Master's Thesis

Machine Learning based Bitcoin Address Classification

Chaehyeon Lee (이 채 현)

Department of Computer Science and Engineering

Pohang University of Science and Technology

2020

머신 러닝 기반의 비트코인 주소 분류

Machine Learning based Bitcoin Address Classification

Machine Learning based Bitcoin Address Classification

by

Chaehyeon Lee
Department of Computer Science and Engineering
Pohang University of Science and Technology

A thesis submitted to the faculty of the Pohang University of Science and Technology in partial fulfillment of the requirements for the degree of Master of Science in the Computer Science and Engineering

Pohang, Korea
12. 18. 2019
Approved by
James Won-Ki Hong (Signature)
Academic advisor

Machine Learning based Bitcoin Address Classification

Chaehyeon Lee

The undersigned have examined this thesis and hereby certify that it is worthy of acceptance for a master's degree from POSTECH

12. 18. 2019

Committee Chair	James Won-Ki Hong	(Seal)
Member	Jae-Hyoung Yoo	(Seal)
Member	Gwangsun Kim	(Seal)

MCSE 이 채 현. Chaehyeon Lee
20182819 Machine Learning based Bitcoin Address Classification,
머신 러닝 기반의 비트코인 주소 분류
Department of Computer Science and Engineering , 2020,
35p, Advisor : James Won-Ki Hong. Text in English.

ABSTRACT

The bitcoin network is a decentralized system that uses a peer-to-peer network structure to enable cryptocurrency transactions without the intervention of a third party. Participating nodes maintain the same transaction data, so that transparent trades can be made and that blockchain data cannot be forged or modulated. A bitcoin address is required for trading and maintains anonymity for the owner. By exploiting this anonymity, various illegal activities are conducted across the network. To detect and deter illegal transactions, this paper proposes a method of identifying the characteristics of bitcoin addresses related to illegal trades. We extracted 80 features extracted from bitcoin transactions. Using machine-learning techniques, we successfully categorized addresses involved with illegal activities with a ∼84% accuracy. We also examined the address features most affecting their classification and distribution and classified two machine-learning models. We also surmised that if we were to apply majority voting based on the results of classification, we could further specify the category to

which a transaction belongs. The results of the experiment showed that bitcoin addresses related to the Silk Road were very precisely classified, demonstrating the possibility of judging the illegality of transactions in the future.

Contents

I. Introduction **1**

II. Background and Related Work **4**

2.1 Background 4

2.1.1 Bitcoin Address 4

2.1.2 Bitcoin Transaction 4

2.1.3 Random Forest Classifier 6

2.1.4 Artificial Neural Network (ANN) 6

2.2 Related Work 7

III. Address Classification Methodology **10**

3.1 Transaction Collection 10

3.2 Address & Feature Extraction 13

3.2.1 Address Extraction 13

3.2.2 Feature Extraction 16

3.2.3 Labeling 17

3.3 Design of Machine-learning Models 20

3.4 Training & Testing of the Machine-learning Models 20

IV. Experiments and Results **22**

4.1 Dataset Configuration 22

4.2 Evaluation 23
4.2.1 Feature Importance 23
4.2.2 Classification Performance Comparison 25

V. Concluding Remarks **28**

Summary (in Korean) **31**

References **32**

List of Tables

3.1 The number of collected transactions by categories 12
3.2 The number of extracted addresses by categories 14
3.3 The label of each categories 17
3.4 The list of extracted features 18

4.1 Accuracy of the random forest classifier 26
4.2 Performance of random forest classifier by category 26
4.3 Accuracy of the ANN 27
4.4 Performance of ANN by category 27

List of Figures

1.1 Bitcoin values (USD) sent to darknet markets from 2011 to 2018. Orange line shows the proportion of darknet Bitcoin transactions over all transactions(Source: Chainalysis) 2

2.1 The types of Bitcoin transactions 5

3.1 Classification Methodology 11

3.2 Comparison of size distribution of extracted addresses by category 15

3.3 The screenshot showing a portion of the extracted features 19

3.4 Prediction results of the test set 21

4.1 Distribution ratio of transactions and addresses 22

4.2 Feature Importance; (a)the top 10 features (b)the top 20 features . 24

I. Introduction

In 2008, Satoshi Nakamoto produced a white paper about a peer-to-peer electronic payment system [1]. Over time, the price of bitcoin has dramatically fluctuated, and people worldwide have made trades. Blockchain [2] is a decentralized transaction technique in which participants maintain duplicate copies of temporally connected ledger data, called "blocks". Anyone in the network can duplicate the blockchain structure and can validate data on the network. Thus, the bitcoin network is autonomously maintained and operated by thousands of participating nodes without a central authority, assuring transparent transactions. This disintermediation has allowed cross-border value transfers between buyers and sellers having very low transaction fees and scant processing. Bitcoin employs a proof-of-work consensus algorithm that makes it impossible to maliciously delete, forge, or modify existing data. One must have a bitcoin address to send bitcoins, and a single user can have multiple addresses.

However, because it is nearly impossible to infer owner information from the bitcoin address, there are frequent cases of illegal transactions. In fact, there have been a variety of darknets that abuse bitcoin for illegal use [3] [4] and statistics show that the total dollar value of bitcoin traversing the "dark net" has steadily increased since 2011 (Fig 1.1 [5]). By 2013, nearly 1M users were trading bitcoin. In 2017, when the trading value reached its highest value of USD 707M, most was traded through darknet markets. The Silk Road [6] was one of the most famous

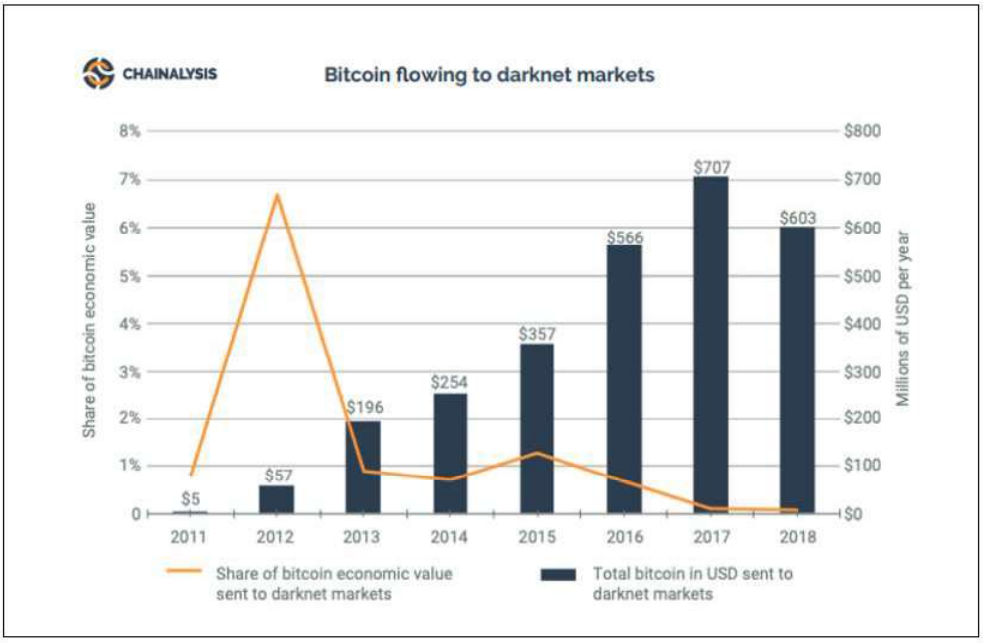

Figure 1.1: Bitcoin values (USD) sent to darknet markets from 2011 to 2018. Orange line shows the proportion of darknet Bitcoin transactions over all transactions(Source: Chainalysis)

