WalletRank

Team 105: Rizki Wicaksono, Mani Narasimhan, Jennifer Evans, Siddharth Seth

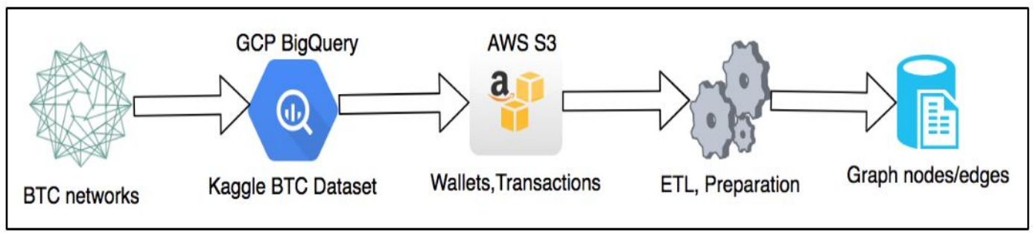
Problem and Motivation:

Bitcoin is very new and experiencing rapid growth and adoption. Its anonymity aspect makes it difficult to regulate or control crimes made with bitcoins (like Ransomware.) People from the average user to banks and corporations struggle to know if they should do business with a particular Bitcoin user (wallet id) or not. The industry today is so immature that it relies on manual reporting of malicious wallet ids on sites like bitcoinabuse.com. Our project strives to remove the uncertainty of doing business with bitcoin, even for unreported malicious wallets. We leverage the full dataset of known wallets, and all reported malicious wallets and perform a calculation known as Personalized PageRank which effectively rates a wallet by the neighbors it has done business with. Wallets that have done business with known malicious wallets will have lower scores. WalletRank reports wallet scores in a consumer-friendly "credit-reporting" scale of 300 (bad) to 800 (good) to allow easy checking on whether to do business with a wallet id or not.

Data Source:

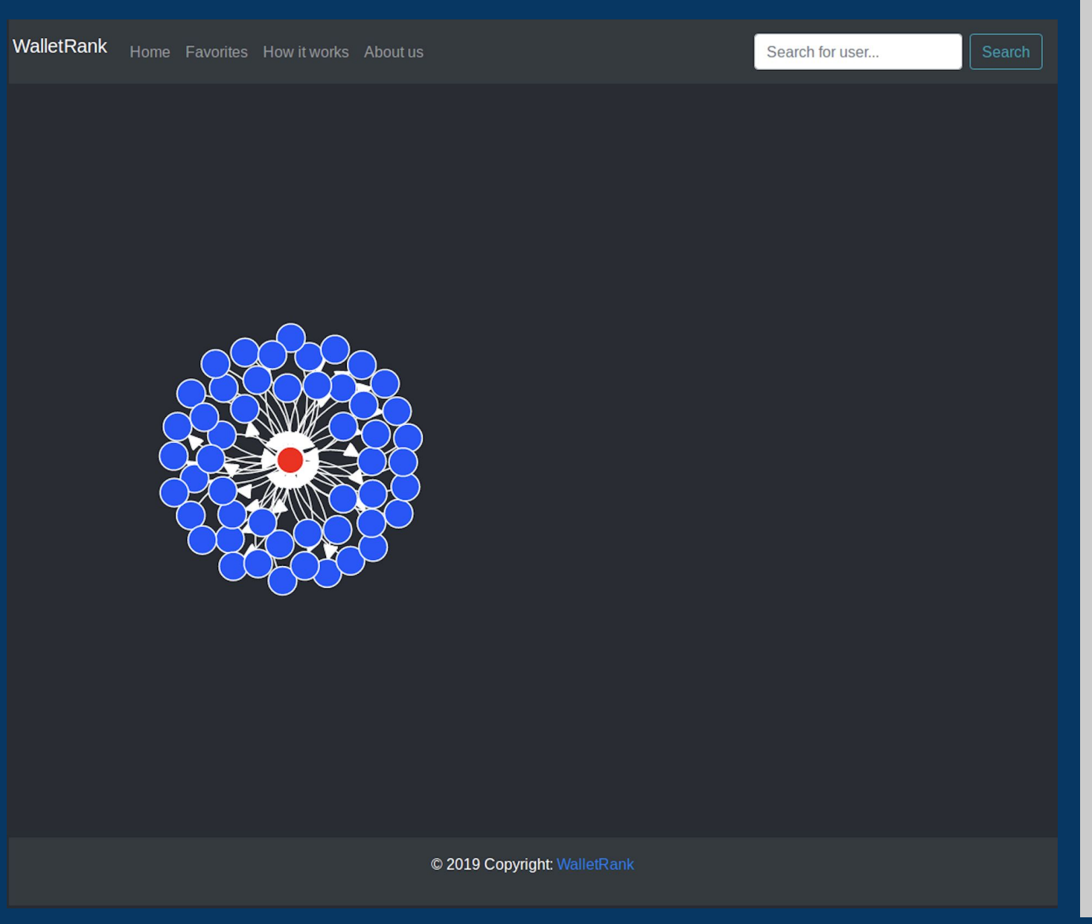
We selected the full bitcoin transaction database for 1 year from Sep 2017 - Sept 2018. We selected the data from the Kaggle dataset using BigQuery (https://www.kaggle.com/bigquery/bitcoin-blockchain). The data represented 698,272,839 records and was stored on 57.18 GB of disk space.

