Bitcoin取引履歴の特徴量に基づくアドレス識別リスクの評価

松本 寛輝[1,a)]　菊池 浩明[2,b)]

概要：Bitcoin を代表とする暗号資産では, プライバシー保護の観点からアドレスの使用を一度限りとし, ユーザは自分のアドレスを明かさないようにすることが推奨されている. しかしながら, 取引履歴を長期間観測することで, アドレスについての特徴量を学習できるので, それに伴う匿名性への影響は明らかになっていない. ユーザは自身が管理しているアドレスを再利用して取引を行う場合, アドレス識別のリスクが不透明となっている. そこで, 本稿では掲示板やマイニングプールなどの Bitcoin アドレスを収集し, 取引構造に基づく特徴量を抽出し,4 つのアドレス集合に分類した. それぞれのアドレス集合に対して Jaccard 再識別手法を用いた識別実験結果を報告し, 実験結果に基づいて取引回数とアドレス識別の関係性を明らかにする.

キーワード：Bitcoin, 暗号通貨

Study on Risk of Bitcoin address to be identified from the features of transaction history

Hiroki Matsumoto[1,a)]　Hiroaki Kikuchi[2,b)]

Abstract: Crypto-currency such as Bitcoin suggests a use of one-time address for the privacy enhancement. However, it is not clear how much risk Bitcoin re-used addresses has when the transaction history is observed. When a Bitcoin address is reused, no one estimates a risk of privacy leakage. Our study tries to solve this problem. We collected several bitcoin addresses, classified into a bulletin board system, a mining pool and so on, and extracted the four features based on the transaction structures. The experimental results reveal the relationship between the transaction histories and the probability to identify addresses using the Jaccard distance.

Keywords: Bitcoin, Crypto-currency

1. はじめに

Bitcoin[1] を代表とする暗号資産は高い匿名性を持つとされ, 国を超えた送金や投資目的など様々な用途で利用されている. しかしながら,Bitcoin が持つ匿名性は Bitcoin アドレスが持つランダムな仮名に基づくものであり, 取引記録の統計情報よりアドレスの識別やユーザ居住地等の属性情報が推定される恐れがある. この問題に対してオープンソースプロジェクトの Bitcoin.org ではプライバシー保護の観点からアドレスの使用を一度限りとし, ユーザは自分のアドレスを明かさないようにすることを推奨している [2]. ユーザは取引ごとに新たなアドレスを作成し, 利用することでユーザの匿名性を高めることが可能となる.

その一方で, 用途によってはアドレスが長期間に渡り, 繰り返し利用されることがある. 代表的な例として, 掲示板

[1] 明治大学大学院先端数理科学研究科
Graduate School of Advanced Mathematical Sciences, Meiji University
[2] 明治大学総合数理学部
School of Interdisciplinary Mathematical Sciences, Meiji University
[a)] cs192026@meiji.ac.jp
[b)] kikn@meiji.ac.jp

や SNS などに自身のアドレスへ寄付を受け付ける目的で公開する場合や Mining pool 事業者などが営利用アドレスを意図的に使い回す場合等がある. Bitcoin アドレスが一定回数取引を行った際のアドレスの匿名性に関して次の研究がある.

Meiklejohn らは取引の入力アドレスを用いた識別手法を提案している [3]. この手法は,Bitcoin の送金を行うアドレス (入力アドレス) が 1 つのトランザクションに複数指定されている場合, 入力アドレスは同一のユーザ (管理者) によって管理されているという仕組みを利用している. この手法は, 送金時に署名を行えるのは全ての入力アドレスの秘密鍵を持っている必要がある, と言うヒューリスティックに基づいた識別手法である.

永田らは取引の送金先アドレス集合を用いた識別手法を提案している [4]. この手法は, ユーザごとに取引を行う相手が決まっているためユーザを追跡することが可能である, ということを仮定している. アドレスの取引頻度と送金先アドレス (取引の宛先アドレス) を用いて, 過去に行った取引記録からアドレスを識別する. (以下, 本稿では永田らの手法で用いられた送金先アドレス集合を宛先アドレス集合と記載する.)

永田らによる宛先アドレスを用いたアドレス識別手法では, 送金時に利用するアドレスはユーザによって変更することが可能であり, アドレスの識別に影響がある. また, アドレスの取引数はアドレスの識別に影響を与えない, と主張を行っている. 一方,Meiklejohn らは, 入力アドレスのヒューリスティックを用いたが, 出力アドレスは考慮していない.

この問題に対して本稿では, 取引回数とアドレスの識別率の関係性について, ある取引回数における識別リスクを定量的に示すことを目的とする. 先行研究 [3] の手法に対して, 我々は Bitcoin の取引構造に着目し, 送金元アドレスと出力アドレスを用いた新たな 2 つのアドレス識別手法を提案する.