online black markets, trading drugs, weapons, child pornography, stolen goods, and malicious code.

In addition to illegal goods transactions, illegal activities such as money laundering and scamming are acting as a factor in hindering the enactment of cryptocurrency laws. Therefore, it is necessary to find a way to detect illegal transactions on blockchains. Although various cryptocurrencies such as bitcoin, Ethereum [7], and Monero [8] are used for illegal transactions in darknet, since bitcoin is the most actively used cryptocurrency on darknet, we focused on bitcoin

and studied the methodology to detect illegal transactions on bitcoin networks.

Because illegal users are likely to repeat transactions, and one user can leverage multiple bitcoin addresses, we classified the characteristics of bitcoin addresses to help detect illegal activities. The address characteristics associated with illegal transactions can be analyzed by collecting the transaction lists of known illegal trades. Machine learning classification models [9] can then be used to train the features so that such activities can be identified.

In this paper, we present a methodology for detecting illegal bitcoin addresses, and we then explain the detailed process of detection. Section II explains the background and several related works. Section III describes the classification process and its implementation. In Section IV, we present the results of several experiments. A broad discussion and conclusion with future works are provided in Section V.

II. Background and Related Work

2.1 Background

2.1.1 Bitcoin Address

Digital keys, addresses, and digital signatures are used to prove ownership of bitcoin ownership [10] [11]. Digital keys, comprising private and public keys, are stored in digital wallets, which are simple databases. A public key is used to receive the bitcoin, and a private key is used to sign the transaction to consume the transmitted bitcoin. The public key is generated from the private key, and in most cases, bitcoin addresses can be generated from the public key. Addresses can be infinitely generated, and the wallets can generate and maintain multiple bitcoin addresses indefinitely. When generating a transaction, the transmitter must specify the recipient's bitcoin address, which is shared with others. Bitcoin transactions transfer the ownership of bitcoin to the address of the recipient, and the blockchain is updated. During this process, personal information is neither collected nor transmitted.

2.1.2 Bitcoin Transaction

There are several types of bitcoin transactions having different input and output values. Fig 2.1 (a) shows the most common type of transaction: one output of bitcoin remittance from one input value, and another output that returns the remaining balance to the original owner. Because bitcoin lacks a

mechanism that automatically returns remaining bitcoin to its original owner, the owner must generate an output that performs this function. The transaction of Fig 2.1 (b) sends multiple inputs to one address, and that of Fig 2.1 (c) allows multiple output values in one transaction for distributing and sending bitcoins to multiple addresses. The first transaction type can be included in the third type. As shown in Fig 2.1 (d), it is also possible to generate transactions having multiple inputs and multiple outputs. The sum of the input values should be greater than or equal to the sum of the output values, and the difference between the two sums becomes the transaction fee.

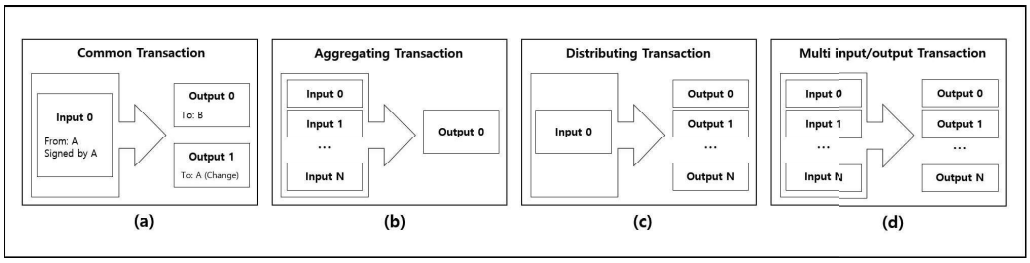

Figure 2.1: The types of Bitcoin transactions

2.1.3 Random Forest Classifier

The random forest [12] algorithm creates decision trees for classification and regression analysis. The decision tree produces various decision paths and results, and the final decisions are made by answering questions at each element from root to leaf. To construct a decision tree, one must decide the features to be included and the depth of the tree. Random forest randomly selects the elements of each tree. Using an ensemble technique, it creates a number of decision trees and determines the final result by majority voting. Random forest is easy to understand and interpret, and it can simultaneously handle numeric and categorical data. For this paper, we implemented a classification model using a random forest algorithm to determine the category of a given address.

2.1.4 Artificial Neural Network (ANN)

An ANN [13] mimics the nervous system of a living organism. The artificial neurons (nodes) in an ANN are abstractions of nerve cells (neurons). As a learning algorithm, it mimics human intelligence by replicating the behaviors of neurons that receive stimuli and conveying them to another neuron. The nodes in an ANN are interconnected via several layers. The data to be learned are inserted through the input layer, processed in one or more hidden layers, and outputs the final result through the output layer.

2.2 Related Work

Several authors have suggested methodologies for detecting types of bitcoin activities by using machine-learning methods, and they have evaluated their performance via experiments. By combining various features of bitcoin networks and machine-learning techniques, they were able to determine whether or not activities were legal.

Zambre and Shah [14] proposed a machine-learning-based system that determined the characteristics of users related to bitcoin thefts and identifies those performing similar actions. To detect bitcoin thefts and fraudulent activity, they analyzed the transaction information of several famous thefts [15] [16]. They extracted 22 features to segregate dishonest users from honest users and clustered them using a k-means [17] clustering algorithm to identify theft behaviors, achieving 76.5 % accuracy.

Toyoda et al. [18] identified bitcoin addresses related to a high yield investment program (HYIP) by analyzing transaction patterns. They manually identified HYIP and non-HYIP addresses and extracted several features, such as the number of transactions associated with the bitcoin address and the number of blocks mined. A pattern was assigned to each transaction, and the frequency of each pattern was utilized as a key feature. They labeled the bitcoin address as "HYIP" or "non-HYIP" for classifying cybercrime groups via supervised learning. About 83% of the HYIP-related addresses were correctly classified.