Using this data we produced an edge dataset with source wallet, destination wallet, as well as node (wallet) indegree and outdegree information. We used the bitcoinabuse.com database of manually reported malicious wallets to create a dataset of 233 known bad wallet IDs and we used the combined information from all datasets running Personalized PageRank, to generate a dataset of walletid, score. This score was converted to a consumer-friendly range of 300 (bad) to 800 (good).

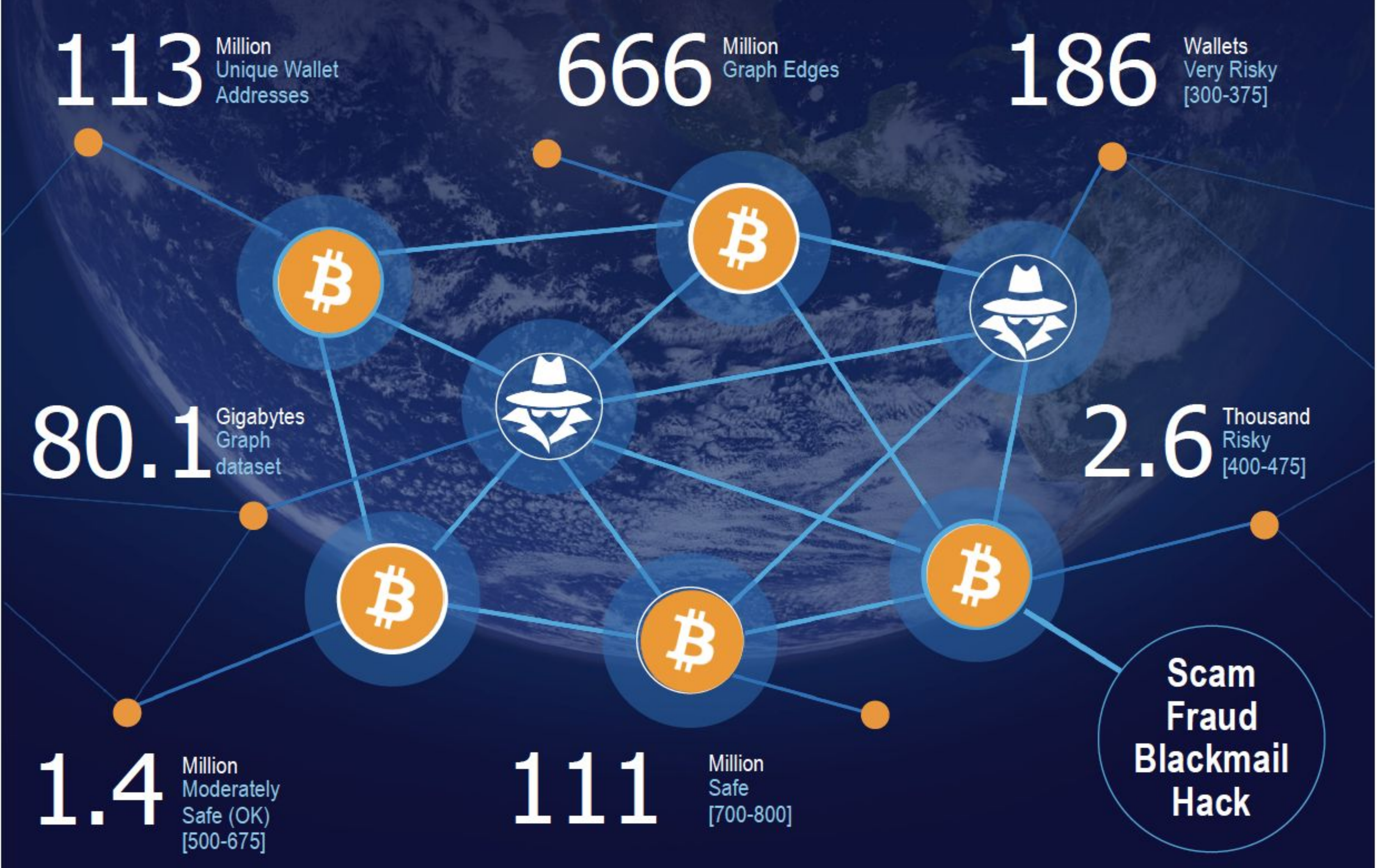


Experiments and Results:

We have collected the raw data, processed and cleaned it, and we have a full-scale run of Personalized PageRank (PPR) on the full bitcoin dataset. We fed the data for the known 233 malicious wallets from BicoinAbuse.com-reported wallets into the algorithm to "tag" the wallets of interest. We also fed the full dataset of all wallets and transactions for the year including Sep 2017 - Sep 2018. These wallets can be considered nodes, and transactions can be considered edges, forming a directed graph. This directed graph and the list of malicious wallets, supplied to the PPR algorithm produces a score for each node in the graph between [0, 1]. This successfully worked when we ran the full algorithm and each node was identified with a score. Scores for known bad (reported) wallets were set to 1, and we looked at the top 25 wallets that were predicted bad by the algorithm.



© 2019 Copyright: WalletRank



113 Million Unique Wallet Addresses

666 Million Graph Edges

186 Wallets Very Risky [300-375]

80.1 Gigabytes Graph dataset

2.6 Thousand Risky [400-475]

1.4 Million Moderately Safe (OK) [500-675]

111 Million Safe [700-800]

Scam Fraud Blackmail Hack

Risk Score for Bitcoin

The anonymity of bitcoin offers positive benefits, but can attract fraudulent or even illegal activities. Unfortunately, the anonymity of bitcoin makes regulatory compliance and enforcement difficult. We propose a Bitcoin Wallet "Credit Score" called WalletRank to measure risk of doing business with a wallet-holder. Similar to the credit score used to measure creditworthiness, WalletRank can measure risk before conducting a bitcoin transaction.

Personalized PageRank

Personalized PageRank (PPR) is used to quantify risk of malicious wallets. We begin with marking few hundreds of wallets as malicious based on open fraud reports found in bitcoinabuse.com and use this as initialization vectors to PageRank algorithms. We then convert probability into risk scores from 300-800.

300-375 Very Risky

400-475 Risky

500-675 Moderately Safe (OK)

700-800 Safe

Evaluation:

Previously unreported malicious/potentially malicious wallets were evaluated and found by manually investigating a number of other sites which were of ill repute and also mentioned these wallets (common bitcoin scammer websites such as bitcoindoubler.com). We found among these 10 previously unknown (not reported on Bitcoinabuse.com) malicious wallets, and 10 previously unknown potentially malicious wallets. Only 5 were found not to be malicious. It seems the PPR algorithm is, in fact, able to predict new bad actors based on association with the known ones, confirming our major hypothesis.

This method and result appears to be state of the art versus the manual ad-hoc after-the-fact reporting methods that exist today. Malicious wallet prediction is now possible.

Observations include the findings for the top 25 highest scoring PPA wallets (i.e. most likely to be malicious) in the table nearby:

Wallet Address	WalletRank	Findings	Types	New
1NDyJtNTjmwk5xPNhjgAMu4HDHigtobu1s	300	Malicious	Blackmail	No
12cgpFdJViXbwHbhrA3TuW1EGnL25Zqc3P	300	Not Malicious	-	-
1KSuWHN6Hc36tJmYtk4RyAbkKDtNjACRX9	300	Malicious	Scam	Yes
1FoWyxwPXuj4C6abqwhjDWdz6D4PZgYRjA	300	Malicious	Scam	Yes
1JUToCyRL5UwgeucjnFAagKs4v1YqhjT1d	300	Potentially Malicious	Anomalous	Yes
1DUb2YYbQA1jjaNYzVXLZ7ZioEhLXtbUru	300	Malicious	Scam	Yes
1Fu3iBR2EMQWeYGi3XvrPmcPUkne8ZWj9h	300	Malicious	Hacked	No
1LhWMukxP6QGhW6TMEZRcqEUW2bFMA4Rwx	300	Potentially Malicious	Anomalous	Yes
1G47mSr3oANXMafVrR8UC4pzV7FEAzo3r9	300	Malicious	Scam	Yes
1LCAJF94Yxin9eWNx19b5BZnrnBSVstV1g	300	Potentially Malicious	Scam	Yes
1Po1oWkD2LmodfkBYiAktwh76vkF93LKnh	300	Malicious	Scam	Yes
392LK4ZQD3gixWg5xJRTv1a24N3YDgCbwP	300	Malicious	Scam	Yes
1L6zTihRVecCjisYkn6BuXKrwvg8hJFC4f	300	Malicious	Scam	Yes
1H6q83MQr9k8c6VezSU8x5oKasABjF4btN	300	Potentially malicious	Anomalous	Yes
3QorgsdWKX2CvaMDPH5PvRchoz8s9GM2by	300	Not Malicious	-	-
39f5XZ1vRB3Lm187psnrwSikLy8Xmm3DCu	300	Potentially malicious	Anomalous	Yes
1PAxSJnxGRWTNSa4NgRbNiQahEMY9KjGJ3	300	Potentially malicious	Anomalous	Yes