(1) 送金元アドレスとは Bitcoin を受け取ったアドレスに対して送金を行ったことのある送金元のアドレスである. ユーザは自身のアドレスに対して送金を行うアドレスを制御できないため, ユーザが意識的にアドレスの識別率へ影響を与えることは困難であり, それ故に, この特徴量に基づいてユーザを識別できる見込みがある.

(2) 出力アドレスは 1 つのトランザクションに指定された複数のアドレスが Bitcoin の受け取り (出力アドレス) アドレスの集合である.

(3) 永田らの取引数はアドレスの識別に影響を与えないという結果は, 宛先アドレス集合を用いた評価結果であったので, 上記にあげた他 3 つのアドレス集合による識別方法については不明であった.

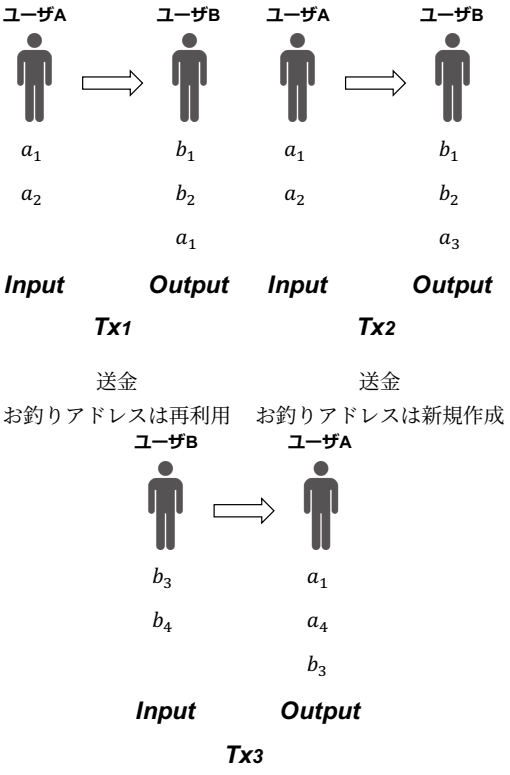

図 1 Bitcoin の送金, 受け取りを行う取引構造

我々は先行研究の手法を踏まえ,4 つのアドレス集合を用いたアドレス識別実験を行う. アドレス識別の評価指標として以下 3 つの課題を明らかにすることを目的とする.

- 取引回数が与えるアドレス識別率への影響
- 4 つのアドレス集合が与えるアドレス識別への影響
- Bitcoin アドレスの利用目的 (5 つの利用方法) が与えるアドレス識別率への影響

本稿では,2 章で Bitcoin アドレスの取引構造とアドレス集合に関する定義を行う. 3 章では Bitcoin アドレスの識別実験を取引回数, アドレス集合, 利用方法の 3 つの観点から実施し, アドレス識別実験結果に関する考察を行う.

2. 取引構造に基づくアドレス集合の定義

本章では,Bitcoin アドレスが行う取引構造に着目し,4 つのアドレス集合を定義する.

2.1 取引構造とアドレス集合

Bitcoin アドレスが行う取引構造の例を図 1 に示す. Bitcoin の取引 (トランザクション Tx) 中に, 送金を行うアドレスは $Input$, Bitcoin を受け取るアドレスは $Output$ に指定されている. 図 1 ではユーザ A が管理しているアドレス a_1,a_2 を用いてユーザ B が管理しているアドレス b_1,b_2 へ送金を行っている. このとき, ユーザ A は送金を行った際のお釣りを受け取るため $Output$ に自身のアドレスを指定することがある. 図 1 の Tx_1 では, 送金時に使用したア

表 1　a_1 についての 4 つのアドレス集合の定義

アドレス集合	定義	図 1 中での アドレス例
宛先アドレス S	a_1 が送金を行うアドレス	$\{a_3, b_1, b_2\}$
送金元アドレス R	a_1 に送金を行うアドレス	$\{b_3, b_4, a_2\}$
入力アドレス I	a_1 の送金時に同時に $Input$ に利用される アドレス	$\{a_2\}$
出力アドレス O	a_1 の受け取り時に同時に $Output$ に利用される アドレス	$\{a_4, b_1, b_2, b_3\}$

ドレス a_1 を再度用いてお釣りを受け取っているが,Tx_2 の様にユーザ A が新たなアドレス a_3 を作成してお釣りを受け取ることもある．Tx_3 では, ユーザ A が管理しているアドレス a_1,a_4 に対してユーザ B が管理しているアドレス b_3,b_4 から送金を行っている．このとき, ユーザ B は送金を行った際のお釣りを受け取るため $Output$ に自身のアドレス b_3 を指定している．