Kanemura et al. [19] analyzed bitcoin transactions and addresses related to darknet markets and proposed a voting-based system that determined the

labels of multiple addresses controlled by the same entity based on the number of the majority labels. They identified the characteristics of transactions related to darknet markets (DNM [20]) that could be used to identify newly generated DNM transactions. They extracted 73 features and used them to train the supervised classifiers. The proposed voting methods achieved an ∼0.8 F1 score.

In a previous work [21], we conducted research to detect illegal transactions based on their characteristics. Although bitcoin addresses and clusters associated with criminal activity have been identified and classified several times, classification from transaction features alone has not been reported. Our previous work extracted nine features and added one label, giving 10 features for each transaction. We used them to train supervised-learning classification models, which ultimately achieved an F1 score of ∼0.9. However, the test set may have been over-fitted, and the number of features used to determine the illegality of the transaction was probably too small.

Following these and previous studies, we have extended our scope to detect illegal transactions using the characteristics of bitcoin addresses rather than transactions. We increased the number of features to be extracted and checked which ones most affected the classification model.

Chainalysis, Inc., a company specializing in cryptocurrency security technology, conducts services to track abnormal cryptocurrency transactions and provides digital forensics [22]. It tracks the details of transactions and monitors whether they are legal. When a transaction occurs, a suspicious pattern or account activity informs the relevant agency. We, therefore, propose a system that

can detect illegal transactions of bitcoin networks such as Chainalysis using our proposed methodology.

III. Address Classification Methodology

To classify bitcoin addresses and detect illegal transactions, we designed a four-step methodology comprising transaction collection, bitcoin address feature extraction, machine-learning training, and testing (Fig 3.1). We collected several types of transaction hash lists and derived bitcoin transmission and reception addresses. We extracted 80 address features that were assigned different labels. Labeled data were learned by the machine-learning classification models, and the trained models were used to determine the classification to which the given bitcoin address belonged. The performance of the classification models was verified using F1 score. The following subsections describe each step in detail.

3.1 Transaction Collection

Before implementing the machine-learning model, we collected transaction hash lists from a publicly available forum, WalletExplorer.com [23], which discloses categories of data used for specific groups (e.g., exchanges, mining pools, services, dark nets). We focused on five categories: mixers, exchanges, gambling, pools, and Silk Road. We built a simple web crawler using Python and the Beautiful Soup library [24] to obtain a list of hash values for transactions. Data in all categories except Silk Road were collected beginning in January 1, 2016. For Silk Road, only data prior to 2018 was collected, because that is when the site was whut down. The number of transactions collected is specified in the table 3.1.

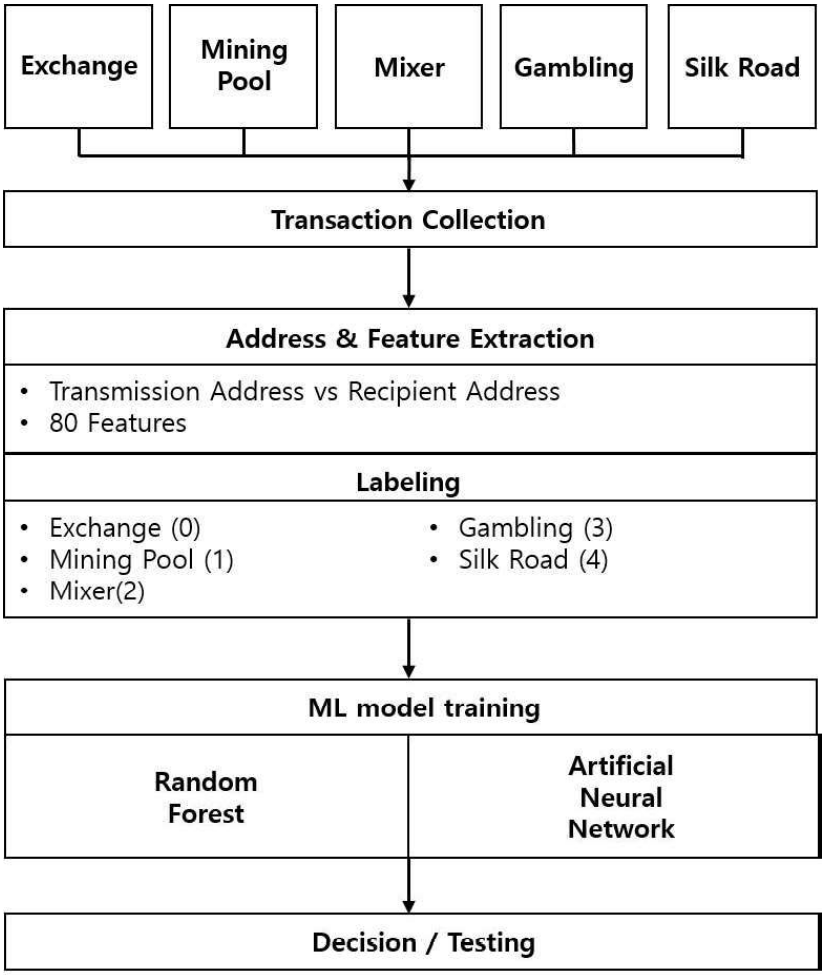

Figure 3.1: Classification Methodology

Table 3.1: The number of collected transactions by categories

Category	**The number of collected transactions**
Exchange	761,494
Mining Pool	325,800
Mixer	93,200
Gambling	752,300
Darknet (Silk Road)	956,186

Note that experimental dataset does not necessarily reflect the proportions of real distributions in the bitcoin network. The collected data constituted only a fraction of transactions, and only a portion of the collected data were learned to alleviate any data imbalance.

3.2 Address & Feature Extraction

3.2.1 Address Extraction

More than one transmission and reception address can be extracted from a bitcoin transaction. We obtained transaction details using JavaScript Object Notation (JSON) remote-procedure calls (RPC) [25] [26]. The transmission addresses are obtained by referring to the [vin] field of transaction details, and the bitcoin reception addresses are extracted by referring to the [vout] field. Depending on the type and category of transactions, the numbers of transmission and reception addresses varied. A particular bitcoin address might only serve one bitcoin, or it might receive only one, but it may also be used to transmit and receive at the same time.

Table 3.2 below presents the number of total transactions by category, the number of bitcoin addresses associated with each transmission, the number of addresses associated with bitcoin receptions, and the number of total addresses. Fig 3.2 shows address distribution per category. For mining pools and mixers, a relatively small number of addresses were extracted, because certain addresses used for these services likely appeared repeatedly across the transactions. However, because the number of users exploiting these service as large, the number of addresses extracted from transactions was very large.

