暗号通貨ビットコインと匿名性

2018年1月仮想通貨取引所「コインチェック」から顧客資産の仮想通貨NEM約580億円分が流出した。侵入者は取引所へのサイバー攻撃を行い不正にコインのアドレスを入手し、自身が管理する外部の9つの口座に対して19分間で送金を行った。その後、問題の口座の監視は続けられているが、複数の口座に分散送金され、不正者の身元をテイクダウンするのは極めて困難とみられている。

そもそも、仮想通貨とは何なのか？ なぜ追跡ができないのか？ ここでは、仮想通貨の原理とその匿名性の強さを解説し、本研究室で行っているビットコインの観察と匿名性を評価する研究成果を紹介する。

ビットコイン（Bitcoin）は、Satoshi Nakamoto氏が2008年にネットワーク上で発表した論文が基になっている代表的な暗号通貨である。特徴は、中央の管理者を持たず、分散型のピアツーピア（P2P）によって取引記録を共有していくことと、ブロックチェインと呼ばれる取引の履歴を束ねた世界共通の台帳（分散台帳）であろう。ブロックチェインは世界中で検証競争が行われ、10分間に1回のペースでその勝者にブロックを掘り当てた（マイニングされた）報酬が支払われている。これが原動力になり、取引の正当性が保証されている。時間を遡って支払額を偽造することが防止され、それがこの通貨の安全性の根拠になっている。

ブロックチェインの匿名性に焦点を絞る。取引を表す図1を見よう。

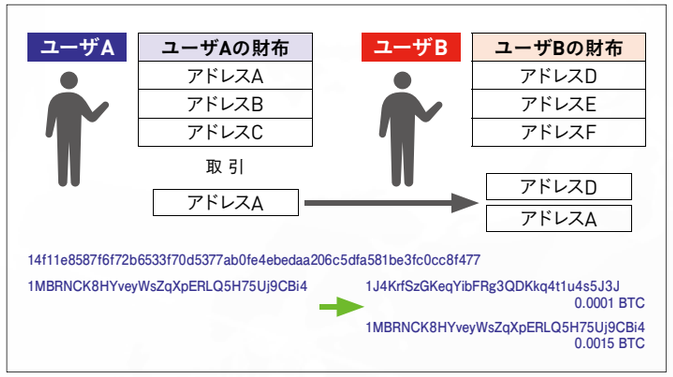

図1:ビットコインにおける取引

ユーザAはウォレット（財布）に複数のビットコインアドレスA、B、Cを管理している。アドレスは、公開鍵をハッシュ関数にかけて符号化（56進数）した値であり、図の例のように、英数約30文字の長さを持つ。乱数によって無料でいくつでも作ることができる。図の取引では、アドレスAからアドレスDに0.0001 BTC（≒100円）送金している。分散台帳を見れば、アドレスAからDへの取引があったことは世界中のピアに公開されている。しかし、1J4Kr…で始まるアドレスDが誰のものかは分からない。仮にそのアドレスの身元が割れても、また新たなアドレスを作ればよい。この乱数による匿名性のしくみを仮名化と呼ぶ。

仮名化による匿名性のレベルは十分ではない。誰のものか分からない乱数ではあるが、複数の取引に同じアドレスが使われれば、同一ユーザの取引であることが分かる。例えば、アドレスを含む取引の時刻分布を調べれば、そのユーザのタイムゾーンが推定できる。

そこで、本研究室では、2012年からの1年半に承認された10万個のブロックを収集し、既知のアドレス約500個の識別実験を行った。ユーザは特定の相手とのビットコインの送受信を繰り返すことに着目し、観測期間における送金先のアドレスの変動を調べた。図2は、送金先アドレス集合の分布を表している。学習観測期間の送金先集合と評価観測期間の送金先集合の類似度を定量化したところ、自分のアドレスの送金先との類似度（図の赤、self）と他人のアドレスとの類似度（青色、others）とは大きな差が見られた。Jaccard係数は二つの集合が等しいと最大値1、互いに素な集合だと0を取るので、他人とはほとんど共通の送金先がないことを表している。この送金先の類似度を特徴量とすると、アドレスからユーザを識別できる精度は80.5%に至った。学習データの割合を変動させても、平均6割が識別された。興味深いことに、取引数が多いユーザが必ずしも識別されるわけではない。

ビットコインの送金先履歴を用いてアドレスからユーザを識別する実験結果を報告した。仮名化による匿名性は高くはないが、100以上提案されて取引されている仮想通貨の中には、暗号プロトコルを用いてより高い匿名性を保証するものもある。理想的な通貨として、将来何が残るであろうか。今後の発展が楽しみである。

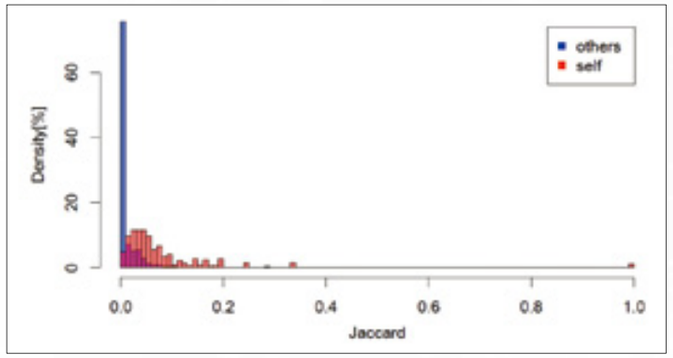

図2:送金先集合の自他の分布

PROFILE

菊池 浩明
Hiroaki Kikuchi

総合数理学部教授
専門：ネットワークセキュリティ、個人情報保護

1965年 秋田県生まれ
1990年 明治大学大学院博士前期課程修了
（株）富士通研究所、東海大学情報通信学部を経て
2013年より現職。博士（工学）

主な著書・論文
『IT Text ネットワークセキュリティ』（共著・オーム社・2017年）
『図解コンピュータ概論ソフトウェア・通信ネットワーク』（共著・オーム社・2017年）

所属学会
電子情報通信学会、日本知能情報ファジィ学会、IEEE、ACM各会員、情報処理学会フェロー

研究 最前線 THE FRONT LINE of RESEARCH