オンチェーンデータ解析を用いた仮想通貨の価格推移に関する考察

A Study of Virtual Currency Price Trends Using On-Chain Data Analysis

徳永 翔 † 水谷 后宏 ‡,§
Sho Tokunaga Kimihiro Mizutani

1. はじめに

近年，分散台帳技術としてブロックチェーン技術が注目されている．分散台帳技術とは，分散型のデータベース（台帳）を実現する技術で，通貨といった交換対象となるデータを特定の企業や組織の中央サーバにて保存する従来の台帳管理とは異なり，異なる企業や組織間の複数のサーバを協調させることで 1 つの台帳管理をする技術である．分散台帳上では，データを送る側と受信する側にて利用されるアドレスは公開鍵にて表現され，公開鍵の値間でデータが送受信された場合，送受信対象となるアドレスとデータの組み合わせを分散台帳に記録する．この時，分散台帳への記録を実施した際に，何かしらのインセンティブを支払うという機構が実装されているアプリケーションとして仮想通貨がある．仮想通貨を管理する分散台帳では，仮想通貨の送受信記録を台帳に書き込むコンピュータに対して，仮想通貨のインセンティブを支払う仕組みを導入しており，インセンティブ目当てで，多数のコンピュータを用意し分散台帳への書き込みを実施している企業や組織が出現してきている．

このインセンティブとして支払われる仮想通貨は高騰しており，かつ社会的な通貨として利用されようとしており，仮想通貨の分散台帳を解析することで，その解析結果が経済動向の指標などに利用されるようになってきている．本研究では，仮想通貨における分散台帳への読み書きを監視するシステムの構築を目指し，ビットコインの保有量が多い組織を発見し，その組織の取引を監視することで，仮想通貨市場へのインパクトを計測することとした．

計測には，ビットコインの取引データをビットコインネットワーク上から抽出し，約 1000 万件の取引データのデータベースを構築した．本データベースにて扱うデータは，鯨と呼ばれる取引高が極めて高い個人・組織のデータとした．本データを解析し，取引額とビットコインの価格を時系列データとして可視化した．その結果，黎明期であった 2021 年前半にて，取引高が極めて高く，その量に応じて，ビットコインの時価総額が急激に高騰したことが判明した．これにより，鯨の取引き活動がビットコインの価格に与える影響を可視化することに成功した．

† 近畿大学大学院総合理工学研究科, Graduate School of Science and Engineering Research, Kindai University
‡ 近畿大学情報学部, Faculty of Informatics (KDIX), Kindai University
§ 近畿大学情報学研究所, Cyber Informatics Research Institute, Kindai University

2. 本研究で用いた技術について

2.1 ブロックチェーン

これまで，クラウドによって，計算機リソースが特定の組織によって集中的に管理され，サービスの創出がなされていた．しかしながら，クラウドへデータを保存することのリスクとして，クラウド事業者による情報流出や改竄があり，これらのリスクを背負った上で，クラウド事業者にデータ管理を委ねなければならないという問題が生じてきている．この問題を解決するため，個々の PC やデバイスが協調して，データを管理する手段としてブロックチェーン技術というものが創出されるようになってきた．ブロックチェーンネットワーク上では，各 PC やデバイスが，保存されているデータを認証し，改竄を検知する．また，分散的にデータを保存するため，クラウドのような単一障害点が発生しにくいという利点も存在している．ブロックチェーンが実現するデータ管理を可視化したものを図 1 に示す．

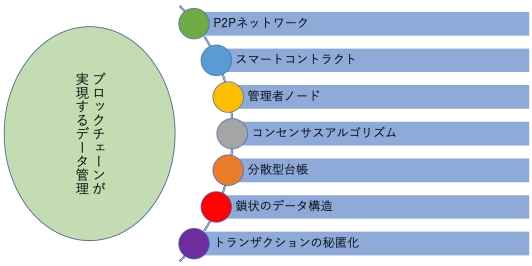

図 1: ブロックチェーンが実現するデータ管理 [1]

図 1 に示すとおり，P2P ネットワークと呼ばれる分散コンピューティング技術上でブロックチェーンが実装されており，スマートコントラクトのようなデータの認証機能であったり，コンセンサスアルゴリズムによって，複数人間で特定の物事の合意 (例: 記録されているデータが正しいかどうかの多数決) が実現されている．ブロックチェーン技術は多岐にわたって応用されるケースがあるが，最も有名な応用例に分散台帳技術というものがあ

