明治大学総合数理学部

2016 年度

卒　業　研　究

Bitcoin 国別ノードランキング

学位請求者　先端メディアサイエンス学科

永田倖大

目次

1 はじめに

1.1 研究背景

Bitcoin は 2009 年に運用が開始され，取引手数料が安いことや匿名性の高さ，第三者機関を介さずに取引できるという特徴がある．さらに，Bitcoin にはブロックチェーンという技術が使われており，その技術を用いたサービスを IBM が他者に対して提供を開始した[1]．

Bitcoin の取引は匿名であると言われているが,2014 年に S.Meiklejohn らにより特定のウォレットを管理しているユーザを判別することができたという研究結果[2]が示され,その匿名性が疑問視されている．

さらに，分散管理されているため誰と誰が通信しているか，どの国で多く使われているかなどが不明であった．

1.2 研究目的

本研究では Full Node のウォレットが Bitcoin ネットワークへコネクションを行う時の様子を長期間観測し，Bitcoin ノードの IP を収集した．収集したデータを解析するシステムを開発し，どの国で多く使われているか，どれくらいの IP が集まるのかを明らかにすることを目的とする．

1.3 論文構成

本論文の構成は次の通りである．第２章では，Bitcoin ネットワークへの接続，Bitcoin に関するパケットの解説,Bitcoin ノード IP の収集および分析について述べる．第 3 章では,Bitcoin 取引の可視化について述べる．

1.3 Bitcoin とは

Bitcoin は Nakamoto の論文[3]に基づいたディジタル通貨である．国を超えて容易に送金できること，取引手数料が安い，匿名性が高いという利点を持つ．Bitcoin には 2100 万 BTC が採掘上限として決められている．この上限には 2140 年ごろに到達する見込みである．ユーザはビットコインを管理するためにウォレットと呼ばれるものを使用する．ウォレットには Multibit や Bitcoin Core や，ペーパーウォレットというものがある