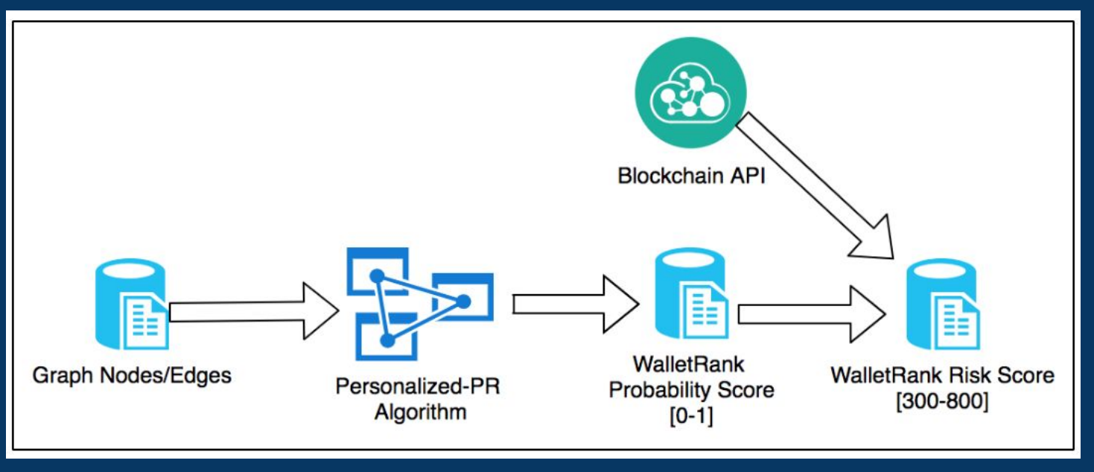
Our Innovation and Contribution:

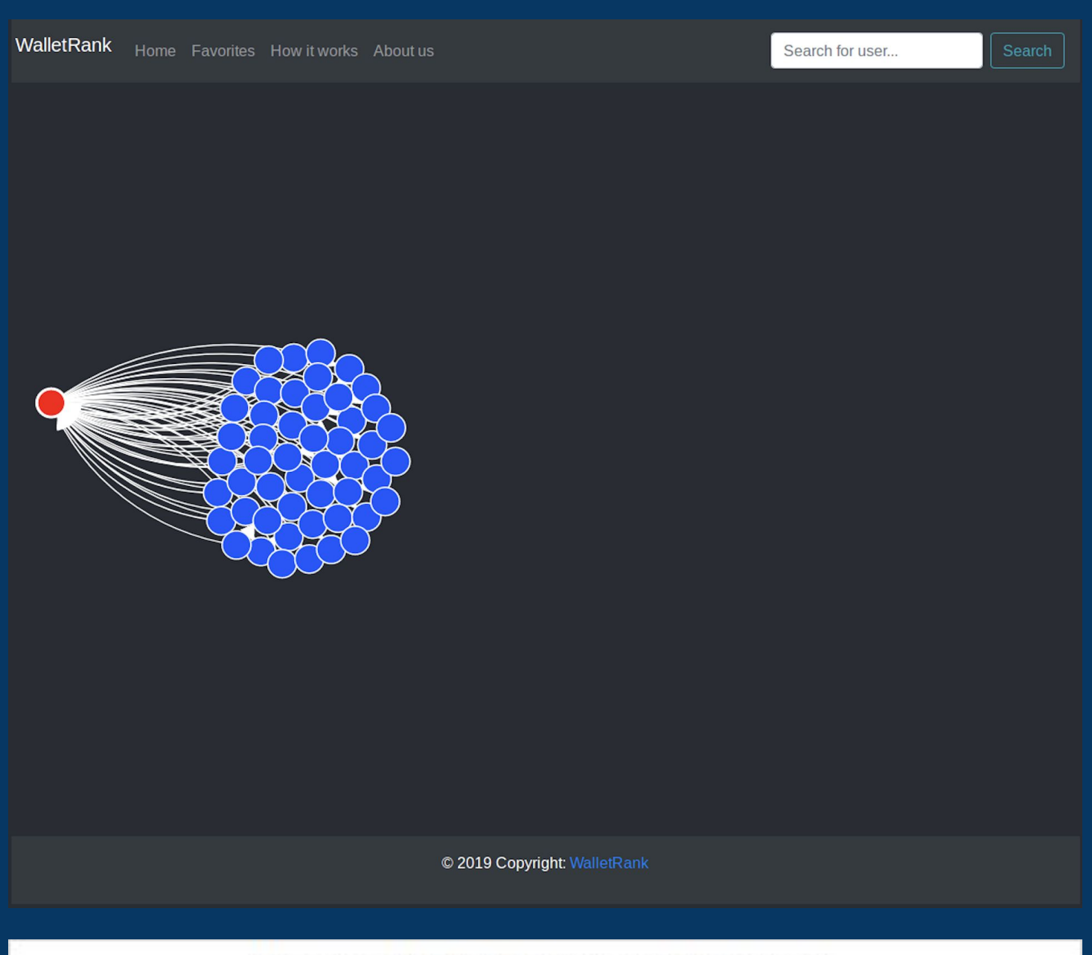
- Intuitive and user friendly risk score even for novice users
- Automated way to quantify risk of every bitcoin wallet address based mainly on personalized PageRank algorithm
- Interactive visualization of bitcoin transactions allows user to inspect interesting wallets and transactions