Table 3.2: The number of extracted addresses by categories

Category	**Transmission addresses**	**Recipient addresses**	**Total addresses**	**Transactions**
Exchange	1,395,325	6,736,265	8,665,943	761,494
Mining Pool	218,476	1,036,143	1,375,327	325,800
Mixer	178,721	480,754	718,915	93,200
Gambling	726,210	3,960,029	5,345,783	752,300
Darknet (Silk Road)	704,376	938,730	2,305,872	956,186

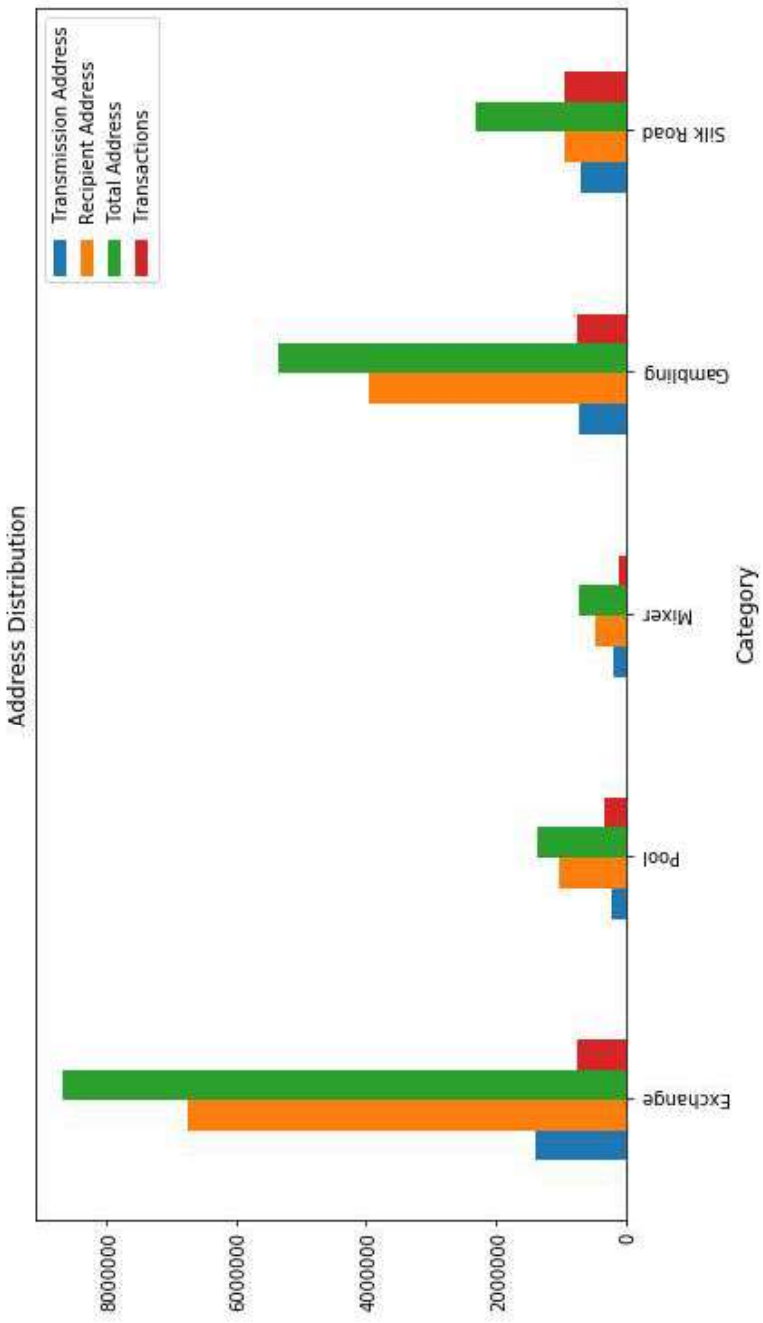

Figure 3.2: Comparison of size distribution of extracted addresses by category

3.2.2 Feature Extraction

After extracting the transaction list, we extracted the key features for training. Illegal transactions exhibited common characteristics, such as high transaction fees in order to have them quickly included in blocks, multiple identical outputs inside one transaction (indicating money laundering), and multiple address distributions. To identify the common patterns associated with illegal transactions, We extracted features related to the addresses obtained from each.

We selected 28 features, and for some, we obtained four values: average, total, minimum, and maximum. When a specific bitcoin address appeared several times in a collected transaction, the feature values were updated per the incremental values and classified either as transmission or the reception. Depending on the category, some feature values were filled with -1. If a specific address was a transmission address, the feature values related to reception were set to -1, and, if a specific address was a reception address, the feature values related to transmission were filled with -1. A bitcoin address not having a value of -1 indicates that it transmitted or received bitcoins. We extracted 80 features, including those related to transmission and reception.

A Python script returned the transaction details of a given transaction hash from the JSON-RPC calls, extracting the relevant features (i.e., bitcoin transmission and reception amounts, transaction fees, number of inputs associated with transmission, number of outputs associated with reception, number of transmission addresses associated with transmission, number of reception addresses associated with reception). Table 3.4 provides simple descriptions of features

extracted for bitcoin-address classification.

3.2.3 Labeling

When training supervised-learning-based classification methods, training data must be labeled. Therefore, after extracting the features of each address, we manually labeled each according to the classification of the corresponding transaction (Table 3.3).

Table 3.3: The label of each categories

Category	**Label**
Exchange	0
Mining Pool	1
Mixer	2
Gambling	3
Darknet (Silk Road)	4

Table 3.4: The list of extracted features

Name	**Description**	**Etc**
Bitcoin amount (transmit/receive)	Transmitted/Received bitcoin amount	
Total bitcoin amount (transmit/receive)	The amount of total bitcoin transmitted/received by the transaction associated with the transmission address	
Transaction fee (transmit/receive)	Transaction fees associated with bitcoin transmission/reception	
Sibling inputs/outputs (transmit/receive)	The number of sibling inputs/outputs	avg
Sibling inputs/outputs_out/in (receive/transmit)	The number of outputs/inputs associated with bitcoin transmission/reception	sum min
Unique address (transmit/receive)	The number of unique transmission/receiving addresses	max
Unique address_out/in (receive/transmit)	The number of unique receiving/transmission addresses associated with bitcoin transmission/reception	
Transaction Size (transmit/receive)	Transaction size associated with bitcoin transmission/reception	
Block Interval (transmit/receive)	The interval of the blocks related to the transmission/reception transaction	
Relevant transaction number (transmit/receive)	The number of transactions associated with the transmission/receiving address	
Lifetime (transmit/receive)	Life time of the transmission/receiving address	
First block (transmit/receive)	Block height where the transmission/receiving address first appeared	
Total transaction number	Total number of transactions associated with the address	
Total life time	Lifetime of the address	
Label	Classification of the address	