る．次節では，その内容について説明する．

2.2　分散台帳技術

分散台帳技術とは，複数の計算機リソースから構成されるブロックチェーンネットワークにて，データを分散的に保存する技術である．扱うデータが，何かしらの台帳データ (取引データ) と仮定した時，分散台帳技術では以下の要素が用いられる．

ハッシュ値
取引データや取引データの塊（ブロック）の識別子．

トランザクション
時系列データとして管理される取引データ．

ブロック
ハッシュ値を識別子とするトランザクションの集合．

これらの要素の中で，分散的に管理される実態はブロックとなる．ブロックのハッシュ値が連続的に管理されるように，隣接するブロックのハッシュ値を，当該ブロックを管理するノードは保持することとなる．本形態は，Java 言語などで用いられる ArrayList のようにリスト構造となっており，そのリストのある領域を分割して，複数のノードで管理することとなる．

トランザクションが複数保存されるブロックが生成されていき，各ブロックのトランザクション内容を参照するためのハッシュ値がブロックに実装されている．ブロックチェーンネットワークにて特定のトランザクションを参照し，そのデータを認証または参照する場合は，ハッシュ値を元にブロックを特定したのち，ブロックからトランザクションデータを探すこととなる．具体例として図 2 を示す．

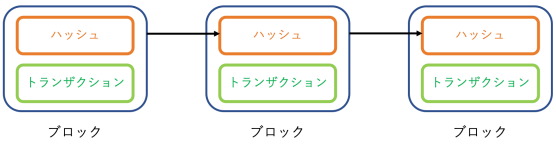

図 2: 分散台帳技術の例 [2]

図 2 のように，ブロックがハッシュ値によって接続されるデータ保存形態が分散台帳であり，特に，分散台帳上で扱われる全ての取引データをオンチェーンデータと呼ばれる場合がある．本研究ではオンチェーンデータを解析する．

2.3　ビットコイン

ビットコインとは，暗号通貨の一種であり，全ての取引データがオンチェーンデータとしてブロックチェーンネットワーク上に記憶されている．ビットコインネットワークでは，前述したブロックの中にハッシュ値，トランザクションデータと主にナンスという値が記録されている．ナンス (nonce) とは「number used once」の略で，「一度だけ使われる数」意味をもち，ブロックをブロックチェーン上に登録するか否かを決定する要素となっている．特定の規則 (例: 先頭 n ビットが 0 であるハッシュ値) を持つブロックのみが，ビットコインネットワーク上にて正式なブロックとして登録される．図 3 にビットコインネットワークの概要を示す．

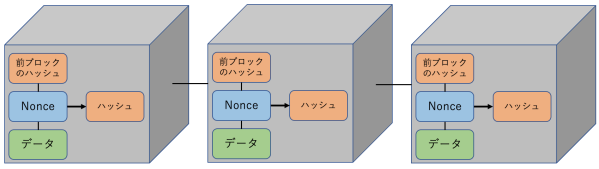

図 3: ビットコインネットワーク上でのデータ管理 [3]

ハッシュ値によってブロックの前後関係が決定されるだけでなく，ナンスを保持することで，正しいブロックを保存している．ビットコインネットワーク上にて保存される取引データは，送受信されるビットコイン量だけでなくタイムスタンプやウォレット (口座番号のようなもの) の送受金記録等々がある．

オンチェーンデータとしてのビットコインは，現在全ての仮想通貨のオンチェーンデータの約 45 パーセントを占める割合となっている．また，2021 年 1 月には，その割合が約 70 パーセントとなっており，オンチェーンデータとして，膨大な取引データが管理されていることが分かる．他の通貨に着目すると，イーサリアムは約 20 パーセントとなっており，ビットコインに次ぐ，オンチェーンデータ量となっている [4]．本研究では，一番取引量が多いビットコインに着目し，ビットコインのオンチェーンデータ解析を行うこととした．

イーサリアムの取引高は全体の約 20 パーセント程度となっているものの，ビットコインの取引高が下落している状況に対して，増加傾向になっている．これは，ビットコインよりもイーサリアムの取引高が増加しているということを示しているだけでなく，ビットコインの取引ユーザがイーサリアムの取引活動を活発化していること示している．つまり，ビットコインとイーサリアムの間で顧客争いのような闘争状態がなされていることがわかる．

