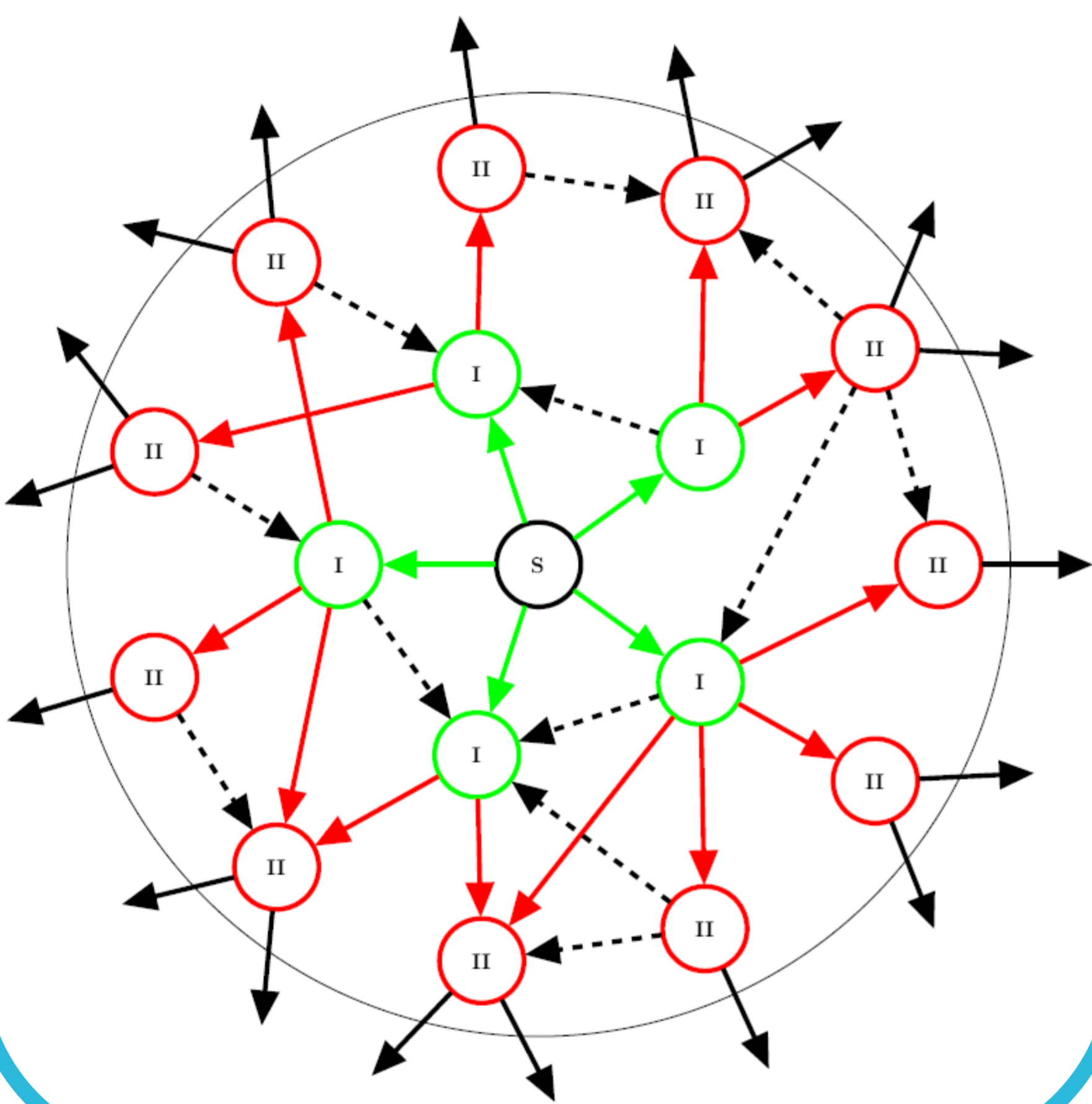


# Identification of Bitcoin Users with Naive Bayes Classifier Method

P. L. Juhász, J. Stéger, D. Kondor

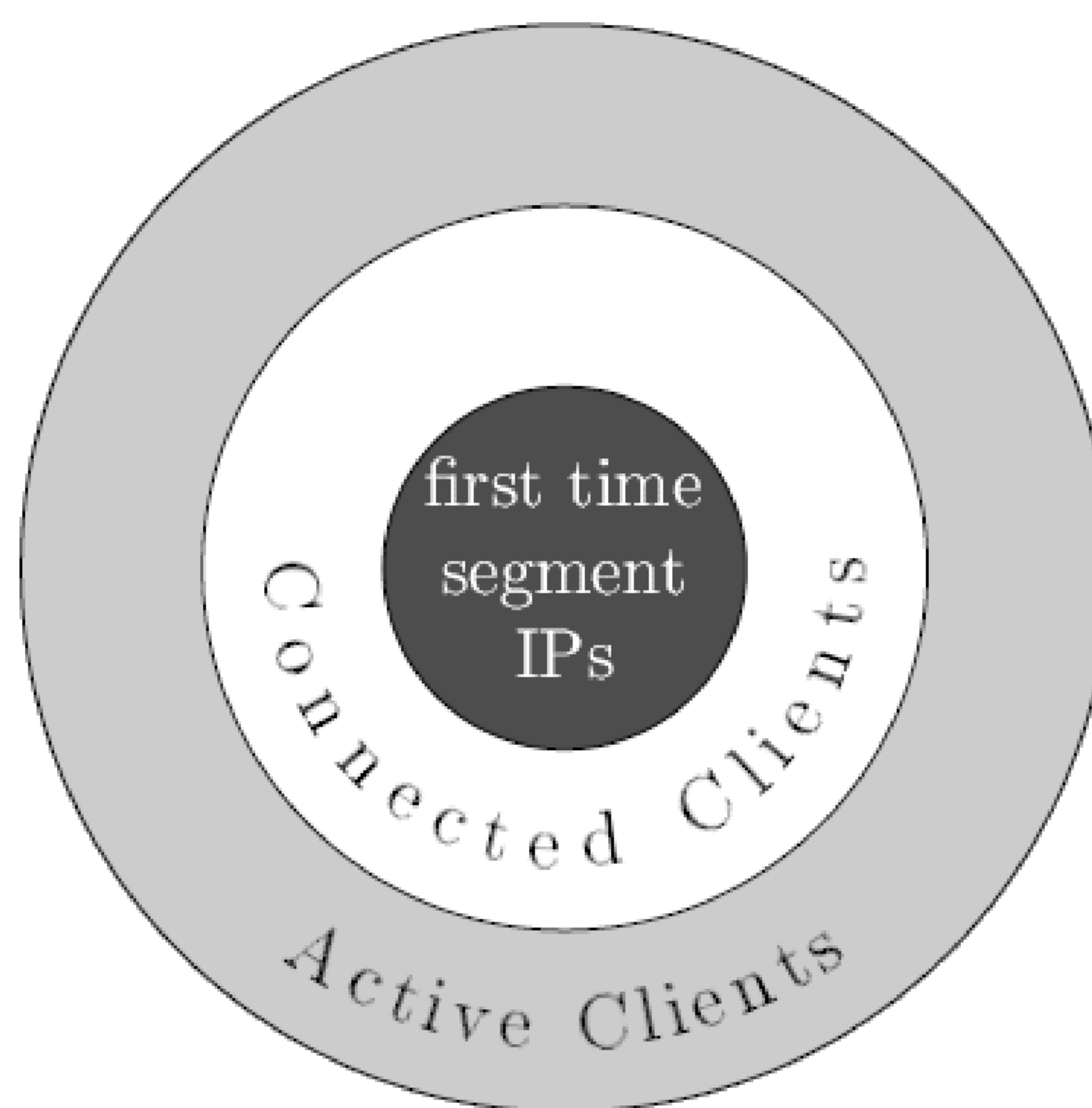
## Information Propagation

- S initiates a transaction
- Information is relayed hop-by-hop on network
- Transaction becomes known in the network



## First Time Segment

Clients advertising messages in the first time segment have higher probability of being the creator of the transaction.