Address	btc_trns_avg	btc_trns_sum	btc_trns_ain	btc_trns_max	btc_recv_avg	btc_recv_sum	btc_recv_ain	btc_recv_max	trns_value_avg	trns_value_sum	trns_value_ain
3MchpCTGdhMrrSXDho4MemUYvQNaSmal42	-1.000000	-1.000000	-1.000000	-1.000000	0.015616	0.015616	0.015616	0.015616	-1.000000	-1.000000	-1.000000
33efq9TxU88PBiQtvFrQ2ZKNnBDrhTpEbK	-1.000000	-1.000000	-1.000000	-1.000000	0.000589	0.031816	0.000231	0.005348	-1.000000	-1.000000	-1.000000
1PxBNjK7tajzufgd12oPW6RoisWt8n1ULL	1.741667	10.450000	0.377068	2.949620	1.492857	10.450000	0.377068	2.949620	184.842314	1109.053887	2.011544
12DRARTmSxEEjHmryxm8vdtHbGng6j9HSR	0.012063	0.012063	0.012063	0.012063	-1.000000	-1.000000	-1.000000	-1.000000	4.012063	4.012063	4.012063
bc1q0ufnl6k64ywqt6kt3cyepglwmxlen8rmrdqk74	15.000000	15.000000	15.000000	15.000000	-1.000000	-1.000000	-1.000000	-1.000000	15.000000	15.000000	15.000000
...	...	...	...	...	...	...	...	...	...	...	...
1PBBWNEZLtWUbFNoU3Rx3ZvsQpByv2yBH	0.290700	0.290700	0.290700	0.290700	-1.000000	-1.000000	-1.000000	-1.000000	0.290700	0.290700	0.290700
38uCkdFa6u7k2QuiSEho8xBE5qRPgkt4T5	-1.000000	-1.000000	-1.000000	-1.000000	0.000148	0.000148	0.000148	0.000148	-1.000000	-1.000000	-1.000000
1L1fd9NG1gExdsiMzEtFB3U1T6ne5z7Vat	-1.000000	-1.000000	-1.000000	-1.000000	0.045459	0.045459	0.045459	0.045459	-1.000000	-1.000000	-1.000000
1H2ZFSSMxqT8CkaUffAovmJcG1qZhmKott	0.007627	0.007627	0.007627	0.007627	0.007627	0.007627	0.007627	0.007627	1.208980	1.208980	1.208980
39gZxUcBw7hiBDFbS7FAZrxUqjXDqsyiJ3Z	-1.000000	-1.000000	-1.000000	-1.000000	0.008295	0.008295	0.008295	0.008295	-1.000000	-1.000000	-1.000000

blk_btw_recv_avg	blk_btw_recv_sum	blk_btw_recv_ain	blk_btw_recv_max	rel_tx_trns_cnt	rel_tx_recv_cnt	total_tx_cnt	lifetime_trns	lifetime_recv	lifetime_total	init_trns_block	init_recv_block	label
0	0	0	0	0	1	1	-1	31123	31123	-1	574458	0
812	43857	0	6354	0	54	54	-1	43857	43857	-1	438283	0
5	37	0	19	6	7	13	536	37	1241	219277	218372	4
-1	-1	-1	-1	1	0	1	374208	-1	374208	230636	-1	4
-1	-1	-1	-1	1	0	1	37047	-1	37047	567878	-1	0
...	...	...	...	...	...	...	...	...	...	...	...	...
-1	-1	-1	-1	1	0	1	196262	-1	196262	405573	-1	1
0	0	0	0	0	1	1	-1	165944	165944	-1	440764	0
0	0	0	0	0	1	1	-1	119160	119160	-1	482454	2
0	0	0	0	1	1	2	0	119788	8	481834	481826	2
0	0	0	0	0	1	1	-1	61839	61839	-1	539782	2

Figure 3.3: The screenshot showing a portion of the extracted features

3.3 Design of Machine-learning Models

For classification, we used two machine-learning models: random forest and ANN. The addresses were classified into one of five categories. The models were implemented on the application programming interface provided by sklearn [27] and Tensorflow [28]. The ANN model comprised one input layer having 80 features, one hidden layers with 50 nodes, and one output layer.

3.4 Training & Testing of the Machine-learning Models

After extracting the relevant address features, we trained our supervised-learning classification algorithms on the assigned labels. When the training phase was complete, the classification model could distinguish the associated feature values for each category. During the test pase, the classifier predicted where the classification of each address in the test set belonged using trained classifiers. To determine whether the model trained the training set well to enable the derivation of the correct classification results, we measured accuracy by comparing the initial labels with those predicted by the models. Fig 3.4 shows the results of the trained machine-learning.

lifetime_recv	lifetime_total	init_trns_block	init_recv_block	curr_trns_block	curr_recv_block	label	predicted
28610	28610	-1	577060	-1	577060	0	0
81636	61	520045	519984	520045	519984	2	2
-1	349145	256237	-1	256237	-1	4	4
371568	76	233367	233291	233367	233291	4	4
26277	26277	-1	466699	-1	492976	2	2
153551	7747	456638	448891	456638	448891	3	3
143460	143460	-1	458148	-1	458148	2	2
-1	380784	224025	-1	224025	-1	4	4
-1	46194	558509	-1	558509	-1	0	3
169877	169877	-1	432065	-1	432065	1	3
70557	70557	-1	531063	-1	531063	2	2
3716	3716	-1	548163	-1	551879	2	0
157921	157921	-1	444491	-1	444491	3	3
-1	33363	572057	-1	572057	-1	0	2
10887	10887	-1	595312	-1	595312	0	0
-1	378697	226127	-1	226127	-1	4	4
21590	21590	-1	584357	-1	584357	0	0
4763	4763	-1	438068	-1	442831	3	3
114093	114093	-1	493579	-1	493579	0	0
192875	192875	-1	408968	-1	408968	1	1

Figure 3.4: Prediction results of the test set

IV. Experiments and Results

4.1 Dataset Configuration

We collected several transactions for each category, and the number of addresses extracted from each transaction differed per category (Fig 4.1). The total extracted addresses was 18M, and, owing to hardware limitations, all data could not be trained. Therefore, the experiments were conducted by randomly selecting datasets. We set the datasets to different sizes to test the model and conducted the experiments several times. Prior to training, we defined the size of the training and test sets. The training:test split was set to 60:40 for each experimental dataset.