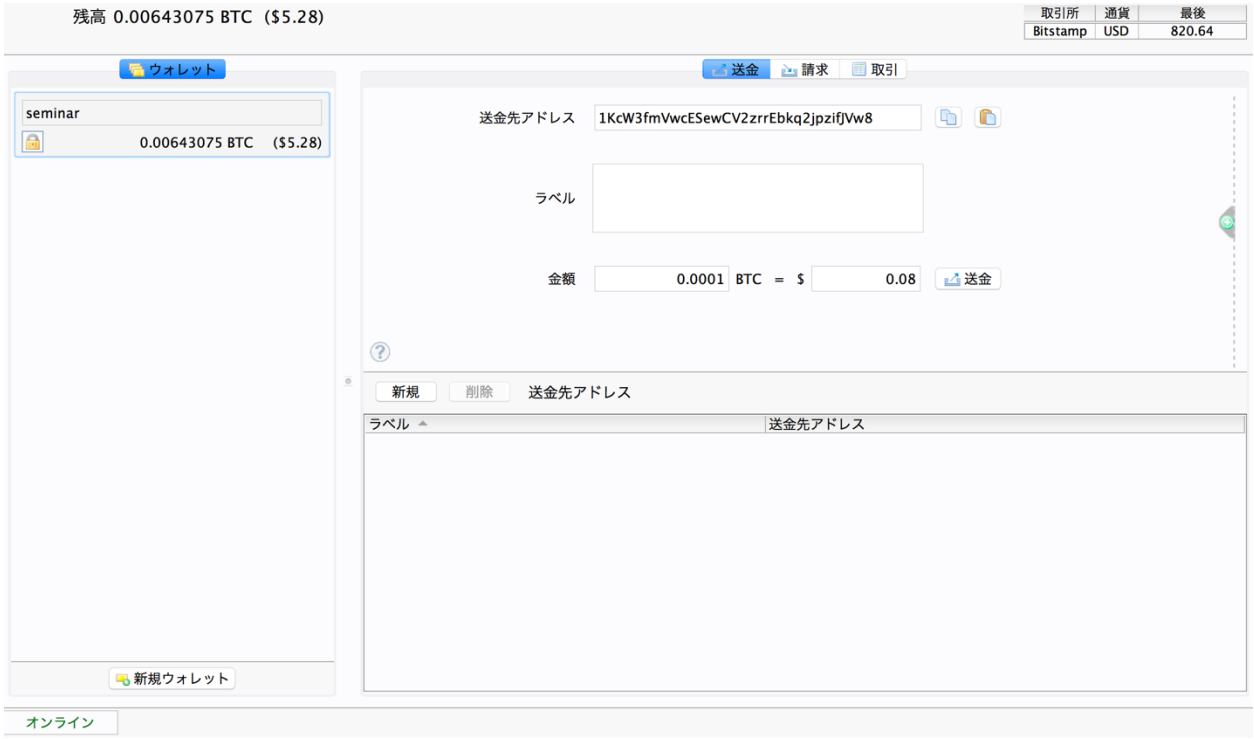

図 1:Multibit 画面

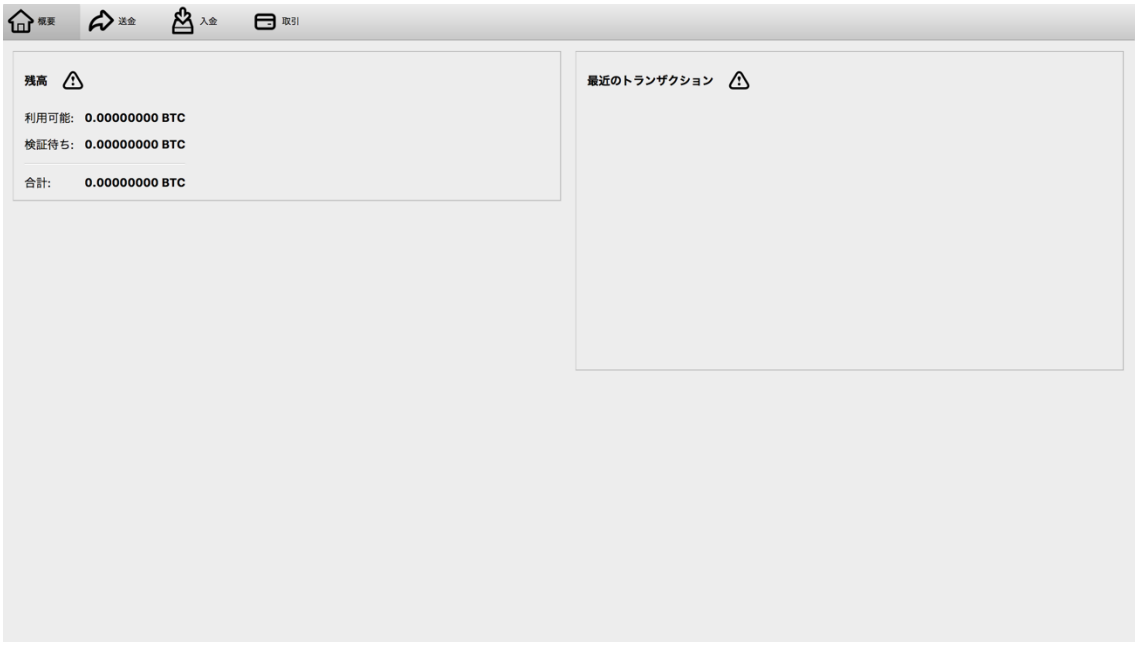

図 2:Bitcoi Core 画面

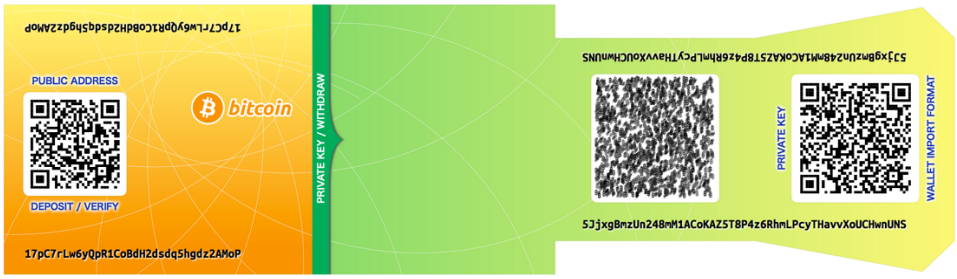

図 3:ペーパーウォレット

取引は各ノードが P2P 方式で行い，取引完了には時間を要する．取引を行うためにはアドレスが必要になる．取引方法は，自分の管理しているアドレスを入力とし，宛先に相手のアドレスを設定することで取引を行うことが可能である．図 4 であるように，ユーザ A がユーザ B にビットコインを送金したいときは,アドレス A とアドレス B を入力として送金先としてアドレス D を指定することで Bitcoin の送金を行うことができる．

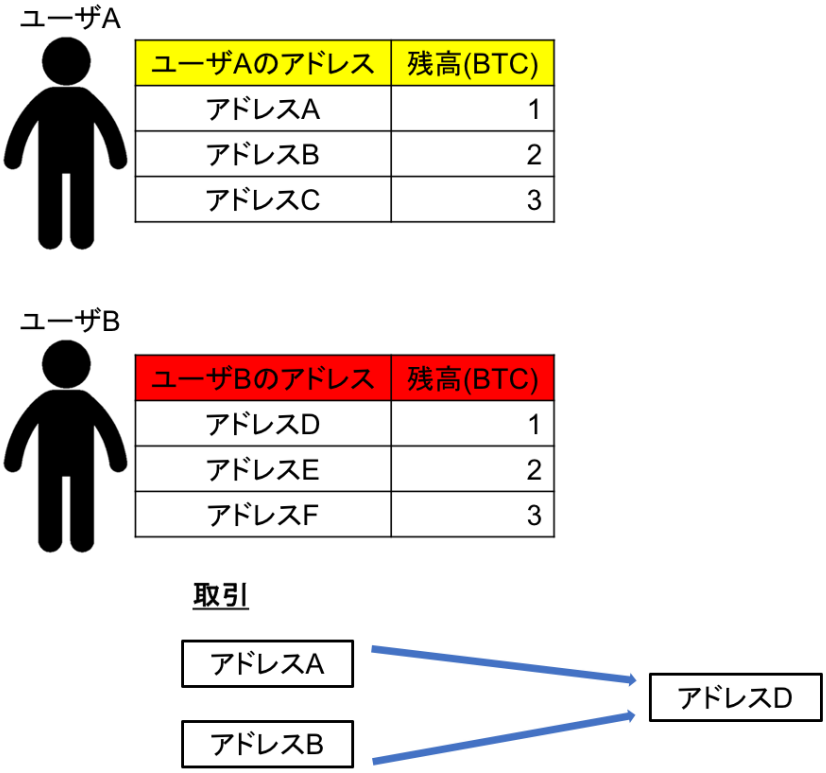

図 4:取引方法

取引はブロックの中に格納される．ブロックは約 10 分に 1 個生成されており，ブロックが生成されることで取引は承認を得る．ブロックを生成するためには膨大な計算量が必要となる．

要約	
取引件数	95
合計出力	2,803.00335174 BTC
推定取引量	129.01099381 BTC
取引手数料	0.0107483 BTC
ブロック高	353204 (主鎖)
タイムスタンプ	2015-04-22 09:01:34
受け取り時刻	2015-04-22 09:01:34
中継所	BTCChina Pool
難易度	47,610,564,513.47
ビット	404166640
サイズ	49.3427734375 KB
バージョン	3
ノンス	1454422704
ブロック報酬	25 BTC

ハッシュ	
ハッシュ	00000000000000000004178fead62073c38508ab1eabcea34bcef0d8ec027cf868
前のブロック	00000000000000000005b5be5ad2edfc8483c936b598e6ae2044f2ade4081934d2
次のブロック（複数可）	00000000000000000242bdc8f6abad659e13ea4f5165dae0aa65b6c460332428
マークルルート	e7bad14d9242353ee7ba80c0de5bf9f005a8678dd0be64570dff2aebb0078070

伝搬ネットワーク (クリックして表示する)

図 5:ブロック情報

ブロックはブロックチェーンを形成する．ブロックチェーンは，ビットコインの二重取引や改ざんを防止する役割がある．

ブロック生成を行うのはユーザであり，最初にブロック生成したユーザには報酬としてビットコンを手に入れることができる．ブロックを生成する作業は Proof of work と呼ばれている．

1.4 ウォレット使用例

Multibit の使用例を紹介する.