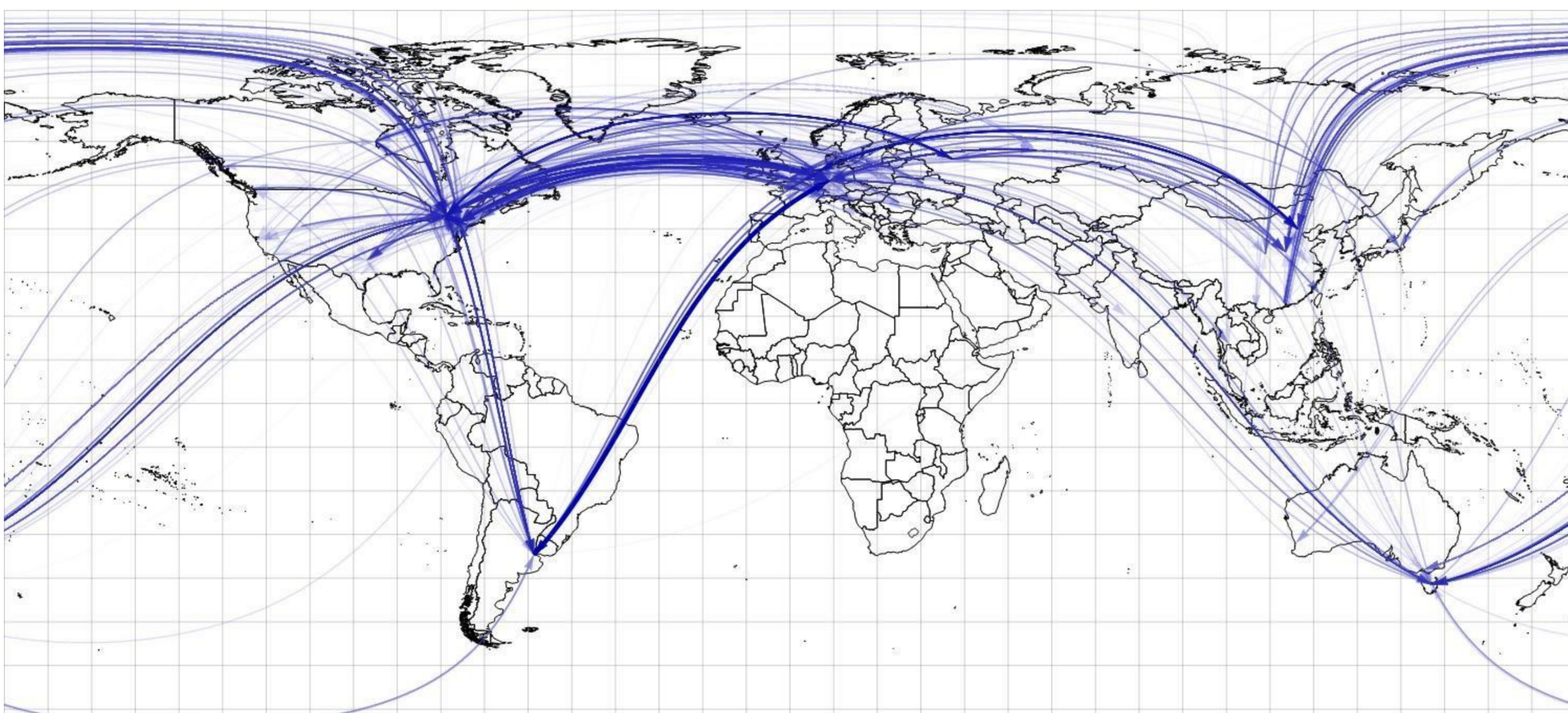
## Naive Bayes Classifier

- User transactions are grouped using Blockchain
- Transaction probabilities combined by Naive Bayes Classification

$$\mathbb{P}(C_k|\mathbf{d}) = \frac{\prod_{i=1}^m \mathbb{P}(C_k|d_i)}{\mathbb{P}(C_k)^{m-1} \sum_{\kappa=1}^K \frac{\prod_{i=1}^m \mathbb{P}(C_\kappa|d_i)}{\mathbb{P}(C_\kappa)^{m-1}}}$$

## Geographical Localization of Users

- MaxMind is used to resolve the locations of the clients based on their IP addresses
- Transactions are assigned to clients by the probability model



## Results

- We managed to track a portion of bitcoin currency and its flow across the most active countries