3. オンチェーンデータの解析手法について

本研究ではビットコインのオンチェーンデータ解析を行うこととした．具体的には，ビットコインネットワーク上の取引量を観測及び解析し，ビットコイン価格との連動性を検証することとした．ビットコインの取引量を観測するためには，膨大な数のウォレットの取引量を解析しなければならない．これは，現実的に不可能なので，ウォレット残高が多い上位 50 位のウォレットを特定し，これらのウォレットの取引データを参照することとした．図 4 にビットコイン残高が高い上位 50 位のウォレットのうちの一部を示す．

Top 100 Richest Bitcoin Addresses

	Address	Balance △1w/△1m	% of coins
1	34xp4vRoCGJym3xR7yCVPFHoCNxv4Twseo wallet: Binance-coldwallet	252,597 BTC ($4,911,575,712)	1.32%
2	bc1qgdjqv0av3q56jvd82tkdjpy7gdp9ut8tlqmgrpmv24sq90ecnvqqjwvw97 wallet: Bitfinex-coldwallet	168,010 BTC ($3,266,836,082)	0.8800%
3	1P5ZEDWTKTFGxQjZphgWPQUpe554WKDfHQ	132,845 BTC ($2,583,068,781) +521 BTC / +5171 BTC	0.6958%
4	3LYJfcfHPXYJreMsASk2jkn69LWEYKzexb wallet: Binance-BTCB-Reserve	125,351 BTC ($2,437,364,744) / +8750 BTC	0.6566%
5	3M219KR5vEneNb47ewrPfWyb5jQ2DjxRP6 wallet: Binance-coldwallet	101,266 BTC ($1,969,046,191) / +48094 BTC	0.5304%
6	bc1qazcm763858nkj2dj986etajv6wquslv8uxwczt	94,643 BTC ($1,840,272,534)	0.4957%
7	37XuVSEpWW4trkfmvWzegTHQt7BdktSKUs wallet: 77604498	94,505 BTC ($1,837,589,947)	0.4950%

図 4: ビットコイン保有量の多い上位 50 位のウォレット（一部抜粋）[5]

本情報を取得するスクリプトを用意し，連続的に鯨のウォレット情報を取得した．また，取得したウォレット情報を用いて，ビットコインネットワークの公開ノードの API を通して，当該鯨のウォレットの取引履歴を取得することとした．図 5 に取引データの一部を記載する．


```
{
  "hash160": "****",
  "address": "****",
  "n_tx": 17,
  "n_unredeemed": 2,
  "total_received": 1031350000,
  "total_sent": 931250000,
  "final_balance": 100100000,
  "txs": [
    "--Array of Transactions--"
  ]
}
```


図 5: 取引データの一部

取引データを見ると，hash160 や address は鯨を識別する値と結論付けた．また，total がつく識別子の値については，当該鯨の送受信量の総量を示しており，balance は現在の残高を示している．また，txs については，実際の取引データの配列が格納されている．txs の中身については，ある鯨のウォレットから送信・受信されたビットコインの量であったり，宛先等々のデータが格納されている．

なお，単一のブロックの情報を取得する場合は，図 6 のような情報を取得することとなる．


```
{
  "hash": "****",
  "ver": 1,
  "prev_block": "****",
  "mrkl_root": "****",
  "time": 1322131230,
  "bits": 437129626,
  "nonce": 2964215930,
  "n_tx": 22,
  "size": 9195,
  "block_index": 818044,
  "main_chain": true,
  "height": 154595,
  "received_time": 1322131301,
  "relayed_by": "108.60.208.156",
  "tx": [
    "--Array of Transactions--"
  ]
}
```


図 6: 単一のブロック情報の一部

単一のブロックのハッシュ値が hash にて指示されており，prev_block が，このハッシュ値のブロックの前に存在するブロックのハッシュ値になっている．時刻 time はこのブロックが生成された時間を意味しており，nonce はこのブロックが認証に必要な値となっている．ブロック内のトランザクションは tx にて保管されている．なお，図 5 にて示した際，トランザクションデータは txs 内に保持していると述べたが，ブロック内のデータでは tx に保存されている．

次に，複数のアドレスに対応するトランザクションデータを取得する場合は，図 7 のようなデータ形式にて受信することとなる．具体的には addresses の構造体内に，複数のアドレスに対応する取引情報が保存されることになる．この時のトランザクションのデータは上位 50 位のデータとなっている．


```
{
  "addresses": [
    {
      "hash160": "****",
      "address": "****",
      "n_tx": 4,
      "total_received": 1401000000,
      "total_sent": 1000000,
      "final_balance": 1400000000
    },
    {
      "hash160": "****",
      "address": "****",
      "n_tx": 0,
      "total_received": 0,
      "total_sent": 0,
      "final_balance": 0
    }
  ],
  "txs": [
    "--Latest 50 Transactions--"
  ]
}
```