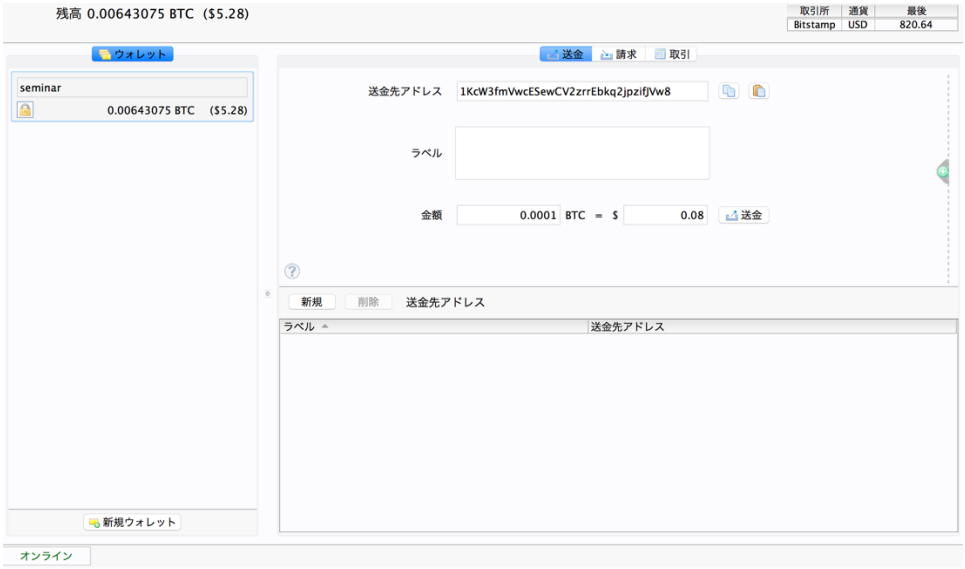

図 6:Multibit 画面 1

図 6 は Multibit の画面である. ウォレットはアドレスを管理するために必要になる. 新規ウォレットから作成することが可能である. ビットコインを送金するには, 送金先アドレスににアドレスを入力し, いくら送るかを金額に設定する.

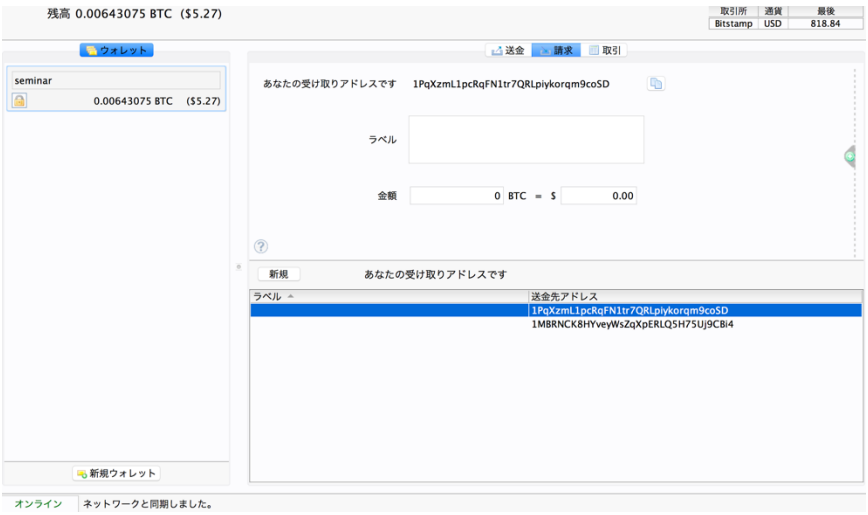

図 7:Multibit 画面 2

自分にビットコインを送金してもらうときは請求ページにあるアドレスを相手に教える必要がある. 新規ボタンを押すことでアドレスを作成することができる.