取引構造より a_1 が行った取引に利用されたアドレスを 4 つのアドレス集合に分類する．アドレス a_1 の**宛先アドレス集合** S は,a_1 から期間内に一度でも送金を行ったアドレスの集合である．宛先アドレス集合は先行研究で永田らが定義した送金先アドレス集合と同一である．アドレス a_1 の**入力アドレス集合** I は,a_1 が送金を行った際に同時に $Input$ フィールドに指定されたアドレスの集合である．入力アドレス集合は Meiklejohn らが定義した入力アドレスのヒューリスティックによって識別対象のアドレスを管理しているユーザが管理しているアドレスの集合と同一である．

既存の 2 つのアドレス集合に対して我々は新たに 2 つの集合を定義する．アドレス a_1 の**送金元アドレス集合** R は,a_1 に対して期間内に一度でも送金を行ったアドレスの集合とする．アドレス a_1 の**出力アドレス集合** O は,a_1 に対して送金が行われた際に, 取引の $Output$ フィールドに指定されたアドレスの集合とする．

4 つのアドレス集合の例を, 図 1 におけるアドレス a_1 のアドレス集合の例を用いて表 1 に示す．本研究ではアドレスの収集期間を 10 年間と半年間の 2 つの期間 D に分けて識別実験を行う．表 1 の 4 つのアドレス集合には識別を行うアドレス a_1 を含めないことに注意せよ．

2.2　データセット

本研究で取得したアドレスと取引の数を表 2 に示す．表 2 のアドレスについて, 我々は 2 回以上取引を行っていた 5 種類の利用方法に基づく Bitcoin アドレスを収集し, 関連する取引記録を Blockchain Explorer[5] より取得した．

Bitcointalk[6] は暗号資産に関する情報を交換する掲示板サイトである．Bitcointalk に登録しているユーザのプロフィールページに記載されているアドレスを収集した．

表 2　収集したアドレスデータ

利用方法	アドレス数	取引数	収集期間 D
Bitcointalk BBS	44,067	3,139,677	2009/1/4 - 2019/11/18
Bitcointalk BBS	1,968	28,832	2019/4/1 - 9/30
Bitcoin ATM	404	26,843	
Dark web	82	35,048	
Exchange	680	33,252	
Mining Pool	96	24,449	

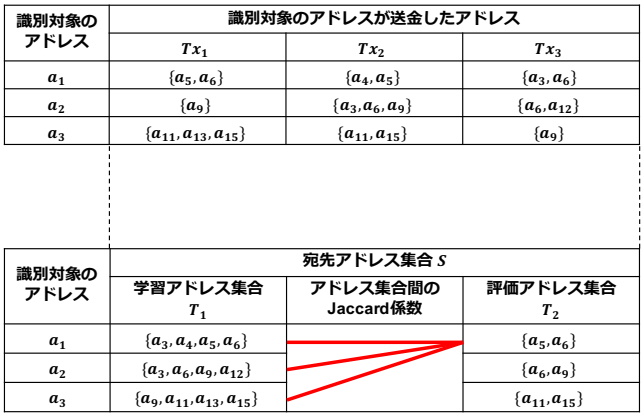

図 2　宛先アドレス集合 S と Jaccard 係数を用いた識別手法

Bitcoin ATM[7] は Bitcoin を預貯金することができるサービスである．カナダに設置された ATM を利用し ATM のアドレスおよび ATM を利用しているユーザのアドレスを収集した．Darkweb は匿名通信路 Tor ネットワークである．特殊なブラウザを用いてアクセスを行うウェブページ上に掲載されているプロモーション用のアドレスや違法商品 (クレジットカード番号等) を取り扱うアドレスを収集した．Exchange(交換所) はユーザの所有する Bitcoin を現金と交換するサービスである．交換所と取引を行っているユーザのアドレスを WalletExploer[8] よりアドレスを収集した．Mining Pool は多数のマイナーが協力し Bitcoin の取引情報をまとめたブロックに対して取引の検証を行い, 報酬を得るための仕組みである．2019 年 4 月 1 日から 9 月 30 日までの間にマイニング報酬を受け取ったアドレスを収集した．

3.　アドレス識別実験

本章では, はじめにアドレス識別成功の定義について説明を行う．次に取引回数やアドレス集合, アドレスの利用方法の 3 つの観点からアドレス識別実験を行い, アドレス識別成功率の結果を示す．