図 7: 受信するデータ形式

価格推移との連動性を検証するためには，取引が行われた時刻とその時の送受信量を取得する必要があるため，ビットコインネットワークにおけるトランザクションデータの仕様を調査した．図 8 に鯨の取引データの一部を記載する．

```
{
  "sequence": 4294967293,
  "witness": "****",
  "script": "",
  "index": 0,
  "prev_out": {
    "spent": true,
    "script": "****",
    "spending_outpoints": [
      {
        "tx_index": 8442428470053686,
        "n": 0
      }
    ],
    "tx_index": 2847304666865775,
    "value": 32350407,
    "addr": "****",
    "n": 8,
    "type": 0
  }
```


図 8: 解析対象のトランザクションデータの一部

取引データの中には，out や in といった構造が見られ，これらは，鯨から送信・受信した取引データの情報を意味することがわかった．このとき，value というキーに紐づく値が散見することがわかる．この value がビットコインネットワーク上での取引高を意味していると結論付け，この値と時刻を取得することとした．なお，取得したデータは 2019 年 4 月から現在に至るまでの内容となっており，取引データ数は 1000 万件を超える．補足として，他の格納情報について説明する．

tx_hash

取引データの ID を示している．tx_hash_big_endian も同様の意味を持つ．

tx_output_n

対応するトランザクションのビットコインの送出量．

tx_index

blockchain.info におけるトランザクション ID．blockchain.info とは，ビットコインの取引データを扱う公開ノード．

script

bitcoin を交換するために満たすべき条件．

上記で示した情報は，ほんの一部であり，膨大な情報がトランザクションデータとして格納されている．なお，これらの仕様についてはビットコインの仕様に基づいたものとなっており，ビットコインの仕様およびプロトコルについて深く熟知していなければ，解析に必要な情報を取得することは難しい．

4. 実験結果

前節で説明した手法を用いて，ビットコインの価格とビットコイン残高の高い鯨 1〜50 位の取引量の推移を計測し，長い期間アクティブな鯨の中から抜粋し，結果を表示する．図 9，10，11 にその結果を可視化した．実験結果から，大きく分けて 2 つの価格推移がわかると考えられる．