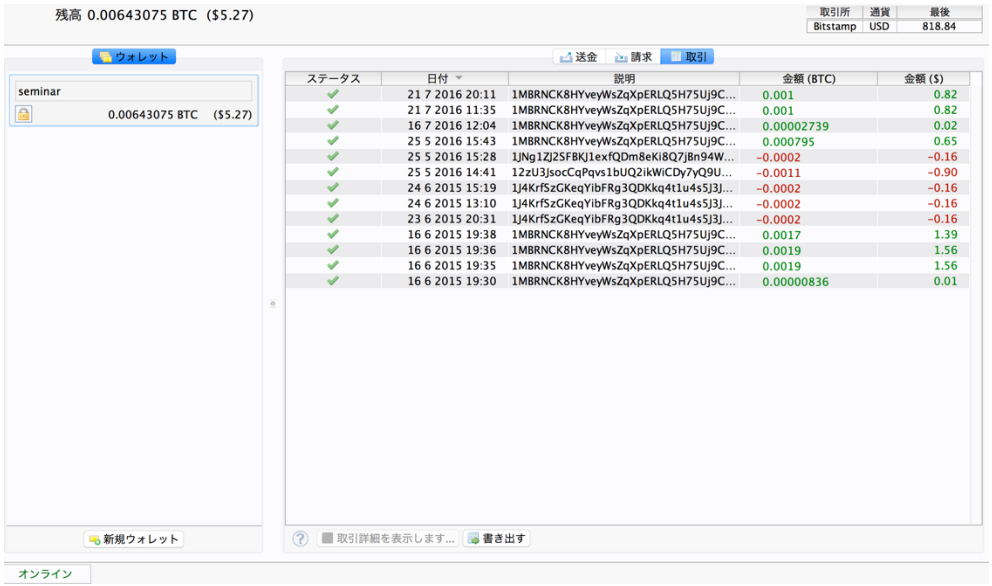

図 8:Multibit 画面 3

取引ページでは，過去に行った取引の一覧を確認することができる．

Bitcoin Core の使い方を紹介する．

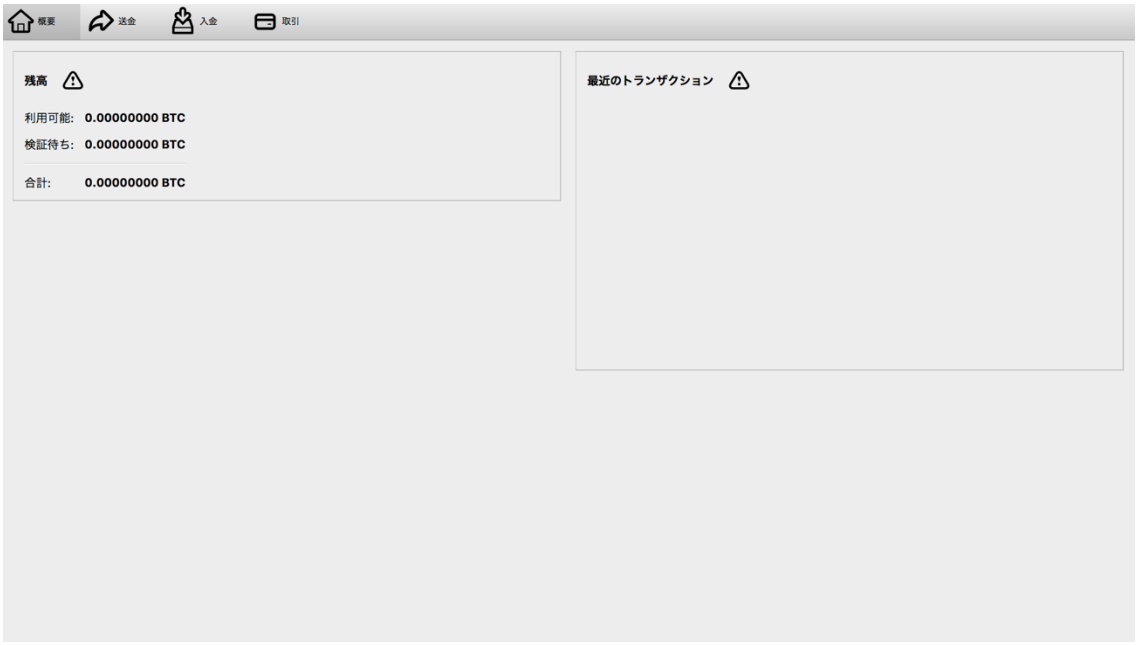

図 9:Bitcoin Core 画面 1

図 9 は Bitcoin Core の画面である．Bitoin Core を起動するとブロック情報のダウンロードを始める．現在ブロック情報は，448308 ブロックで容量は 112.4GB になる．

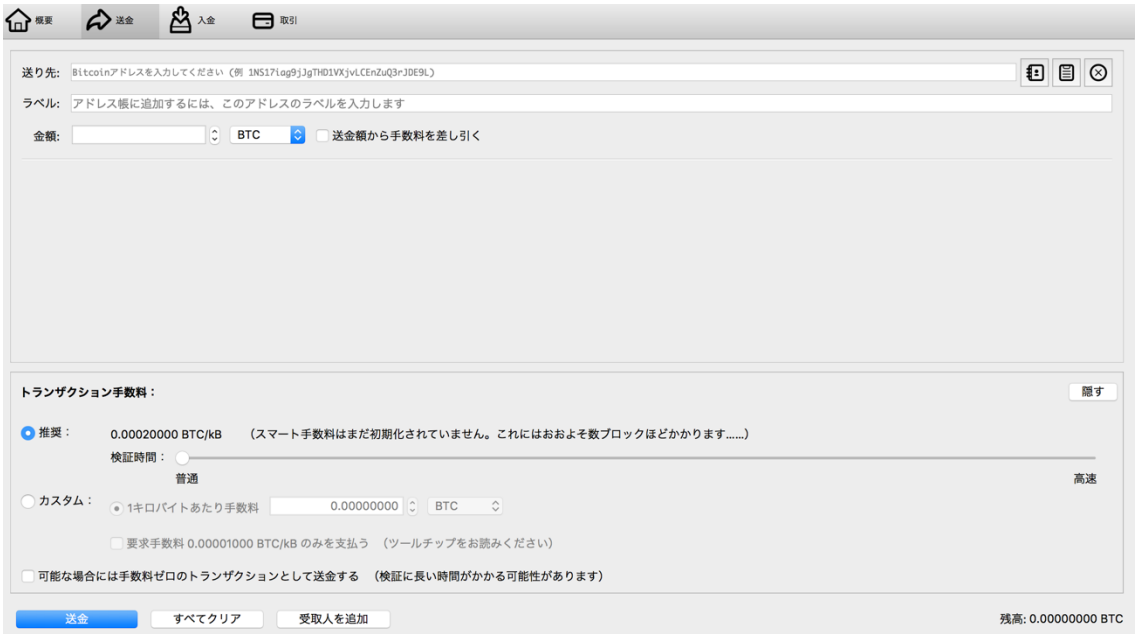

図 10:Bitcoin Core 画面 2

ビットコインを送金するには，送り先にアドレスを指定する．Bitcoin Core では取引手数料を指定することができる．

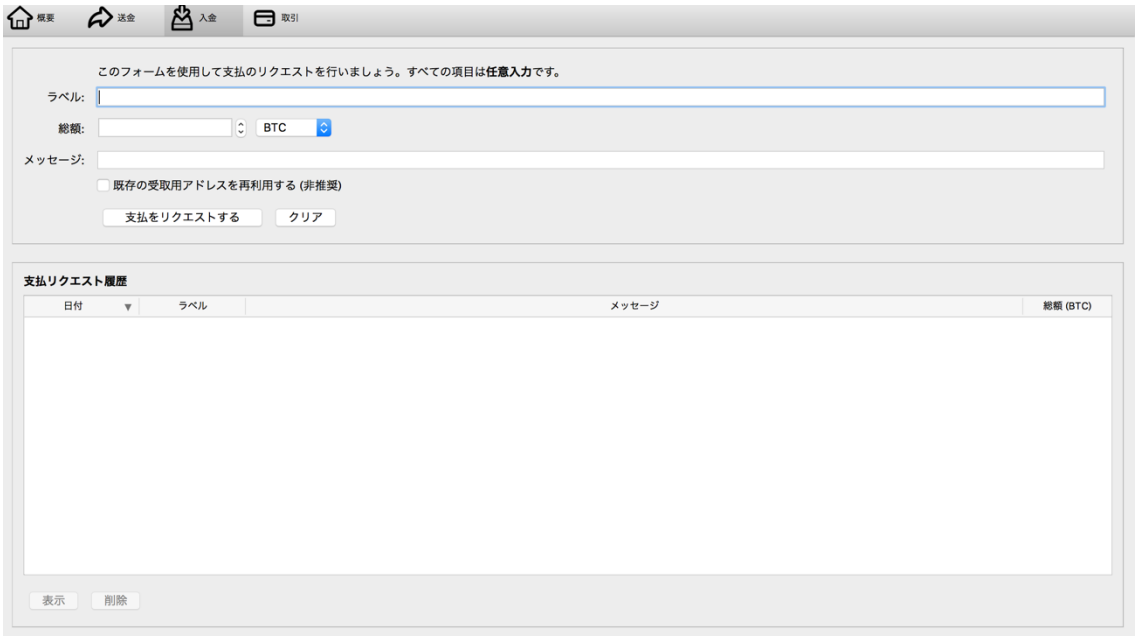

図 11:Bitcoin Core 画面 3

図 12:Bitcoin Core 画面 4

図 11 の画面で，支払いをリクエストするをクリックすることで図 12 の様な画面が出てくる．この画面に書いてあるアドレスを相手に教えることで，ビットコインを受け取ることができる．

1.5 先行研究紹介

S.Meiklejohn らの先行研究[2]は，ビットコインの取引の際に使用される，アドレスをクラスタリングし再識別攻撃を行うことで管理者を明らかにすること，Bitcoin 市場の長期的な変化の分析，その変化による Bitcoin システムへの影響，安西や詐欺目的で使用されたビットコインの検知を目的としている．