3.1　Jaccard 係数を用いたアドレス識別手法

アドレス識別の評価に Jaccard 係数を用いた集合の類似度を利用する．Jaccard 係数とは, ある集合 A と別の集合

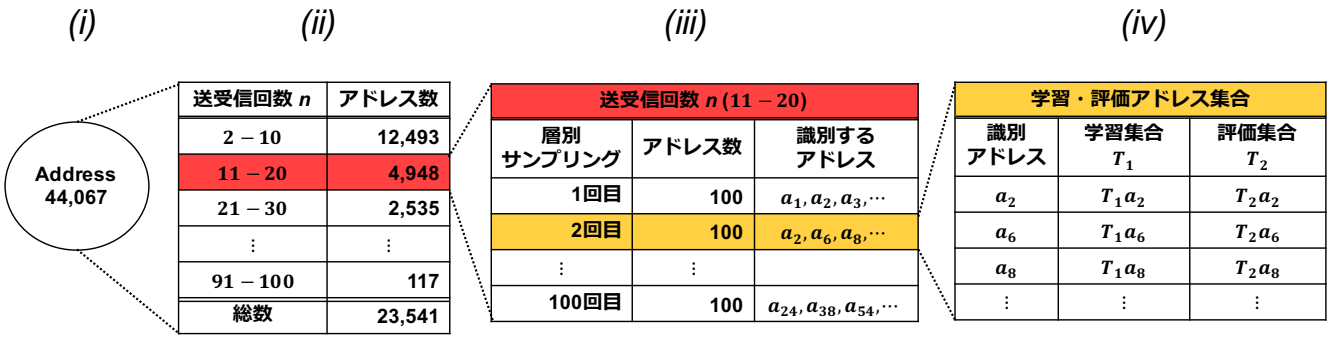

図 3　$D = 10$ 年間のアドレスサンプリング

表 3　$D = 10$ 年の間に利用された Bitcointalk アドレスと送受金回数

送受金回数 n	アドレス数
$2-10$	12,493
$11-20$	4,948
$21-30$	2,535
$31-40$	1,408
$41-50$	842
$51-60$	499
$61-70$	335
$71-80$	211
$81-90$	153
$91-100$	117
合計	23,541

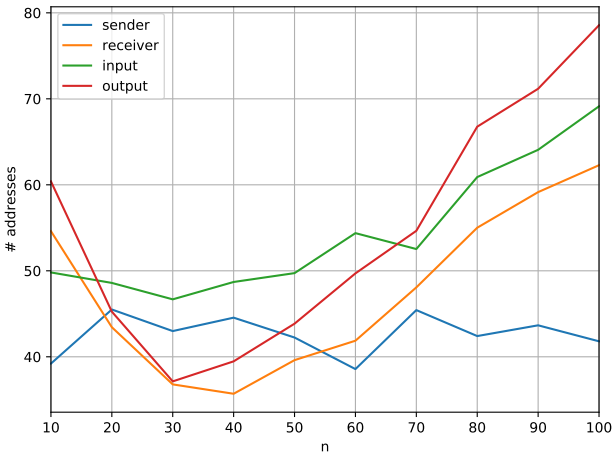

図 4　送受金回数についての平均アドレス個数の推移

B について $J(A,B) = \frac{|A \cap B|}{|A \cup B|}$ で定められる類似度である.

宛先アドレス集合 S を用いたアドレス識別手法の例を図 2 に示す. 識別対象のアドレス a_1, a_2, a_3 が送金を行った 3 つの取引 Tx_1, Tx_2, Tx_3 より, 識別対象のアドレスが送金を行った宛先アドレスを取得する. 取得したアドレスを学習アドレス数と評価アドレス数が 7 対 3 となるようにランダムにサンプリングし, 学習アドレス集合 T_1 および評価アドレス集合 T_2 を作成する.

図 2 において, アドレス a_1 を識別する場合の Jaccard 係数の値は以下のようになる.

$$J(T_1a_1, T_2a_1') = \frac{\{a_5, a_6\}}{\{a_3, a_4, a_5, a_6\}} = 0.50$$
$$> J(T_1a_2, T_2a_1') = \frac{\{a_6\}}{\{a_3, a_5, a_6, a_9, a_{12}\}} = 0.20$$
$$> J(T_1a_3, T_2a_1') = \frac{\{\phi\}}{\{a_5, a_6, a_9, a_{11}, a_{13}, a_{15}\}} = 0$$

ここでは学習アドレス集合と評価アドレス集合が共に a_1 となる $J(T_1a_1, T_2a_1')$ が他のアドレスと比較し Jaccard 係数の値が最も高いため, アドレス識別に成功している. また,Jaccard 係数が最も高い値を持つアドレスが複数存在する場合はアドレス識別が失敗したとみなす.

