die Blockchain im Detail einzusehen. Die folgende Ansicht zeigt eine Transaktion, mit jeweils zwei Adressen auf Input- und Output-Seite.



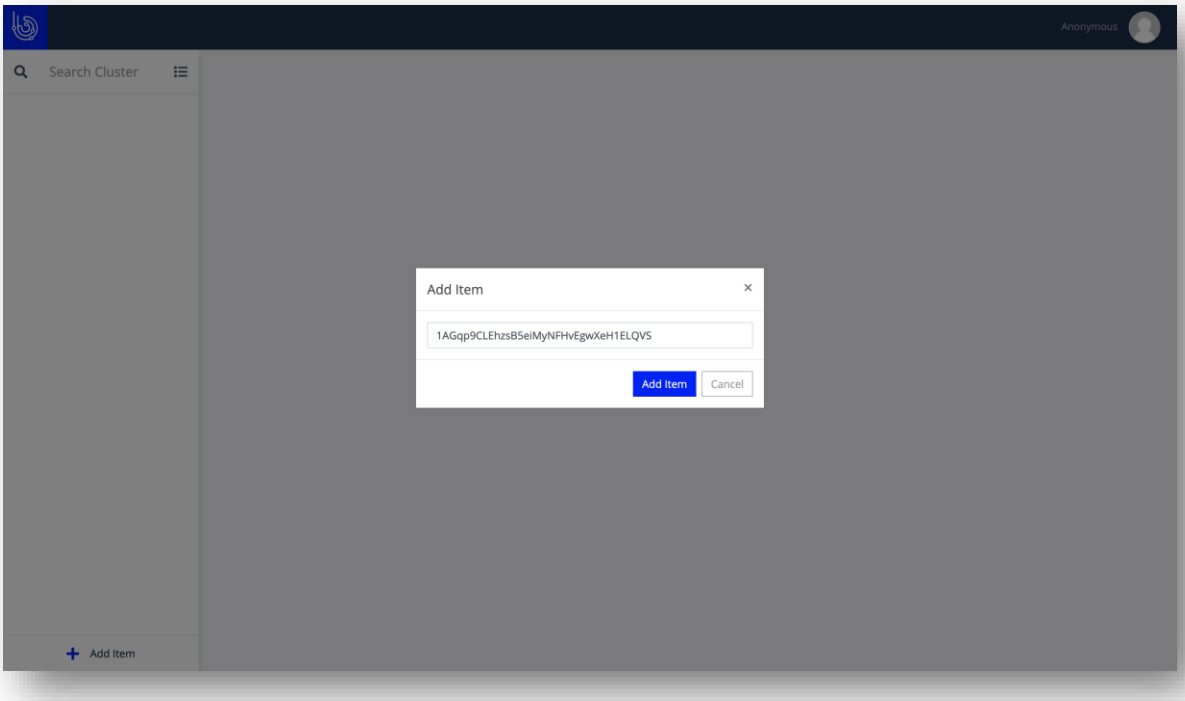
2.6 Transaction Graph Explorer

Zentraler Bestandteil des Frontends bildet der Transaction Graph Explorer. Sie bietet dem Benutzer die Möglichkeit, Transaktionen und in Verbindung stehende Adressen in Form eines Graphen zu visualisieren.

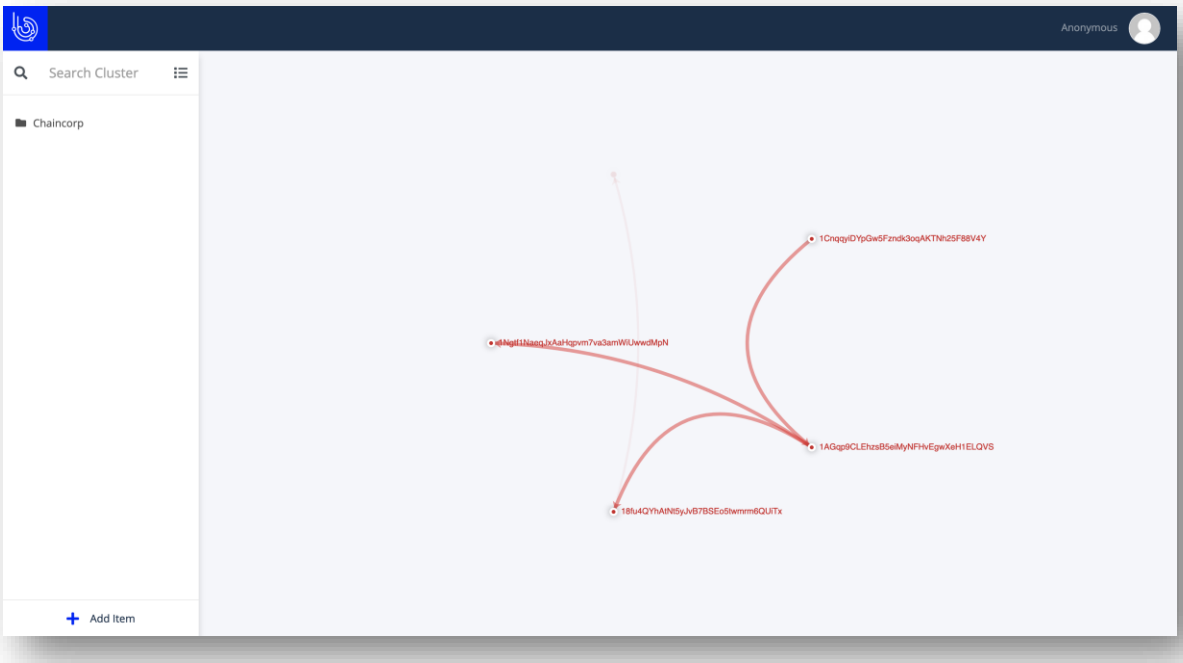
Die Benutzerführung beginnt mit der Verwaltung von *Workspaces*. Ein Workspace kann dazu verwendet werden um Adressen und Transaktion einer bestimmten Ermittlungs-Tätigkeit zu gruppieren:



Wird ein neuer Workspace angelegt, oder ein bereits vorhandener geöffnet, können nun die zu visualisierenden Adressen in den Workspace eingefügt werden:



Nachdem mehrere Adressen in den Workspace eingefügt wurden, erscheinen verbindende Kanten im Graphen um Geldflüsse zwischen einzelnen Adressen hervorzuheben:



3 Verlauf der Arbeitspakete

3.1 Arbeitspaket 0 – *netidee Startbericht*

Den Beginn unseres Projektes markierte der Startbericht. Eine zentrale Zielstellung des Projektes stellte die Erarbeitung neuer Analyse-Möglichkeiten von Transaktionen in Blockchains dar, sowie deren prototypische Implementierung für die Bitcoin Blockchain. Die zentrale Herausforderung dieser Projektphase bestand darin, den Umfang der dafür notwendigen Tätigkeiten abzuschätzen, sowie die Features des geplanten Systems zu definieren.

3.2 Arbeitspaket 1 – *Website für Projektpräsentation*

Zur Präsentation unserer Projektergebnisse, haben wir zusätzlich zur netidee Webseite eine eigene Projektwebseite eingerichtet, welche unter https://www.blockninjas.io abgerufen werden kann.

3.3 Arbeitspaket 2 – *Research Blog + Testdatenset*

Um die im Rahmen des Projektes gesammelten Erkenntnisse einer breiteren Öffentlichkeit zuführen zu können, wurde ein Research Blog unter https://www.netidee.at/blockninjas sowie auf unserer Website unter https://www.blockninjas.io zugänglich gemacht.

Weiters wurde für die Implementierung und Evaluierung der geplanten Strukturerkennungs-Algorithmen ein Testdatenset definiert und wie in den Projektergebnissen ersichtlich auf GitHub veröffentlicht.

3.4 Arbeitspaket 3 - *Zwischenbericht*

Dieses Arbeitspaket beinhaltet die Verfassung des Zwischenberichts sowie die Anfertigung der damit zusammenhängenden Dokumente und Auswertungen.

3.5 Arbeitspaket 4 – *Live System*

Dieses Arbeitspaket umfasste die folgenden Tätigkeiten:

- Prototyp
- AWS Konfiguration
- Webinterface Sketch
- Readprozess
- Webinterface v1
- Webinterface Sketch Structures
- Rollout Konzept
- Structure Detection Integration
- Clustering
- API v2 (inkl. Clustering)
- Structure Detection auf Clusterdaten
- Webinterface v2 (Graph)
- Webinterface v3 (Structure)

Eine zentrale Herausforderung stellte die Anfertigung des Designs der in Abschnitt 2 ersichtlichen User Interfaces, um eine übersichtliche Visualisierung der Bitcoin Blockchain als Graph zu ermöglichen.

Als weitere Herausforderung bei der Umsetzung des Live-Systems stellt sich der Entwurf einer geeigneten Architektur heraus, welche über die Infrastruktur auf *Amazon Web Services (AWS)* abgebildet werden sollte. Die Vielseitigkeit sowie Komplexität von AWS erforderte eine zeitintensive Einarbeitung. Aufgrund der Kosten der eingesetzten AWS Komponenten, mussten Experimente zeitlich begrenzt werden und erforderten eine gründliche Planung der Deployments.

Aufgrund der erheblichen laufenden Kosten der AWS Infrastruktur, kann das System derzeit leider nicht öffentlich zugänglich betrieben werden.

3.6 Arbeitspaket 4 – *Source Code*

Dieses Arbeitspaket umfasste die folgenden Tätigkeiten:

- Bugfixing
- Veröffentlichung der Software auf Github
- Abschließende Arbeiten
- Dokumentation

4 Umsetzung Förderauflagen

Keine speziellen Auflagen definiert.