アドレスの管理者を明らかにするために，アドレスにタグ付けを行う手法を用いている．

図 13:クラスタリング手法 1

図 13 はアドレスのクラスタリングの手法の一つとして紹介されている．ビットコインの取引では入力アドレスの管理者は同一であるため，図 13 のように入力アドレスが 2 つ以上あるような取引 1,取引 2 があった場合，アドレス A,B,C の管理者は同一であると定めている．

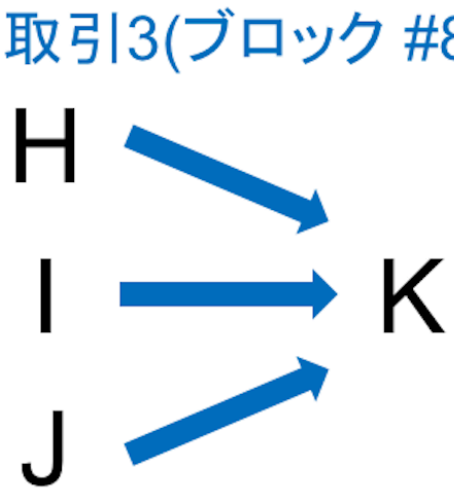

図 14:クラスタリング手法 2

アドレス間の変化を見るために，Change Address というクラスタ手法 2 も用いている．

Change Address とは複数のアドレスで管理しているビットコインを 1 つのアドレス送金する方

法である．

研究結果として，3,384,179 個のクラスターができ，その中で管理者が分かったものは 2197 個であり，1,800,000 個のアドレスの管理者を特定したといっている．

2 実験

2.1 Bitcoin ネットワークへのコネクション

本節では，Bitcoin ウォレット，Bitcoin ネットワークに接続するプロトコル，Bitcoin に関するパケットの説明を行う．

2.1.1 Multibit と Bitcoin Core の違い

Bitcoin ノードには様々な種類があり[4]，そのなかに Full Node と SPV(Simplified Payment Verification)がある.そのクライアントであるウォレットには，表 1 で示す種類がある．

表 1.MultiBit と BitcoinCore の比較

	MultiBit	Bitcoin Core
種類	SPV	Full Node
ブロック情報	ヘッダーのみ	全て所持

Bitcoin Core と Multibit はビットコインの取引を行うことが可能である．しかしノードによって取引を行う機能を持たないクライアントを稼働しているものもある．

2.2.2 コネクションの様子

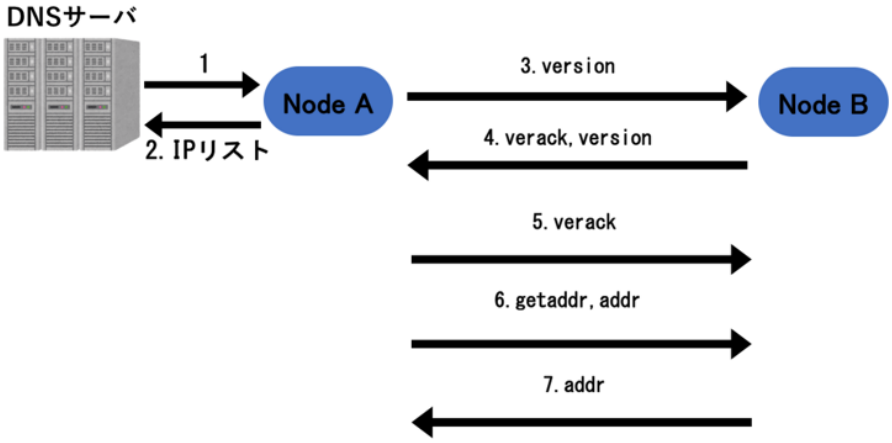

図 15:コネクション方法

図 15 に Bitcoin ネットワークに接続するプロトコルを示す．1.DNS サーバに問いかけ Bitcoin ノードを稼働しているノードの IP アドレスリストを受け取る．そして，ノード A はリストの中から任意の IP を選び，3.version パケットを送信する．3.version パケットには A のクライアント情報や，どのブロック情報まで保持しているかなどが含まれている．3.version パケットを受けとるとノード B は 4.verack パケットを A に返信する．その後，ノード B に対して 6.addr パケットと 6.getaddr パケットを送信する．6.addr パケットには，A のノード情報や，他の Bitcoin ネットワークのノード情報が含まれている．6.getaddr パケットを B に送ることによって，7.addr パケットを返答として受け取る．7.addr パケットには，1 つだけノード情報が入っているものや，1000 個入っているものまで存在する．

2.2.3 各パケットの説明

図 16 に version パケットを内容を示す.

```
Bitcoin protocol
    Packet magic: 0xf9beb4d9
    Command name: version
    Payload Length: 102
    Payload checksum: 0x40538f0e
    Version message
        Protocol version: 70012
        Node services: 0x0000000000000005
        Node timestamp: Dec 12, 2016 18:09:28.000000000 JST
        Address as receiving node
        Address of emmitting node
        Random nonce: 0x3bbd12a066ee8ce8
        User agent
            Count: 16
            String value: /Satoshi:0.12.0/
        Block start height: 443099
```

図 16:version パケット

Protocol version は送信ノードが使用しているプロトコルバージョンを示しており, String value では使用しているソフトウェアのバージョン, Block start height には, 現在どのブロック情報まで所持しているかが含まれている.

図 17 に addr パケットの内容を示す.

```
Bitcoin protocol
    Packet magic: 0xf9beb4d9
    Command name: addr
    Payload Length: 31
    Payload checksum: 0x3245bb2d
    Address message
        Count: 1
        Address:
5a694e5805000000000000000000000000000000000000ffff...
            Node services: 0x0000000000000005
            Node address: ::ffff:209.188.18.142
(::ffff:209.188.18.142)
            Node port: 8333
            Address timestamp: Dec 12, 2016 18:09:46.000000000
JST
```

図 17:addr パケット

Count は addr パケットに含まれる IP アドレスの数を示し, Node address には Bitcoin ノードを稼働している IP アドレスが記されている.

表 2.addr パケット統計情報

アドレスの個数	パケットの数
1	351
2	25
3	1
524	1
1000	9

表 2 に 387 個の addr パケットの統計情報を示す.

```
Bitcoin protocol
    Packet magic: 0xf9beb4d9
    Command name: tx
    Payload Length: 191
    Payload checksum: 0xae049e31
    Tx message
        Transaction version: 1
        Input Count: 1
        Transaction input
        Output Count: 1
        Transaction output
            Value: 2435310
            Script Length: 25
            Script:
76a91469feb683ce891c508786fdb1d2ef29eb5304d34d88...
        Block lock time or block ID: 0
```

図 18:tx パケット

図 18 にビットコインの送受信の取引を表す tx パケットを示す. Value は送金したビットコインの額であり, 10^{-8} の額を表記している.