Approach to Algorithm and Visualization:

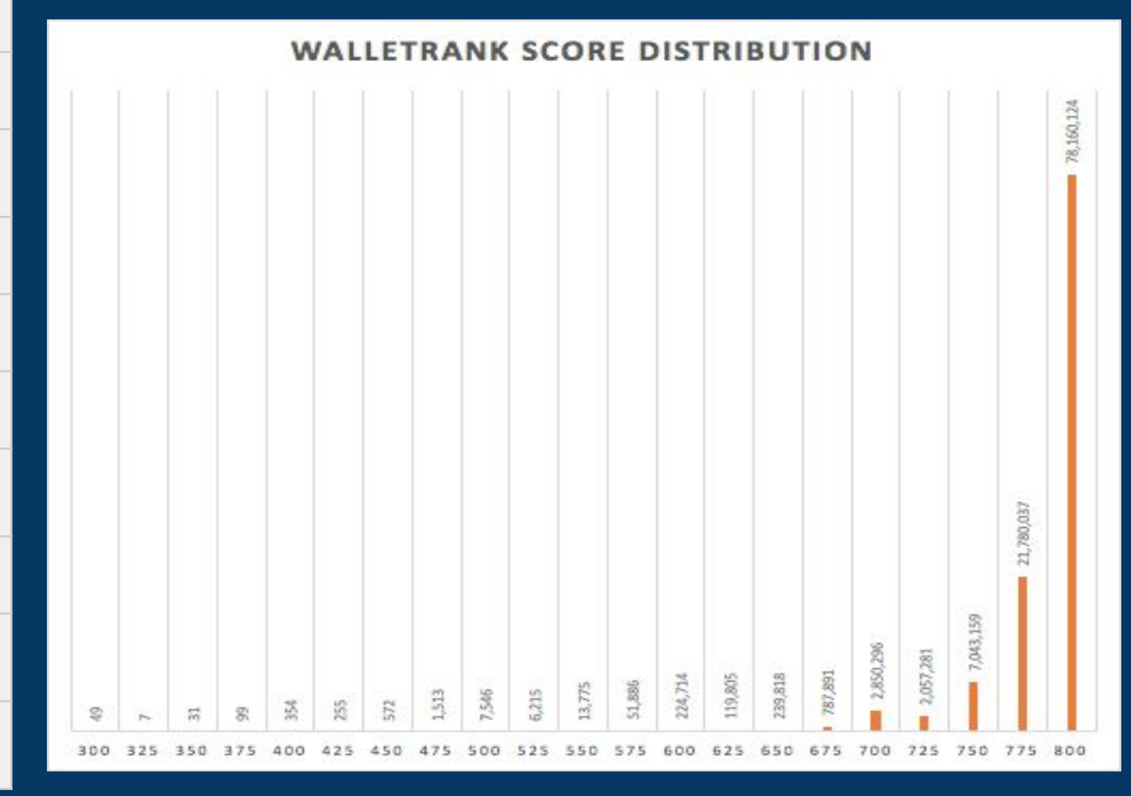
Currently, the mechanism for identifying bad actors within the Bitcoin community is to rely on manual reporting by people who have experienced a problem with a particular wallet (end-user). These are listed on bitcoinabuse.com for people to check before doing business with a particular wallet. The list includes those wallets used to collect money from Ransomware, extortion schemes and other criminal activity. There is an API to both report automatically, and to check if a wallet is on the reported list. However, even though an API exists, this process still requires an actual effort to identify and report a wallet, or to check if someone has reported. It does not "predict" potentially bad wallets, nor does it produce a risk score for them.Our approach is to assess transaction risk with particular wallet addresses in a quantitative way. In other words, we are trying to quantify risk for each bitcoin wallet address such that a user would be able to gauge the risk of doing business before sending bitcoin to the wallet. We propose using Personalized PageRank (PPR) algorithm to find vertices of interest and those nearby vertices of interest. We label known "bad" vertices with a score of 1.0 and allow PPR to "taint" nearby vertices with their bad reputation.