3.2　取引回数に基づくアドレス識別実験方法

識別対象のアドレスが送金, 受け取りを行った送受金回数を n で表す. 例として, あるアドレス a_1 の送受金回数 n が $21-30$ となる時, アドレス a_1 は 21 回から 30 回の送金, 受け取りをそれぞれ行っている.

識別対象のアドレスのサンプリング手法を図 3 に示す.

(i)　表 2 に示した約 10 年間分の Bitcointalk アドレス 44,067 個を対象とする.

(ii)　44,067 個の Bitcointalk アドレスのうち送受金回数が 2 回以上,100 回以下となる 23,541 個のアドレスを識別対象のアドレスとして使用する. 識別対象のアドレスについて, 送受金回数を表 3 に示す.

(iii)　識別に使用するアドレスは表 3 の送受金回数毎に 100 個のアドレスを 100 回, 層別サンプリングする.

(iv)　3.1 節の手法を用いてアドレス識別を行う.

3.3　実験結果

送受金回数 n 毎のアドレス識別結果を表 4, 表 5 に示す. 表 4 では識別対象のアドレス 100 個のうち識別に成功したアドレス数を送受金回数毎にまとめた. 識別できたアドレス数が最も多い送受金回数 n は $91-100$ の場合となり,4 つのアドレス集合を用いた識別結果の平均個数は 62.9 個であった. 4 つのアドレス集合のうち, 出力アドレス集合 O の 547 個が最も識別に成功したアドレス数が多い.

表 4 の取引回数とアドレス識別に成功した平均アドレス個数の変化を図 4 に示す. 取引回数とアドレス識別率の関

表 4　送受金回数についての平均アドレス個数

アドレス集合	送受金回数 n										合計
	$2-10$	$11-20$	$21-30$	$31-40$	$41-50$	$51-60$	$61-70$	$71-80$	$81-90$	$91-100$	
S	39.2	45.5	43.0	44.6	42.2	38.6	45.4	42.4	43.7	41.8	426
R	54.6	43.4	36.8	35.7	39.6	41.9	48.1	55.0	59.2	62.3	477
I	49.8	48.6	46.7	48.7	49.7	54.4	52.5	60.9	64.1	69.1	545
O	60.4	45.3	37.1	39.5	43.9	49.7	54.7	66.8	71.2	78.6	547
平均	51.0	45.7	40.9	42.1	43.9	46.1	50.2	56.3	59.5	62.9	

表 5　送受金回数についてのアドレス識別率

アドレス集合	送受金回数 n									
	$2-10$	$11-20$	$21-30$	$31-40$	$41-50$	$51-60$	$61-70$	$71-80$	$81-90$	$91-100$
S	0.39	0.46	0.43	0.45	0.42	0.39	0.45	0.42	0.44	0.42
R	0.55	0.43	0.37	0.36	0.40	0.42	0.48	0.55	0.59	0.62
I	0.50	0.49	0.47	0.49	0.50	0.54	0.53	0.61	0.64	0.69
O	0.60	0.45	0.37	0.39	0.44	0.50	0.55	0.67	0.71	0.79