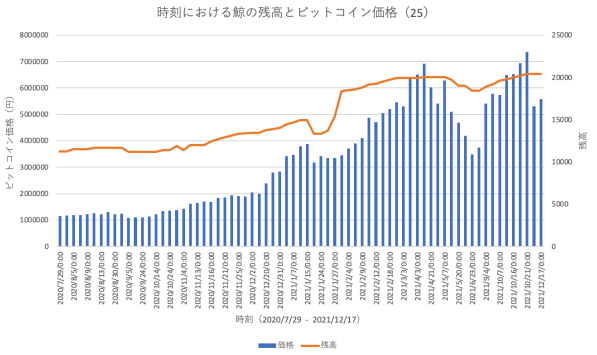

図 9: 25 番目の鯨の残高とビットコインの価格の推移（抜粋 1）

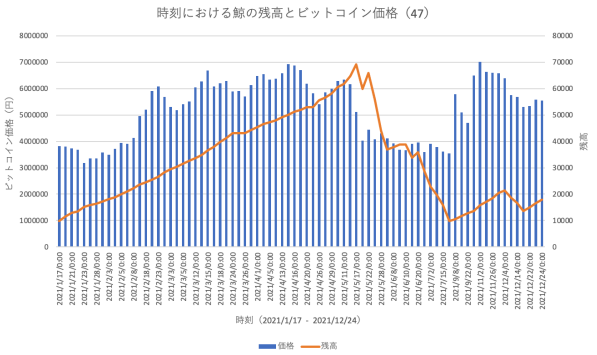

図 10: 47 番目の鯨の残高とビットコインの価格の推移（抜粋 2）

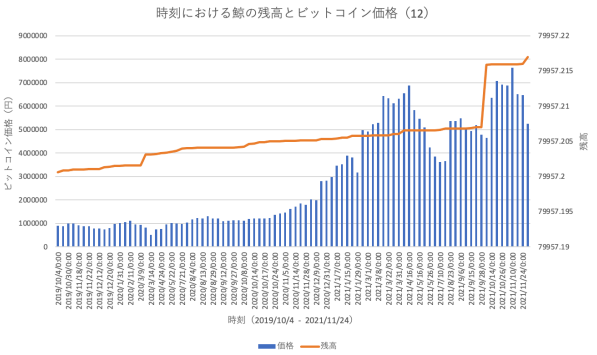

図 11: 12 番目の鯨の残高とビットコインの価格の推移（抜粋 3）

2021 年 1 月まで

2021 年 1 月までは，価格の推移は一定の範囲内で収まっている．2021 年 1 月以降と比べて鯨の取引残高が膨大になっていることがわかる．

2021 年 1 月以降

2021 年 1 月以降については，鯨の取引残高がそれ以前に比べて少ないように見受けられる．これは，鯨が，価格上昇を期待し，あまり残高を移動させないように行動していることがわかる．つまり，価格が暴騰しているため，静観しているという行動が可視化で判明した．

図 9，10 を再度調査すると，価格の上昇局面や下降局面にて，活発的に鯨の取引が行われていることがわかる．上昇局面では，特に 2021 年 1 月から 2021 年 5 月にかけて，取引高は少ないが，鯨の取引回数が増加していることが見受けられる．一方で，取引額がグラフの上界まで達しているケース，つまり 2020 年前半と 2021 年後半では，大きく値段が下がる直前に鯨の取引高が上昇していることがわかる．これは，鯨が，ビットコインを取引所に送りビットコインを現金化することで，ビットコインの価格が下落したという事象が発生したと考えられる．

5. 結論・今後の課題

近年，分散台帳技術としてブロックチェーン技術が注目されている．分散台帳技術とは，通貨といった交換対象となるデータを特定の企業や組織の中央サーバにて保存する従来の台帳管理とは異なり，異なる企業や組織間の複数のサーバを協調させることで 1 つの台帳管理を実現する技術である．分散台帳上では，データを送る側と受信する側にて利用されるアドレスは公開鍵にて表現され，公開鍵の値間でデータが送受信された場合，送受信対象となるアドレスとデータの組み合わせを分散台帳に記録する．この時，分散台帳への記録を実施した際に，何かしらのインセンティブを支払うという機構が実装されているアプリケーションとして仮想通貨がある．仮想通貨を管理する分散台帳では，仮想通貨の送受信記録を台帳に書き込むコンピュータに対して，仮想通貨のインセンティブを支払う仕組みを導入しており，インセンティブ目当てで，多数のコンピュータを用意し分散台帳への書き込みを実施している企業や組織が出現してきている．このインセンティブとして支払われる仮想通貨は高騰しており，かつ社会的な通貨として利用されようとしており，仮想通貨の分散台帳を解析することで，その解析結果が経済動向の指標などに利用されるようになってきている．本研究では，仮想通貨における分散台帳への読み書きを監視するシステムの構築を目指し，ビットコインの保有量が多い組織を発見し，その組織の取引を監視することで，仮想通貨市場へのインパクトを計測することとした．本研究で実施した解析では，ビットコインの取引データをビットコインネットワーク上から抽出し，約 1000 万件の取引データのデータベースを構築した．本データベースにて扱うデータは，鯨と呼ばれる取引高が極めて高い個人・組織のデータとした．本データを解析し，取引額とビットコインの価格を時系列データとして可視化した．その結果，黎明期であった 2021 年前半にて，取引高が極めて高く，その量に応じて，ビットコインの時価総額が急激に高騰したことが判明した．これにより，鯨の取引き活動がビットコインの価格に与える影響を可視化することに成功した．今後の課題としては，多種多様な統計的手法を用いて，鯨の挙動を定量化したいと考えている．また，他のオンチェーンデータに対しても同様の手法によって解析を実施し，その傾向が，本研究成果とどのように類似しているのかを解析したいと考えている．

参考文献

[1] Blockchain の概要: https://home.kpmg/jp/ja/home/services/advisory/risk-consulting/it-advisory/blockchain-base-system.html.

[2] 分散台帳技術の仕組み: https://greenapple-investment.com/system-of-blockchain.html.

[3] Blockchain network の仕組み: https://millionaireblog.net/2021/05/16/.

[4] オンチェーンデータの割合: https://coinmarketcap.com/ja/charts/.

[5] ビットコイン残高の多い上位 50 位のデータ: https://bitinfocharts.com/top-100-richest-bitcoin-addresses.html.