5 Liste Projektergebnisse

1	Website blockninjas.io	CC-BY-SA	https://www.blockninjas.io
2	Testdatenset-Visualisierung	CC-BY-SA	https://netidee.at/blockninjas/testdatenset-visualisiert https://blockninjas.io/preview.html
3	Testdatenset	CC-BY-SA	https://github.com/blockninjas/bitcoin-test-dataset
4	Strukturerkennungs-algorithmus	CC-BY-SA	https://github.com/mrqc/graph-spectrum
5	Blogeinträge	CC-BY-SA	https://netidee.at/blockninjas
6	Graphtool	CC-BY-SA	https://mrqc.github.io/grphr
7	Source Code Explorer	GPL v3	https://github.com/blockninjas/blockchain_explorer
8	Source Code API	GPL v3	https://github.com/blockninjas/blockchain_api
9	Source Code Datenimport und Cluster-Analyse	GPL v3	https://github.com/blockninjas/blockchain_analyzer
10	Entwickler-Doku	GPL v3	https://netidee.at/blockninjas/unsere-datenstruktur
11	Anwenderdoku Import und DB Setup	GPL v3	https://github.com/blockninjas/blockchain_analyzer/blob/master/README.md
12	Anwenderdoku API	GPL v3	https://github.com/blockninjas/blockchain_api/blob/master/README.md
13	Anwendungerdoku Explorer	GPL v3	https://github.com/blockninjas/blockchain_explorer/blob/master/README.md
14	Zusammenfassung	CC-BY-SA	https://netidee.at/sites/default/files/2019-08/prj2200_Call12_Zusammenfassung_V03.pdf

15	Zwischenbericht	CC-BY-SA	https://netidee.at/sites/default/files/2019-06/prj2200_Call12_Zwischenbericht_V01.pdf
16	Endbericht	CC-BY-SA	https://netidee.at/sites/default/files/2019-08/prj2200_Call12_Endbericht_V02.pdf

6 Verwertung der Projektergebnisse in der Praxis

Ziel ist es mit zusätzlichen potenziellen Anwendern in Kontakt zu treten. Feedback aus dem produktiven Einsatz soll den Product-Market Fit des Systems sicherzustellen und Möglichkeiten zur Weiterentwicklung aufzeigen.

Weiters steht die Suche strategischer Partner, sowie das Ansuchen um weitere Förderungen im Fokus, um einen weiteren Betrieb sowie die Weiterentwicklung zu ermöglichen.

7 Öffentlichkeitsarbeit/Vernetzung

- Im Zuge des i²c Networking Fridays durften wir unser Projekt im Kuppelsaal der TU Wien präsentieren und konnten dabei den Audience Award gewinnen. Der Preis bestand aus einem Ticket zum Pioneers-Festival in Wien welches wir nutzen konnten, um andere Firmen und Projekte aus der Startup-Szene kennenzulernen. Weitere Informationen unter https://www.trendingtopics.at/das-sind-die-gewinner-startups-des-i%C2%B2c-networking-friday-an-der-tu-wien/
- Newseintrag TU Wien Webseite: https://www.tuwien.at/tu-wien/aktuelles/news/news/erfolge-bei-netidee-fuer-projekte-der-fakultaet-fuer-informatik/
- #glaubandich Challenge 2018 Wien: https://www.trendingtopics.at/glaubandich-challenge-kandidaten-wien-wko-sky-lounge/
- #glaubandich Challenge 2019 Wien: https://www.trendingtopics.at/glaubandich-wien-fintech-startups-teilnehmer-2019/

8 Eigene Projektwebsite

Es wurde zusätzlich zur netidee Webseite eine eigene Webseite implementiert, welche unter folgender Adresse abgerufen werden kann: https://www.blockninjas.io

9 Geplante Aktivitäten nach netidee-Projektende

Wir wollen in Zusammenarbeit mit externen Organisationen unser Produkt weiterentwickeln und weitere Clustering-Methoden identifizieren. Der Algorithmus zur Strukturerkennung, der

weiter oben veranschaulicht wurde, wird bei der nächsten Deadline zum Journal der IEEE International Conference on Blockchain eingereicht.

Zu Abgabe des Endberichts konnten wir insgesamt 19 Clustering-Methoden identifizieren. Ein Teil dieser Clustering-Verfahren wurde im Rahmen dieses Projekts implementiert. Im Zuge der Weiterentwicklung der Software, sollten weitere Verfahren implementiert werden, um in einer Gegenüberstellung die Wirksamkeit der einzelnen Verfahren zu vergleichen.

10 Anregungen für Weiterentwicklungen durch Dritte

Wir freuen uns über jegliche Anregungen und Feedback und würden und besonders freuen, falls jemand die Weiterentwicklung des Projekts unterstützen möchte. Konkret bieten sich folgende Möglichkeiten der Weiterentwicklung:

- Implementierung weiterer Clustering-Verfahren
- Erweiterung der GUI und API um neue Abfragemöglichkeiten
- Recherchen zur Deanonymisierung von Bitcoin Adressen
- Integration weiterer Blockchains

11 Danksagung

Wir möchten uns auf diesem Weg auch noch bei netidee für die großzügige Unterstützung bedanken, welche es ermöglicht hat, unsere Ideen zu verwirklichen und einer breiteren Öffentlichkeit zugänglich zu machen!