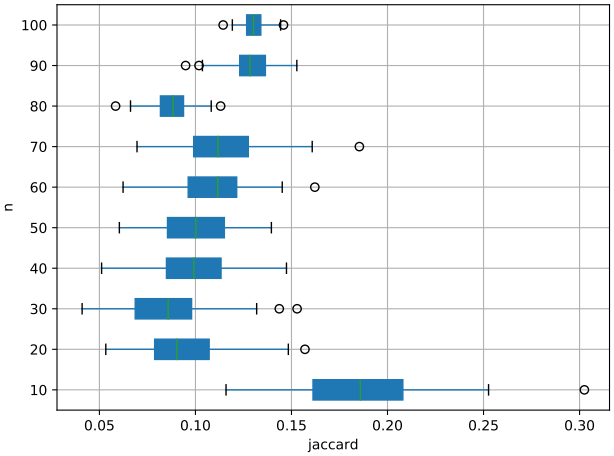

図 5　出力アドレス集合 O における送受金回数と Jaccard 係数の分布

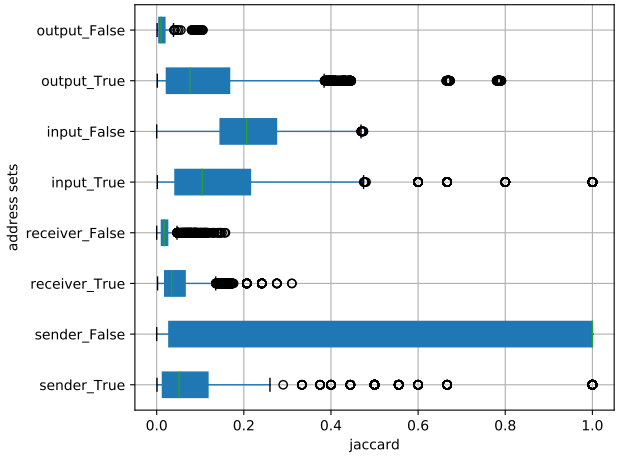

図 6　4 つのアドレス集合についての Jaccard 係数の分布

表 6　4 つのアドレス集合と平均アドレス個数の標準偏差

アドレス集合	標準偏差
S	2.2
R	9.1
I	7.2
O	13.4

係として送受金回数 $n = 40(31-40)$ のアドレスが最も識別率が低く,$n = 100(91-100)$ のアドレスが最も識別率が高い結果となった. 3 つのアドレス集合 R,I,O では n が 41 回以上を超えるとアドレス識別率が増加し, 出力アドレス集合 O では最大で 2.1 倍まで識別率が増加する.

表 4, 表 5 の結果より, 最も識別したアドレス数が多い集合 O の Jaccard 係数の分布を図 5 に示す. ここで, アドレス識別に成功した際の Jaccard 係数の分布を示している. 送受金回数 n の増加に伴い, アドレス識別に成功した際の Jaccard 係数の分布は小さくなっていた. 識別に成功したアドレスにおける学習アドレス集合と評価アドレス集合の類似度は 0.05 から 0.15 の間に 75%以上のアドレスが分布していた.

取引回数とアドレスの識別率の関係性として永田ら [4] の先行研究では取引回数とアドレスの識別率に相関がない結果を示していた. 表 4, 図 4 の結果より,4 つのアドレス集合の取引数と識別に成功したアドレス数の標準偏差を表 6 に示す. 集合 S は取引数に依存することなく識別率が推移しているように見えたが, 識別したアドレス数の標準偏差は 2.2 で安定していた. 集合 R, I, O の標準偏差の値がいずれも 7 以上の値であり, 集合 S と比較して 3 倍以上の変動がある.

3.4　アドレス集合に基づくアドレス識別実験結果

4 つのアドレス集合の Jaccard 係数の分布を図 6 に示す. 表 4, 表 5 の結果より最も識別したアドレス数が多い送受金回数 n が $91-100$ の Jaccard 係数を対象とした.

識別対象のアドレスの取引間隔を 1 日ごとに示した散布図を図 7 に,1 年毎のアドレス集合の大きさの推移を図 8 に

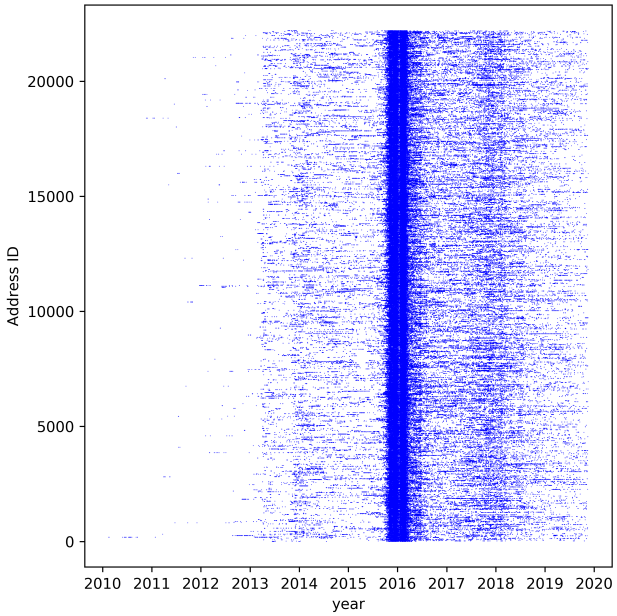

図 7　$D = 10$ 年の間におけるアドレス取引分布

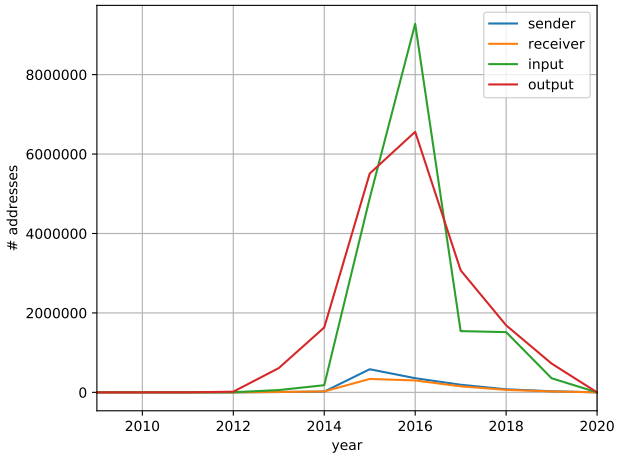

図 8　$D = 10$ 年の間におけるアドレス集合の大きさの変化

表 7　5 種類の利用方法の識別アドレス数

利用方法	アドレス数
Bitcointalk BBS	844
Bitcoin ATM	106
Dark web	49
Exchange	274
Mining Pool	85
合計	1,358

示す. 図 7, 図 8 共に 2016 年前後での取引数, アドレス数が最も多い.

