Summary of this thesis

Crypto-currency or crypto-assets can provide a unique opportunity to perform a detailed study on financial transactions and interactions among users. Publicly available and big data accessible by the recent technology of distributed ledger, blockhain, help us to understand statistical properties and dynamics of economic network, in which users are interconnected with each other through money flow of transactions among each other and also in the exchange markets of crypto-currencies as well as fiat currencies. Users, who play dominant roles with respect to their frequencies and amounts of transactions, must have vital roles in the entire system of crypto-currencies. While anonymity of users is a core technology of blockchain, de-anonymization, if possible and even partially, helps to reveal various aspects in the ledger system of blockchain.

The purpose of this thesis is the de-anonymization of users, in particular, what we call big players and persistently-active ones, and the understanding significant properties in the dynamics of crypto-currency flow. I employ the blockchain of Bitcoin, in which all the transactions are recorded with a list of addresses, which are anonymous wallets, but can be partially identified as individual users. I constructed graphs or networks comprising of users or addresses as nodes and transactions or money flow as edges. Then I performed exploratory data analysis and network analysis in order to find significant patterns and interesting dynamics of the activities in the money flow. The thesis has the following three parts.

First I studied the daily time-series of transactions in their daily numbers and volumes during 2013 to 2018, when the generation of Bitcoin mining blocks was relatively stable. I focus on significant spikes in the transactions in the total number of transactions and total sum of volumes. By using smoothed periodogram or power-spectrum analysis for the time series, I found weekly pattern of these two variables, which implies that the financial organizations' trading systems are dominant roles giving higher activities during weekdays compared to weekends, which is similar to the exchange market of fiat currencies.

Second, following the above observation, I constructed daily networks and analyzed the network properties of the users as nodes and money flow attributed as edge flow circulated among users to focus on weekdays and weekends activities. I then performed an analysis using threshold for the flow of Bitcoin to define "big players" by proposing a method to identify financial institutions as those users satisfying certain criteria. The criteria concern about high frequency of appearance, in other words, appearing persistently on daily big transactions and showing a distinct weekly pattern of total average network flow. We were actually able to find known financial institutions as well as others.

Third I applied the method of non-negative matrix factorization (NMF) which can decompose the matrix of numbers and volumes of transactions into a certain number of components with relative weights. The purpose of such an analysis is to reveal hidden components in which users play different patterns of sending and/or receiving money. I proved that the NMF can be interpreted by a stochastic model. Then I performed simulations for a toy problem and estimated the parameters involved in the stochastic model in a framework of Bayesian estimation. From this result of

simulation, one can understand that the results of NMF can be interpreted as the probabilities of relative weights and the vectors corresponding to main senders and receivers. In the real data of Bitcoin, I found that there are actually big players that were already identified as financial institutions and also as others in the second part above. Moreover, I applied the method to temporal change of the network and found that the dynamics has a stable structure corresponding to the same components as well as a slowly changing dynamics.