2.2 IP 収集および分析

本節では，Bitcoin ネットワークに接続し長期間観測することによって IP 収集を行い，それらのデータの分析を行う．

2.2.1 実験目的

Bitcoin ノードの IP 収集を行い，どの国で多く使われているか，どれくらいの数の IP が集まるのか，さらに Bitcoin ネットワークの接続プロトコル，Bitcoin に関するパケットの分析を目的とする．

2.2.2 実験方法

Bitcoin の Full Node Client である Bitcoin Core を起動し，その時に通信をおこなった Bitcoin に関するパケットを観測するために解析システム BitconPcap を開発した．BitcoinPcap を用いて，パケット分析を行い，様々なデータ収集を行う．

2.2.3 IP 収集方法

IP 収集するためには 3 つの手法がある．1 つ目は，Bitcoin ネットワークに接続し，Bitcoin に関するパケットから IP を抜き出す手法．2 つ目は，addr パケットに含まれる IP アドレスを抜き出す手法．3 つ目は，DNS サーバに問いかけ，IP リストを取得する手法である．今回は，1 つ目の手法で IP 収集を行う．

2.2.4 作成プログラム BitcoinPcap 説明

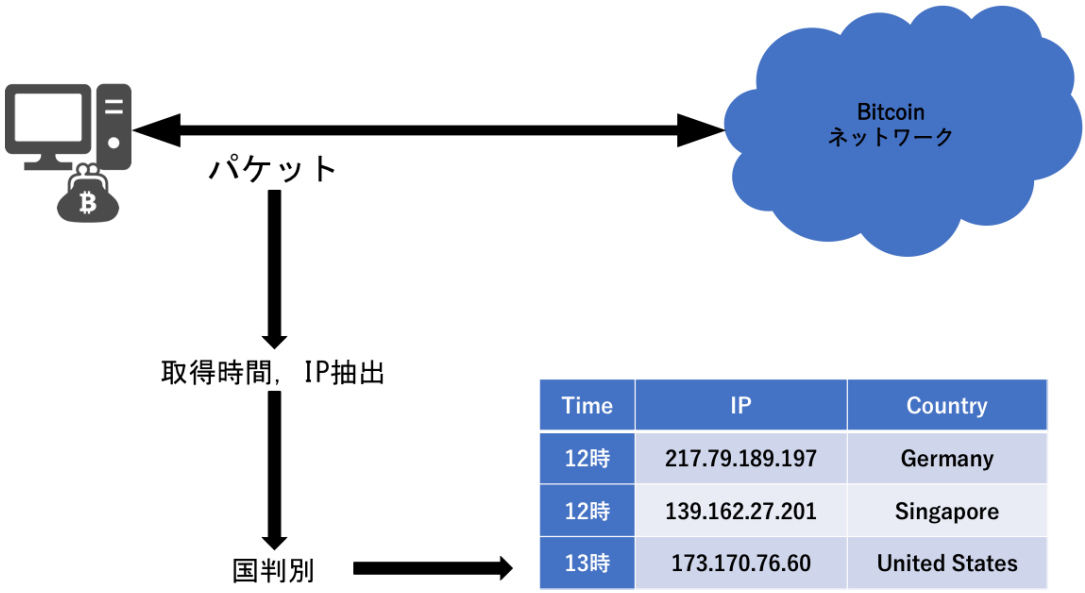

図 19:BitcoinPcap 概要

図 19 に BicoinPcap の説明を行う．はじめにウォレットである Bitcoin Core を起動し，Bitcoin ネットワークに接続する．次に他の Bitcoin ノードと通信を行う．その通信パケットから IP データを抽出し，国判別を行う．そして，それらのデータを表 3 のような csv ファイルとして出力する．BitcoinPcap は特定のパケットだけを抽出することも可能である．例えば，図 17 の様な addr パケットだけを集め，addr パケットの Count を抜け出すことも可能である．

2.2.5 実験環境

実験環境を表 3 に示す．

表 3. 実験環境

OS	macOS Sierra 10.12.1
メモリ	8GB
言語	Python 2.7.12

BitcoinPcap には python のライブラリである pyshark を使用した．pyshark はパケットキャプチャーソフトの thsark を python の環境で利用できるようにしたものである．プログラムの起動はターミナルで行う．

2.2.6 収集データ解説

表 4 に収集情報を示す.

表 4.収集情報

観測期間	2016 年 11 月 17 日～11 月 27 日(10 日間)
起動ウォレット	Bitcoin Core
ユニーク IP 数	1715 個
国数	69 国
観測場所	自宅(東京都国分寺市)
ISP	アルテリア・ネットワークス
bandwidth	最大 1Gbps
IP 上位オクテット	122.219.218.0/24

収集した IP の総数は 2524 個であり，重複を除くと 1715 個である．図はユニーク IP 数の分布図である.

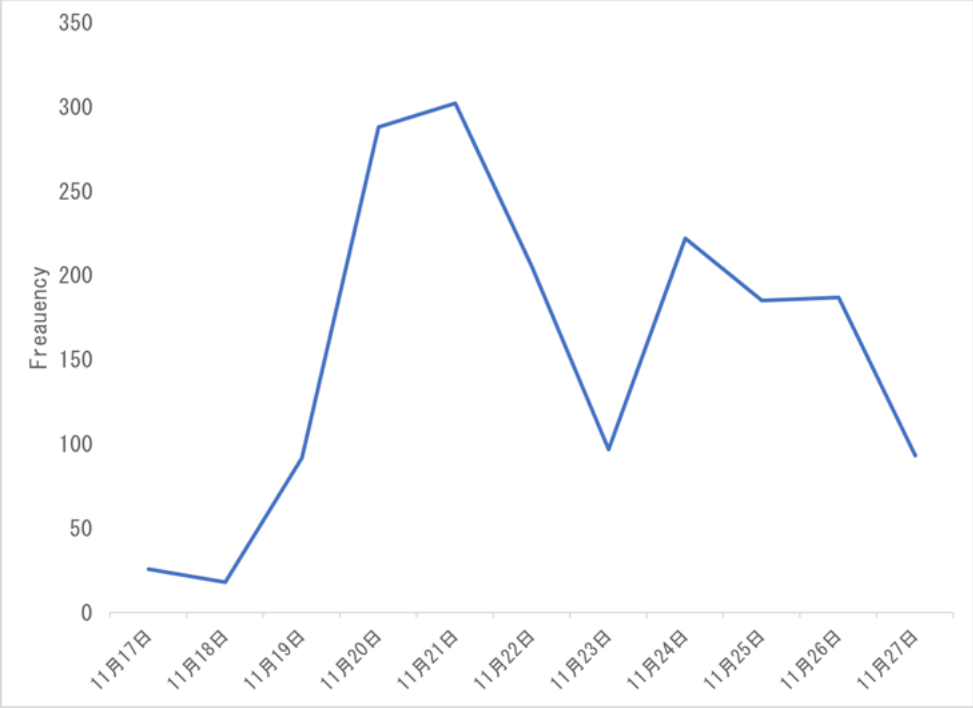

図 20:IP アドレス観測分布

図 20 にユニーク IP の分布図を示す．11 月 17 日から 20 日の観測期間は，日に 7 時間ほど断続的に観測していたため収集 IP が少ない．11 月 20 日から収集数が増えたのは 24 時間連続して観測したためである．