3.5　5 種類の利用方法に基づくアドレス識別実験方法

表 2 で示した 2019 年 4 月 1 日から 9 月 30 日までの 5 種類の利用方法に分類された計 3,230 個のうち 2 回以上同じアドレスと取引を行っていた 1,358 個のアドレスを識別対象とする. 識別対象のアドレスの 5 種類の利用方法の内訳を表 7 に示す.

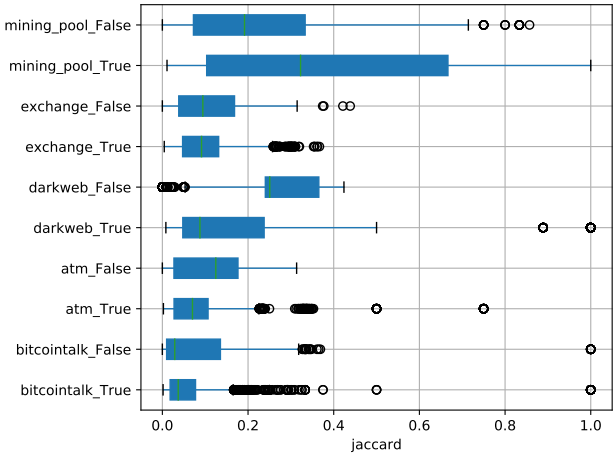

図 9　入力アドレス集合 I における 5 つの利用方法と Jaccard 係数の分布

識別に使用するアドレスは, 表 7 の 5 種類の利用方法別に 30 個のアドレスを 100 回, 層別サンプリングする. 層別サンプリングによって取得したアドレスに対して,3.1 節の Jaccard 識別を用いて識別実験を行った.

3.6　実験結果

5 種類の使用方法別のアドレス識別結果を表 8 に示す. ここで, 識別対象のアドレス 30 個のうち識別に成功したアドレス数を示す. 識別できたアドレス数が最も多い利用方法は Dark web であり,4 つのアドレス集合を用いた識別結果の平均個数は 22.3 個であった. また,4 つのアドレス集合のうち入力アドレス集合 I は 84 個と最も識別に成功したアドレス数が多い.

表 8 の結果より, 最も識別したアドレス数が多い集合 I の Jaccard 係数の分布を図 9 に示す. $D = 10$ 年間の送受金回数を用いて評価を行った Bitcointalk は識別成功時の Jaccard 係数の値は 0 より大きく 0.2 以下の範囲で 75%以上が分布していた. 最もアドレス識別率の高い Darkweb は 0 より大きく 0.5 以下の範囲に分布していた.

3.7　考察

本稿の実験結果より, 取引回数とアドレス識別率の関係は 3 つのアドレス集合 R,I,O で送受金回数が 31 回以上を超えるとアドレス識別率が増加し, 出力アドレス集合 O では最大で 2.1 倍まで識別率が増加した. 我々の実験結果は [4] と同様に, 集合 S は取引回数 (送受金回数) にアドレスの識別率が依存していないことを示した. 他の 3 つの集合 R, I, O は送受金回数に正の相関が見られた. 最も識別率の高い集合 O は, 表 6 より標準偏差の値が 13.4 であり, 取引回数の増減が識別率に大きな影響を与えていると考えられる. 集合 O の識別率が高い要因の一つに交換所との取引回数が多

表 8 5 種類の利用方法に関しての平均アドレス個数と識別率

アドレス集合	利用方法										合計
	BBS		ATM		Darkweb		Exchange		Mining pool		
	個数	%	個数	%	個数	%	個数	%	個数	%	
S	12.8	0.43	16.6	0.55	23.9	0.80	4.1	0.14	17.8	0.59	75
R	17.2	0.57	3.6	0.12	22.2	0.74	14.7	0.49	5.0	0.17	63
I	17.6	0.59	16.5	0.55	22.3	0.74	12.6	0.42	15.0	0.50	84
O	19.5	0.65	4.3	0.14	20.9	0.70	22.6	0.75	5.0	0.17	72
平均	16.8	0.56	10.2	0.34	22.3	0.74	13.5	0.45	10.7	0.36	