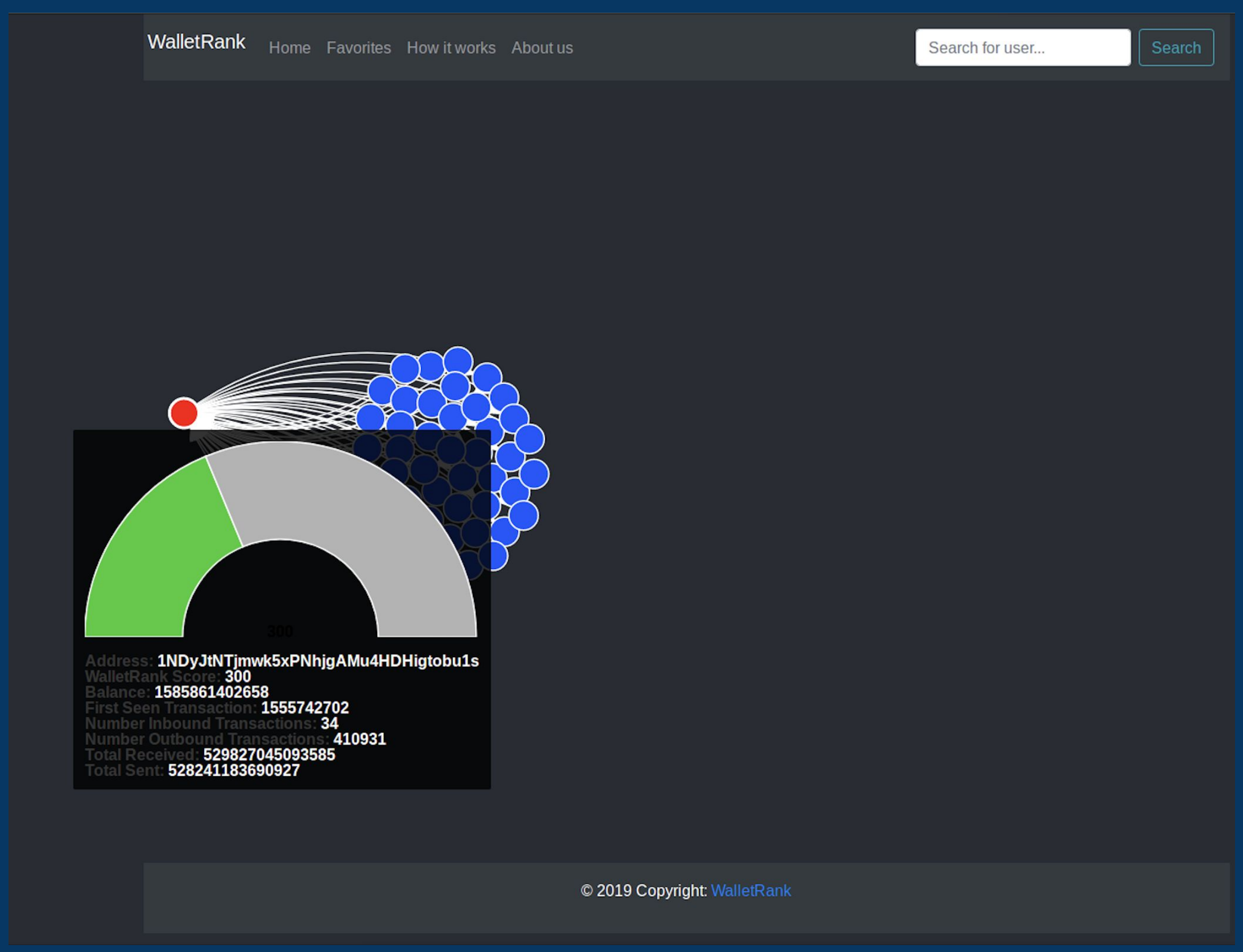
When the algorithm achieves a measure of stability, we have a majority of the vertices which are good near 0.0, those that are bad at 1.0 and those which are suspect somewhere above 0.2. Once we learn the PPR score for each wallet, we can calculate a Consumer-Friendly risk score [300 - 800].