表 5.収集データ一部

Time	IP	Country	Day
12 時	217.79.189.197	Germany	11/17
12 時	139.162.27.201	Singapore	11/17
13 時	173.170.76.60	United States	11/17
14 時	66.96.199.201	Singapore	11/17
15 時	50.206.138.178	United States	11/17
16 時	208.66.68.127	Canada	11/17
12 時	64.121.102.206	United States	11/18
12 時	46.101.192.63	Germany	11/18

表 5 に収集した IP の一部を示す．Country の判別には GeoIP を用いた．

2.2.7 国別 IP 分析

表 6 に観測した IP 上位 10 カ国を示す.

表 6. 上位 10 カ国 IP

順位	国	割合[%]	個数
1	United States	33	447
2	Germany	13	240
3	United Kingdom	6	115
4	France	5	93
5	Netherlands	5	86
6	Russia	4	81
7	Canada	4	74
8	China	3	53
9	Sweden	2	42
10	Ukraine	2	37

表 6 に含まれる IP は，自分のノードと通信を行ったものだけであり，Full Node と SPV のどちらも含まれる．1 つだけ観測できた国は，モナコやヨルダンがあった.

2.3 考察

実験結果より，Bitcoin ネットワークへのコネクション方法，Bitcoin に関するパケットの内容，Bitcoin ノードがどの国で多く使われているかが明らかになった．

本研究で調査を行なったユニーク IP の 33 パーセントはアメリカのものであったので，世界で一番 Bitcoin ノードを稼働している国はアメリカではないかと考える．

IP アドレス 2524 個中 809 個が重複していた．収集 IP の 32 パーセントと高いことから，ウォレットを起動し，通信を確立する相手ノードは常に稼働し続けていると考えられる．

2.4 おわりに

Bitcoin に関するパケットを収集し，情報を取得するプログラム BitcoinPcap を開発した．BitcoinPcap は取得したいパケットの種類や，パケットの一部を抽出することが可能であるため Bitcoin に関するパケットを観測，および解析するのに有益である．

本実験で収集した IP アドレスは，自分のノードと通信を行ったものだけであったが，DNS サーバーへ通信を行い IP リストを受け取る方法や，他のノードに向けて getaddr パケットを送り集める手法など，より効率よく多くの IP を集めることを検討中である．

本実験では tx パケットの解析は行わなかった．tx パケットに含まれる取引額と国の関係や，どれくらいの tx パケットが収集できるか，どの時間帯に取引がよく行われているかを分析することで，ビットコインの使用の特徴が明らかになるだろう．

3 Bitcoin 取引の可視化

本節では Bitcoin 取引の可視化についての説明を行う.

3.1 データ収集方法

世界中の Bitcoin 取引の情報や，ブロック情報などを記録しているサイト Bitcoin ブロックエクスプローラー[5]を用いて，取引データを収集する．データ収集は手作業で行い，自分でデータ量は決定する.

本実験では 8 行の取引を収集した．1 の取引データは 2016 年 1 月に，Bitcoin ブロックエクスプローラーサイトで公開されている実在する取引データを収集した．表 7 の取引データ以降にも取引は続いているが割愛している.

図 21 に Bitcoin ブロックエクスプローラーで公開されている取引例を示す.

図 21:取引データ例

ここで 201417 が取引が格納されているブロック，5c76eb4dfb0941856a22983ef05b2f5c669dadc98ea34ea11974cacba9dc7 が取引 ID，1MdYC22Gmjp2ejVPCxyYjFyWbQCYTGhGq8，1E86A5E6ANEVPuay2XLGVsXjaxT5MbRm，19PphSFxzmsSZ3JRacQArEgN1b67ar83 がアドレスである．合計インプットは取引で使用されたビットコインの総額である.

1MdYC22Gmjp2ejVPCxyYjFyWbQCYTGhGq8 から 1E86A5E6ANEVPuay2XLGVsXjaxT5MbRm へ 50.53036298BTC，19PphSFxzmsSZ3JRacQArEgN1b67ar83 へ 0.10481202BTC を送金している.

3.2 取引データ形式

取引データには，ID，親番号，深さ，ブロック番号，が含まれる．ID はユニークであり，ブロック番号は取引が含まれているブロックの番号で，深さは図のツリーのどの深さにいるのか，親番号はノードの親の ID を表している．親番号が 0 の時は，親ノードは存在しない．表 7 にデータ例を示す．

表 7. 取引データ例

ID	親番号	深さ	ブロック番号	送金額
1	0	1	200000	50.635175
2	1	2	201417	50.635175
3	2	3	201419	50.53036298
4	2	3	201742	0.10481202
5	3	4	201548	50.1615802
6	3	4	201420	0.36878278
7	4	4	207133	0.00009994
8	4	4	201755	1.4777

3.3 実験環境

実験環境を表 8 に示す．

表 8. 実験環境

OS	OS X El Capitan 10.11.2
メモリ	8GB
言語	Processing 3.0.1

3.4 作成プログラム BitcoinTxVisualize 説明

収集したデータを可視化するプログラム BitcoinTxVisualize を用いて，表 7 の csv ファイルを読み込み，可視化した結果を図 22 に示す.

Algorithm ： BitcoinTxVisualize	
	Csv ファイル読み込み ノードに情報を格納 ノードの表示位置，枝の表示位置を決定 ノード，枝を描画

3.5 実験結果

図 22 に表のデータを使ったプログラムの実行結果を示す.