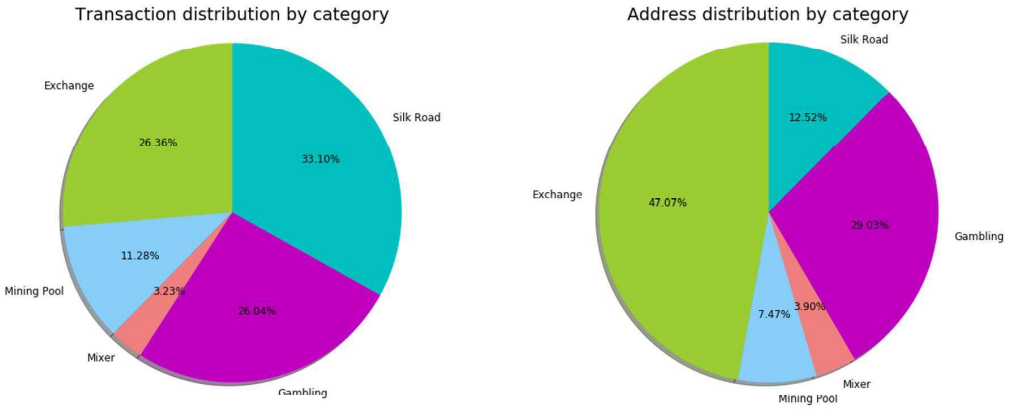

Figure 4.1: Distribution ratio of transactions and addresses

4.2 Evaluation

4.2.1 Feature Importance

We investigated feature importance [29], of which 80 features most affected classification performance. Fig. 4.2 shows the top 10 and 20 features considered to be the most important of the 80.

We did not measure the feature importance of all datasets because of hardware limitations. Therefore, we collected 2,000 data items per category and examined feature performances of 10,000 datasets. Fig. 4.2(a) shows the top 10 features that were most important. Fig. 4.2(b) shows the importance of the top-20 features.

The experimental results show that the characteristics related to bitcoin reception were greatly affected. This can be attributed to the fact that the received address occupied a large part of the collected dataset. Lifetime was the most influential feature and indicated whether the address was used continuously and how long it was active. When a service having the same address is repeatedly used, it has a relatively long lifetime. Additionally, the second-most important feature was the amount of bitcoin received by the address. The third-most important feature was the total amount of bitcoin received by the transaction generated when the address received the bitcoin. The size of the transaction and the transaction fee had the greatest impact on the classification model.

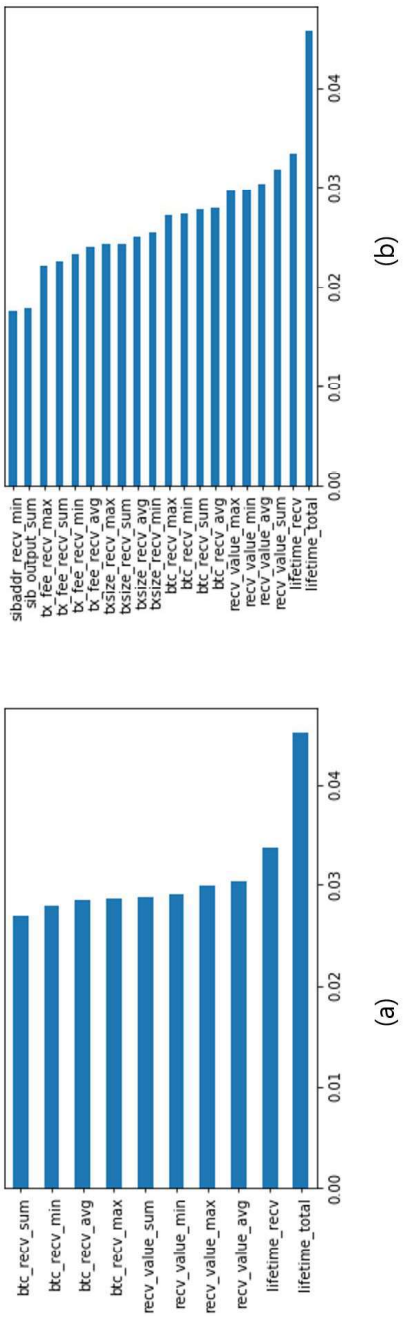

Figure 4.2: Feature Importance; (a)the top 10 features (b)the top 20 features

4.2.2 Classification Performance Comparison

After training the extracted data, we used the two models to classify address. We repeated the experiment several times using different dataset sizes from the five categories, and we checked the differences of performance according to dataset size. We measured the accuracy as a performance index and checked whether the classification was well done using precision, recall, and F1 score values [30] [31] [32].

Random Forest Classifier

We randomly selected 1,000 to 200,000 data items for each category. Therefore, the accuracy of the random forest classifier was measured by setting the total dataset between 5,000 and 1,000,000. As a result of the experiments, we found that the accuracy increased steadily as the size of the dataset increased (Table 4.1). Our experiments showed an accuracy of ∼0.84, and it is expected that the accuracy could be better if dataset size were extended.

Table 4.2 shows the precision, recall, and F1 scores of the results of the experiment from 200,000 data items for each category. These values show how well the classification was done for each category. In particular, in the case of the address corresponding to Silk Road, the scores had a value of 1.0, indicating that the address corresponding to Silk Road was well classified without error.

Table 4.1: Accuracy of the random forest classifier

Each data set	**Accuracy**
1,000	0.741
3,000	0.782
5,000	0.789
10,000	0.804
30,000	0.825
50,000	0.833
100,000	0.838
200,000	0.844

Table 4.2: Performance of random forest classifier by category

Category	**Precision**	**Recall**	**F1-Score**
Exchange	0.82	0.74	0.78
Mining Pool	0.86	0.85	0.86
Mixer	0.78	0.86	0.82
Gambling	0.77	0.77	0.77
Silk Road	1.0	1.0	1.0

ANN Model

We randomly selected 10,000 to 30,000 data items for each category to evaluate the ANN model. The total datasets were set from 50,000 to 150,000. The results are shown in Table 4.3. In the case of ANN, accuracy and F1 score were relatively lower than those of the random forest classifier. This shows that the result was not related to the increase of the size of the dataset, and the highest accuracy was 64%. Although the addresses associated with Silk Road were nearly as precisely classified as the random forest classifier, the mixer and gambling-related addresses were only classified at ∼50%(Table 4.4).

Table 4.3: Accuracy of the ANN

Each data set	**Accuracy**
10,000	0.646
20,000	0.620
30,000	0.614

Table 4.4: Performance of ANN by category

Category	**Precision**	**Recall**	**F1-Score**
Exchange	0.62	0.52	0.56
Mining Pool	0.77	0.56	0.65
Mixer	0.45	0.45	0.45
Gambling	0.51	0.46	0.48
Silk Road	0.99	0.98	0.99

V. Concluding Remarks

We classified various categories of bitcoin addresses using machine learning-based classification models. A transaction list was collected and sorted by five categories : exchange, mining pool, mixer, gambling, darknet. The associated bitcoin addresses were obtained from the transaction list. By extracting 80 features of bitcoin addresses and learning those extracted from the classification model, we successfully classified specific addresses. We used random forest and ANN algorithms as classification models, and the accuracy of random forest was 84%, which was relatively higher than that of ANN. We confirmed that the bitcoin addresses related to Silk Road were very well classified by both models.