いことが想定される．交換所を経由して Bitcoin を購入する際に, 交換所のアドレスは取引の手数料を抑えるため複数のアドレスに対してまとめて送金を行うことがある．識別に使用した Bitcointalk のアドレス集合 O の識別率が高い要因は, 交換所との取引によって *Output* フィールドに複数のアドレスが指定されることが原因であると考える．表 8 の Exchange のユーザアドレスについても, 集合 O の識別率が最も高い振る舞いを示しているため,Bitcointalk で公開されているアドレスは交換所との取引を多く行っているアドレスと見られる．ユーザ毎に利用している交換所は異なり, 交換所との取引回数の増加に伴い,*Output* フィールドに指定されるアドレスも増加するため, 集合 O の識別率が高い結果を示したと考えられる．

表 8 のアドレスの利用方法を考慮した識別実験はアドレス集合毎に識別が容易なもの, 難しいものがあることを示している．送受金回数毎の識別実験を行った Bitcointalk アドレスの 4 つのアドレス集合の平均識別率は約 56%である．最も識別率の高い集合 O の値は 0.65 であり, 最も識別率の低い集合 S の値は 0.43 であることからアドレス集合の識別率に大きな差異は見られない．Bitcointalk では属性情報の異なる複数のユーザが登録を行っているため, アドレスの識別率に対して利用方法による偏りが少ないと考えられる．

Bitcointalk を除く 4 つの利用方法では,Exchange アドレスの集合 O の識別率が高く, 集合 S の識別率が低い．また,Bitcoin ATM や Darkweb,Mining pool において集合 S の識別率が最も高い．3 つの利用方法のアドレスには web 上などで公開されている事業者のアドレスが含まれており, 特定のアドレスに対して複数回送金を行うことがある．例えば,Mining pool では採掘の報酬を特定のマイナーに配当している．特定の宛先アドレスに対して複数回取引を行うことが識別率に影響を与えていると考えられる．

4. 結論

本稿では Bitcoin アドレスの取引回数がアドレス識別へ与える影響を明らかにするために 4 つのアドレス集合と Jaccard 係数を用いたアドレス識別実験を行った．我々は取引構造から新たに 2 つのアドレス集合を提案した．本稿の実験結果よりアドレスの取引回数の増加に伴い, 最大で 2.1 倍アドレス識別率が上昇することが明らかになった．4 つのアドレス集合では我々の提案した出力アドレスによる識別手法が最も精度が高い．5 つの利用方法においては Dark web で利用されているアドレスの識別率が最も高いことが明らかになった．

我々が行ったアドレスの利用方法を推定する研究ではアドレスの管理者による属性 (ユーザと事業者) によって推定率が異なっていた [9]．実験結果より 5 つの利用方法において集合ごとに識別率が異なる結果となったが, アドレスの属性情報 (ユーザの取引目的など) を考慮することで識別率に影響を与えることができると推測される．今後は本稿で使用した 5 つの利用方法に含まれる属性情報の有無が Bitcoin アドレスの識別にどれだけ影響を与えるかについて検討することを課題とする．

参考文献

[1] Satoshi Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System”, `https://bitcoin.org/bitcoin.pdf`

[2] Bitcoin.org プライバシーの保護 (`https://bitcoin.org/ja/protect-your-privacy`)

[3] Sarah Meiklejohn, Marjori Pomarole, Grant Jordan, Kirill Levchenko, Damon McCoy, Geoffrey M. Voelker, Stefan Savage, “A fistful of bitcoins: characterizing payments among men with no names”, In Proceedings of the 2013 conference on Internet measurement conference(IMC ’13), Pages 127–140, 2013.

[4] 永田倖大, 菊池浩明, “Bitcoin アドレスの送金先集合に基づく匿名性の評価”, 情報処理学会 第 80 回コンピュータセキュリティ研究発表会 (CSEC-80),pp. 1-6, 2018.

[5] Blockchain Explorer (`https://www.blockchain.com/ja/explorer`)

[6] bitcointalk (`https://bitcointalk.org/`)

[7] Coin ATM Radar Bitcoin ATM Map (`https://coinatmradar.com/`) 2020 年 2 月 14 日 16:32 参照

[8] WalletExplorer.com (`https://www.walletexplorer.com/`)

[9] 松本寛輝, 井垣秀星, 菊池浩明, “Bitcoin サービス業者と利用者アドレスの種類の推定と評価”, 情報処理学会 第 182 回マルチメディア通信と分散処理・第 88 回コンピュータセキュリティ合同研究発表会 (CSEC-88), 2020.