Blockchain API
Graph Nodes/Edges
Personalized-PR Algorithm
WalletRank Probability Score [0-1]
WalletRank Risk Score [300-800]



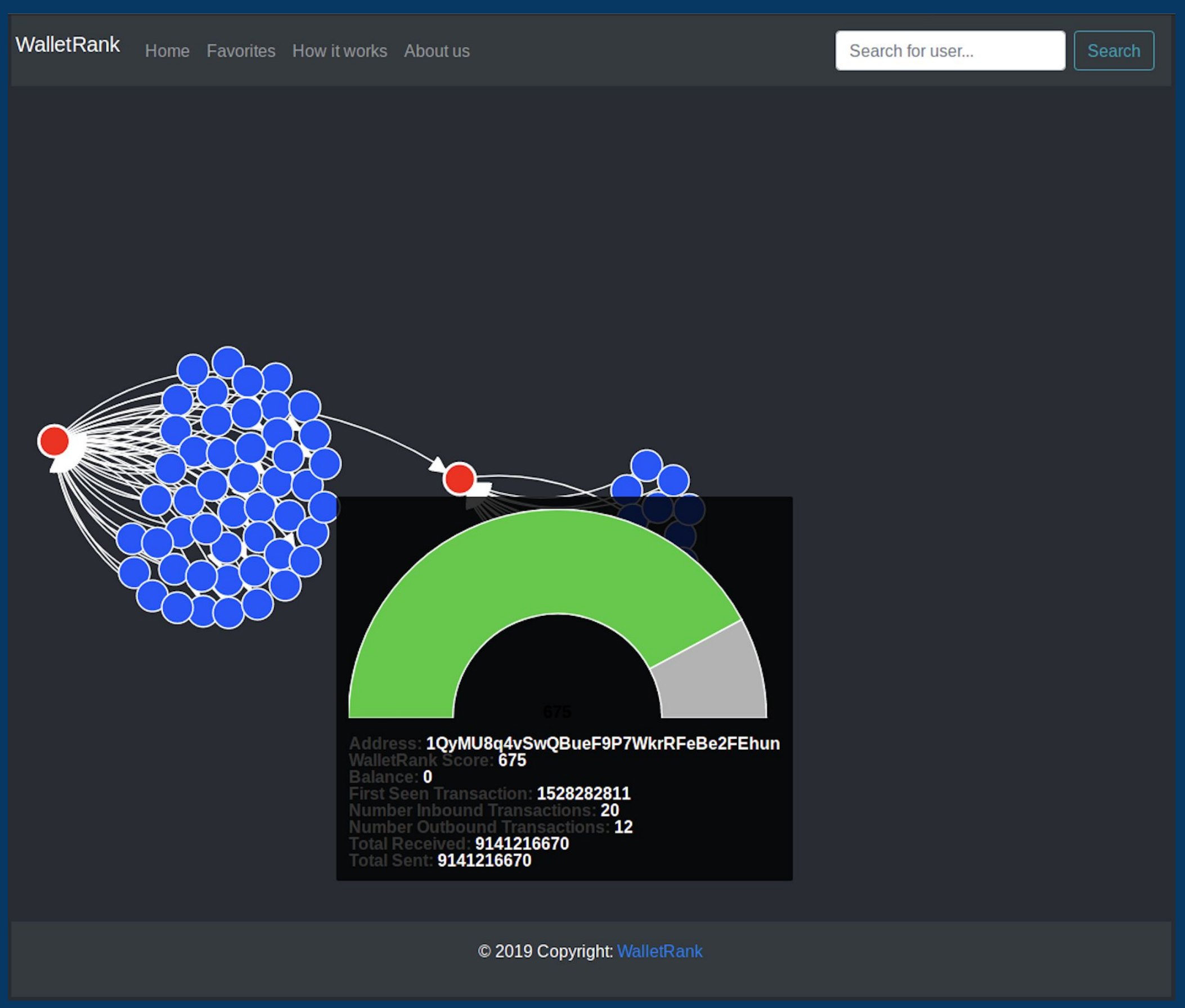
© 2019 Copyright: WalletRank





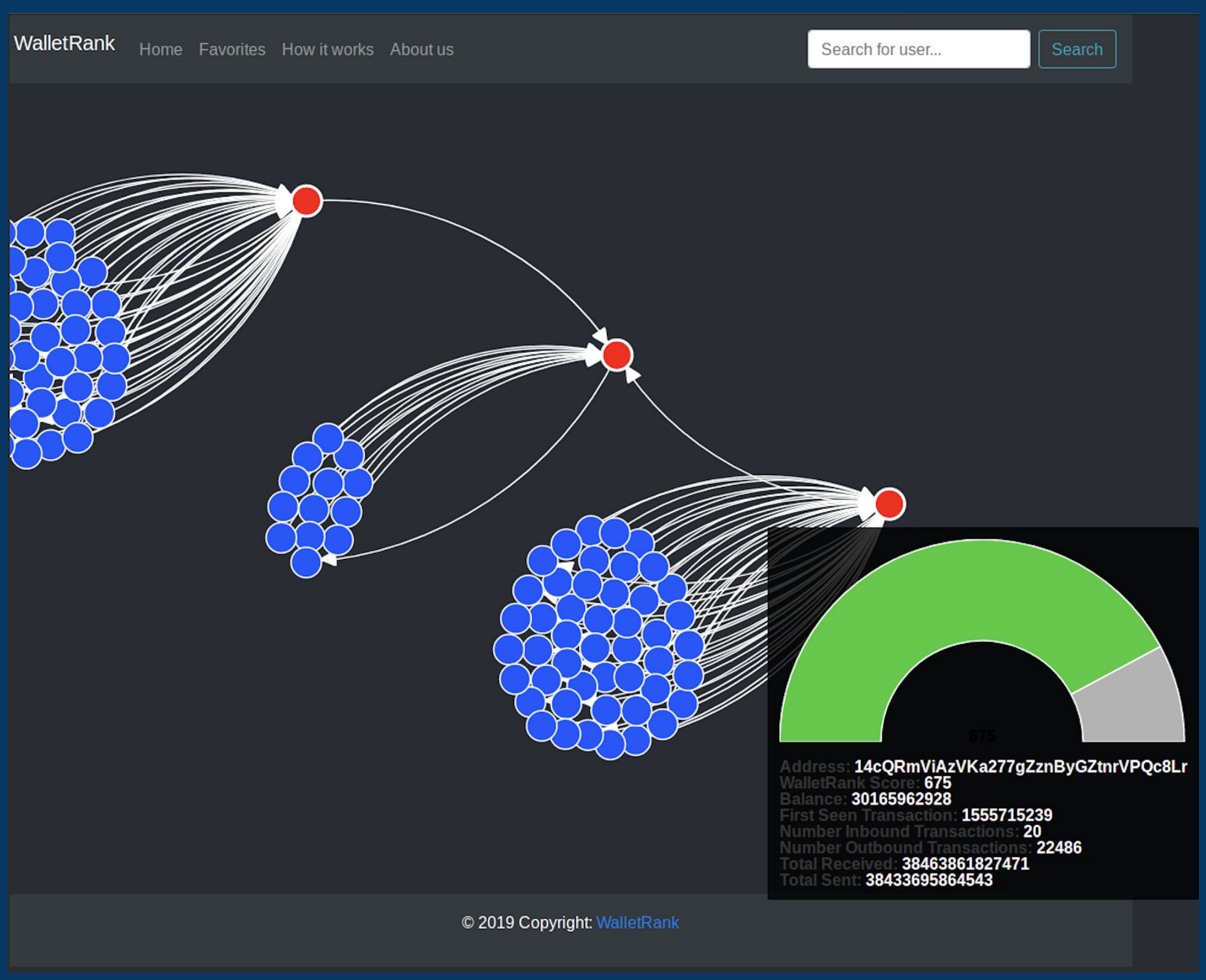
Address: 1NDyJtNTjmwk5xPNhjgAMu4HDHigtobu1s
WalletRank Score: 300
Balance: 1585861402658
First Seen Transaction: 1555742702
Number Inbound Transactions: 34
Number Outbound Transactions: 410931
Total Received: 529827045093585
Total Sent: 528241183690927

© 2019 Copyright: WalletRank



Address: 1QyMU8q4vSwQBueF9P7WkrRFeBe2FEhun
WalletRank Score: 675
Balance: 0
First Seen Transaction: 1528282811
Number Inbound Transactions: 20
Number Outbound Transactions: 12
Total Received: 9141216670
Total Sent: 9141216670

© 2019 Copyright: WalletRank



Address: 14cQRmViAzVKa277gZznByGZtnrVPQc8Lr
WalletRank Score: 675
Balance: 30165962928
First Seen Transaction: 1555715239
Number Inbound Transactions: 20
Number Outbound Transactions: 22486
Total Received: 38463861827471
Total Sent: 38433695864543

© 2019 Copyright: WalletRank