This research contributes to two aspects of related studies. The related works used binary classification to classify bitcoin data. However, this study subdivided and specified classification by applying multiple classification. Furthermore, most previous studies about detecting illegal transactions used bitcoin address-clustering techniques. Bitcoin address clustering clustering technique means to bind bitcoin addresses, which are determined to be controlled by the same entity, to a single cluster. Most of the clustering algorithms are based on heuristic algorithms. However, in this case, the classification results differed depending on which heuristic algorithm was used, but the algorithm could not reflect the overall situation of the bitcoin network. Therefore, it was judged that the clustering techniques reduced the reliability and accuracy of each classifica-

tion. This study did not depend on heuristic algorithms, but instead utilized the characteristics of bitcoin addresses. Thus, we derived a relatively consistent classification result.

There were some limitations to study. First, the proposed ANN model delivered a low F1 score compared to the random forest classifier. This limitation might be overcome by adopting machine-learning based techniques. We could increase performance by adjusting the number of hidden layers or the number of nodes. It is also possible to reevaluate the performance by training the model using characteristics having high importance without using all 80 extracted characteristics. We also could apply other available deep-learning methods. Second, the obtained transaction data had already been labeled prior to acquisition. The test dataset in our experiments was not exposed during model learning, but it might have been previously trained by similar algorithms. In other words, the test dataset might have been exposed to a similar model. Because we obtained the test set using the same method as the training set, it may have been overfitted. Therefore, if we were to test the model on incoming/live transactions from the Bitcoin network, the measured F1 score might be lower than the experimental values reported here.

We should next predict address classifications associated with certain transactions by applying the proposed methodology while predicting the category of transactions by applying majority voting to the results. In future works, we plan to access the dark nets and collect a transaction list on currently operating sites, because the Silk Road has been closed for years. Then, we plan to apply the

proposed methodology to check whether the addresses related to dark net markets are accurately classified and to check whether transactions are generated for those markets by adopting majority voting. This is not limited to data in the darknet market, but can be applied to other services such as mixers. Furthermore, we plan to predict whether a transaction is legal depending on the category of transaction. If the research related to illegal transaction detection is successfully performed, we will design and develop a system that classifies a real-time address category, a real-time transaction category, and real-time illegal transaction detection.

요 약 문

블록체인을 기반으로 하는 비트코인은 P2P 네트워크 구조의 탈 중앙화 시스템으로, 제3자의 개입 없이 암호화폐의 거래가 가능하다. 참여자가 동일한 데이터를 유지함으로써 투명한 거래가 가능하며 데이터의 위/변조가 불가능하다는 특징이 있어 크게 주목받고 있다. 비트코인을 거래하기 위해서는 비트코인 주소가 필요한데, 이 주소는 실 사용자의 정보를 연결하지 않는다는 익명성을 띠고 있다. 이러한 익명성을 악용하여, 비트코인 네트워크에서 다양한 불법 거래들이 활발하게 일어나고 있고 피해가 막심하다.

비트코인에서 발생하는 불법 거래를 탐지하기 앞서, 본 논문에서는 거래와 관련된 비트코인 주소의 특성을 파악하고 주소의 분류를 예측하는 방법론을 제안한다. 여러 서비스(거래소, 마이닝 풀, 믹서, 겜블링, 암거래 시장 - 실크로드)에 활용된 트랜잭션을 카테고리 별로 수집하고, 수집된 트랜잭션으로부터 연관된 비트코인 주소와 80개의 특징을 추출한다. 그리고 머신 러닝 모델을 이용해 특정 비트코인 주소가 어떤 카테고리에 속하는지 분류해보았고 최고 약 84%의 분류 정확도를 가짐을 확인했다. 특정 트랜잭션과 연관된 비트코인 주소들의 분류를 예측한 결과를 바탕으로 Majority voting을 적용한다면, 더 나아가 해당 트랜잭션이 어떤 카테고리에 속하는지 판단할 수 있다. 실험 결과, 암거래 시장과 관련된 비트코인 주소가 거의 정확히 분류됨을 확인할 수 있었고, 이는 추후 트랜잭션의 불법 여부를 판단할 수 있는 가능성을 보여준다.

References

[1] Satoshi Nakamoto et al. Bitcoin: A peer-to-peer electronic cash system. 2008.

[2] Melanie Swan. *Blockchain: Blueprint for a new economy.* " O'Reilly Media, Inc.", 2015.

[3] Campbell R Harvey. Bitcoin myths and facts. *Available at SSRN 2479670*, 2014.

[4] Bitcoin Magazine. Bitcoin magazine: Bitcoin news, bitcoin charts, events. Available at `https://bitcoinmagazine.com/articles/darknet-markets-cant-live-with-or-without-bitcoin`.

[5] Chainalysis. Chainalysis: The blockchain analysis company. Available at `https://www.chainalysis.com/`.

[6] Wikipedia. Silk road (marketplace). Available at `https://en.wikipedia.org/wiki/Silk_Road_(marketplace)`.

[7] Gavin Wood et al. Ethereum: A secure decentralised generalised transaction ledger. *Ethereum project yellow paper*, 151(2014):1–32, 2014.

[8] monero. Zero to monero. Technical report, 2018. Available at `https://www.getmonero.org/library/Zero-to-Monero-1-0-0.pdf`.

[9] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, 160:3–24, 2007.

[10] Stefan Brands. *Rethinking public key infrastructures and digital certificates: building in privacy.* Mit Press, 2000.

[11] Addison M Fischer. Public key/signature cryptosystem with enhanced digital signature certification, September 19 1989. US Patent 4,868,877.

[12] Mahesh Pal. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1):217–222, 2005.

[13] Jacek M Zurada. *Introduction to artificial neural systems*, volume 8. West publishing company St. Paul, 1992.

[14] Deepak Zambre and Ajey Shah. Analysis of bitcoin network dataset for fraud. *Unpublished Report*, 2013.

[15] Bitcoin.com. Bitcoin history part 11: The first major loss of coins. Available at `https://news.bitcoin.com/bitcoin-history-part-11-the-first-major-loss-of-coins/`.

[16] Techcrunch. Binance says more than $40 million in bitcoin stolen in 'large scale' hack. Available at `https://techcrunch.com/2019/05/07/binance-breach/`.

[17] John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

[18] Kentaroh Toyoda, Tomoaki Ohtsuki, and P Takis Mathiopoulos. Identification of high yielding investment programs in bitcoin via transactions pattern analysis. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pages 1–6. IEEE, 2017.

[19] Kota Kanemura, Kentaroh Toyoda, and Tomoaki Ohtsuki. Identification of darknet markets' bitcoin addresses by voting per-address classification results. In *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pages 154–158. IEEE, 2019.