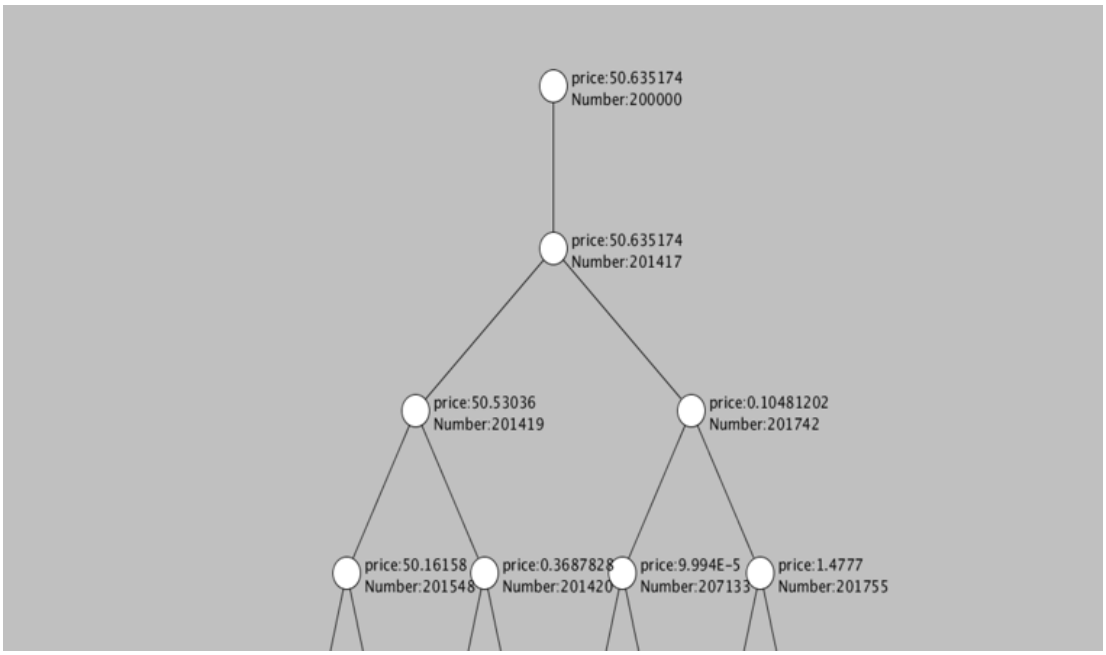

図 22:実行例

ノードの右に表示されている price は受け取った Bitcoin の総額である. Number はその Bitcoin を他のアドレスに向けて送金した取引が含まれているブロック番号であり, そしてノードから出ている枝の本数が、取引先件数である. 線が出ていないノードは, それより先で取引が行われておらず, 保有している状態になっている.

例えば図 22 の深さ 2 のノードは, 50.635174BTC を受け取りその Bitcoin を 2 つのアドレスに向けて, 50.53036,0.1481202 ずつ送金している. そしてこの取引が格納されているブロック番号が 201417 である.

図 22 では深さが 4 までしか表示されていないが割愛している.

3.6 改善点

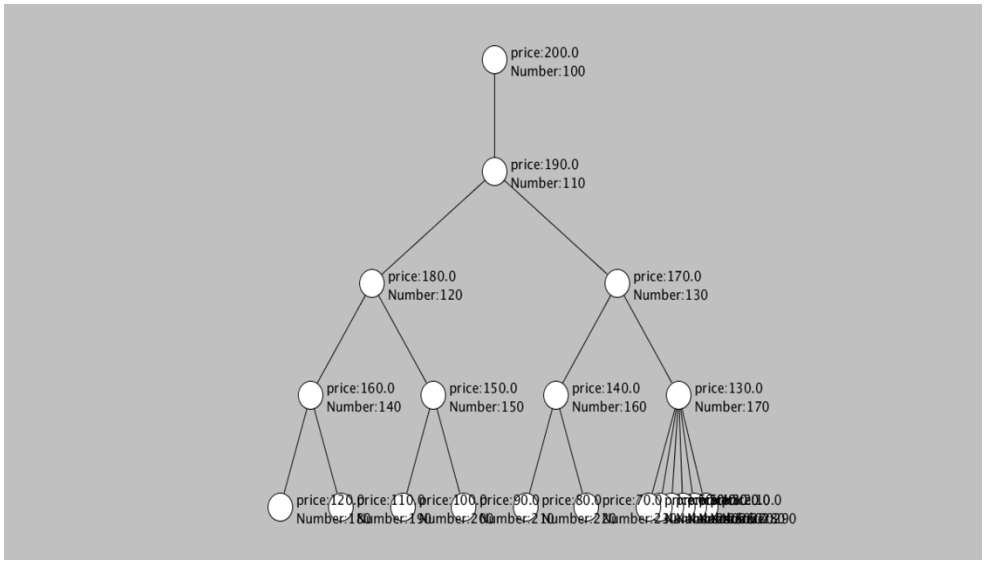

図 23:改善が必要な例

図 23 の様に，深くなるにつれて，ノード 1 つを表示できるスペースが狭くなってしまう．右下のノードのブロック番号が重なってしまい見づらい．さらに，1 つのノードに対して多くの子ノードが繋がっている時もブロック番号が重なってしまう．

本実験では，データ収集は手動で行っているため大量のデータを集めるのは困難である．取引データは図 21 のページに 1 つずつしか公開されておらず，例えば 1000 件の取引データを収集するためには手動で約 1000 回クリックしなければならないからである．

3.7 考察

Bitcoin の取引情報に基づいて，Bitcoin のブロックの繋がりを可視化する BitcoinTxVisualize を開発した．少ないデータ量では可視化した図は見やすく，繋がりを理解することは容易であるが，データ量が大きい時は，可視化した図はノードが重なってしまい図 23 の様に見づらくなってしまう．他にもデータ量が増えると price と Number の表示が重なってしまい確認しづらくなる．さらにデータ収集は手動で行っているため，多くのデータを集めるのは困難である．

なので，多くのデータにも対応し，よりブロックの繋がりが理解しやすいようにプログラムを改善し，データ収集を自動化することを今後の課題とする．

謝辞

本研究に際して，様々なご指導をいただきました菊池浩明教授に深く感謝します．最後に，本研究に対し熱く議論を交わした菊池研究室の皆様に感謝の意を表すると共に，謝辞にかえさせていただきます．

参考文献

[1]	IBM Blockchain （ https://www.ibm.com/blockchain/ ， 2016 年 4 月参照）
[2]	S.Meiklejohn, M.Pomarole,G.Jordan,K.Levchenko, D.McCoy, G.M.Voelker, “A Fistful of Bitcoins:Characterizing Payments Among Men with No Names”,(Internet measurement conference),pp127-140,IMC’13,2013.
[3]	S.Nakamoto , “Bitcoin: A Peer-to-Peer Electronic Cash System” （ https://bitcoin.org/bitcoin.pdf ,2016 年 4 月参照 ）.
[4]	Andreas M.Antonopoulos,“Mastering Bitcoin”,O’REILLY,2014.pp140-143.
[5]	Bitcoin Block Explorer - Blockchain.info(https://blockchain.info/en/ ， 2016 年 4 月参照)