[20] Wikipedia. Darknet market. Available at `https://en.m.wikipedia.org/wiki/Darknet_market`.

[21] Kyungchan Ko Chaehyeon Lee, Sajan Maharjan and James Won-Ki Hong. Toward detecting illegal transactions on bitcoin using machine-learning methods. appear.

[22] Chainalysis. Chainalysis reactor. Available at `https://www.chainalysis.com/chainalysis-reactor/`.

[23] WalletExplorer. Walletexplorer: smart bitcoin block explorer. Available at `https://www.walletexplorer.com/`.

[24] BeautifulSoup. Beautiful soup documentation. Technical report. Available at https://www.crummy.com/software/BeautifulSoup/bs4/doc/.

[25] Bitcoin.org. Bitcoin core. Available at https://bitcoin.org/en/bitcoin-core/.

[26] Bitcoin.org. Bitcoin core json apis. Available at https://bitcoin.org/en/developer-reference#bitcoin-core-apis.

[27] scikit learn.org. scikit-learn: Machine learning in python. Available at https://scikit-learn.org/.

[28] tensorflow.org. tensorflow: An end-to-end open source machine learning platform. Available at https://www.tensorflow.org/?hl=en.

[29] Miron B Kursa, Witold R Rudnicki, et al. Feature selection with the boruta package. *J Stat Softw*, 36(11):1–13, 2010.

[30] Jesse Davis and Mark Goadrich. The relationship between precision-recall and roc curves. In *Proceedings of the 23rd international conference on Machine learning*, pages 233–240. ACM, 2006.

[31] David Martin Powers. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. 2011.

[32] Cyril Goutte and Eric Gaussier. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *European Conference on Information Retrieval*, pages 345–359. Springer, 2005.

Acknowledgements

먼저 석사 학위를 성공적으로 이수할 수 있도록 언제나 저를 지지해주시고, 응원해주신 사랑하는 부모님과 하나뿐인 동생 주은 그리고 유정에게 진심으로 감사드립니다. 때로 흔들리고, 지칠 때도 있었지만 조건 없이 보내주신 사랑과 무한한 응원에 힘을 낼 수 있었습니다. 항상 감사하고 사랑합니다. 또한 좋은 연구 환경에서 부족함 없이 공부할 수 있고, 올바른 지도로 무탈한 석사 생활을 이어나갈 수 있게 애써주신 홍원기 교수님께 진심으로 감사의 말씀을 전합니다. 교수님께서 제공해주신 여러 좋은 기회들과 진심 어린 지도 덕분에 많이 배우고, 많이 성장할 수 있는 시간들이었습니다. 그리고 저의 은인인 지은 언니와 먼 곳에서 항상 힘이 되어준 숭실대학교 13학번 동기를 비롯한 소중한 친구들에게 감사를 전합니다. 내색하지 않았지만 의지할 곳이 있어 든든했고, 지난 시간들을 견딜 수 있게 해주어 고맙습니다.

석사 과정 동안 과분한 도움을 주셨던 경찬 선배, 디펜스 과정에서 많은 도움을 주셨던 지범 선배, 연구실을 위해 묵묵히 고생하셨던 랩장 도영, 세영 선배, 그리고 연구실 생활을 하면서 좋은 추억을 함께 만들었던 연구실 선후배님들, 동기 희곤, 가연, Sajan에게도 감사의 말씀을 전합니다. 함께 고생했던 타 학교 대학원 학생들을 포함하여 모두가 의미 있는 성과를 얻고, 항상 행복한 일만 있기를 바랍니다.

포항공과대학교에 입학하여 낯선 환경에 적응하는 것이 힘들기도 했고, 순간순간 목표를 잃을 때도 있었지만 지난 2년의 시간은 제 스스로를 학업 또는 생활 측면에서 성장시켜 준 뜻깊은 시간이었습니다. 여전히 부족하지만 앞으로 더 나아갈 수 있는 용기와 끈기를 갖게 해준 시간들이었기에 값지고 의미 있는 2년이었다고 생각됩니다. 앞으로도 스스로 한 선택에 책임을 질 수 있고, 제가 목표하는 바를 이룰 수

있도록 끊임없이 노력하는 사람이 되겠습니다.

Curriculum Vitae

Name : Chaehyeon Lee

Research Interest

Blockchain Transaction Monitoring; Blockchain Network Analysis;

Education

2013 – 2018	Department of Information Communication and Electronic Engineering, Soongsil University (B.S.)
2018 – 2020	Department of Computer Science and Engineering, Pohang University of Science and Technology (M.S.)

Research/Project Experience

2018. 4. – 2020. 2. 블록체인의 트랜잭션 모니터링 및 분석 기술개발 (Funded by IITP)

Publications: International Conference

1. **Chaehyeon Lee**, Sajan Maharjan, Kyungchan Ko, James Won-Ki Hong "Toward Detecting Illegal Transactions on Bitcoin using Machine-Learning Methods", 2019 International Conference on Blockchain and Trustworthy Systems (BlockSys'2019), Guangzhou, China, Dec. 7-8, 2019.
2. Kyungchan Ko, Taeyeol Jeong, **Chaehyeon Lee**, Sajan Maharjan, and James Won-Ki Hong "Prediction of Bitcoin Transactions Included in the Next Block", 2019 International Conference on Blockchain and Trustworthy Systems (BlockSys'2019), Guangzhou, China, Dec. 7-8, 2019.
3. **Chaehyeon Lee**, Heegon Kim, Sajan Maharjan, Kyungchan Ko, James Won-Ki Hong, "Blockchain Explorer based on RPC-based Monitoring System", 1st IEEE International Conference on Blockchain and Cryptocurrency (ICBC 2019), Seoul, Korea, May. 14-17, 2019.
4. Kyungchan Ko, **Chaehyeon Lee**, Taeyeol Jeong, James Won-Ki Hong, "Design of RPC-based Blockchain Monitoring Agent", 9th International Confer-

ence on Information and Communication Technology Convergence (ITRC 2018), Jeju, Korea, Oct. 17-19, 2018.

Publications: Domestic Conference

1. **이채현**, Sajan Maharjan, 고경찬, 홍원기, "비트코인 네트워크의 불법거래 탐지 연구", KNOM Conference 2019, Daegu, Korea, May 30-31, 2019, pp. 99-101.
2. 고경찬, **이채현**, 홍원기, "비트코인 노드 메모리 풀 유사도 분석", KNOM Conference 2019, Daegu, Korea, May 30-31, 2019, pp. 16-18.
3. 고경찬, **이채현**, 홍원기, "이더리움 컨트랙트 모니터링 및 분석시스템", KNOM Workshop 2018, Seoul, Korea, Nov 30, 2018, pp. 4-5.