荒野のビットコイン

無名の男達が行う取引の形を求めて

SARAH MEIKLEJOHN, MARJORI POMAROLE, GRANT JORDAN, KIRILL LEVCHENKO, DAMON MCCOY, GEOFFREY M. VOELKER AND STEFAN SAVAGE

Geoffrey M. Voelkerはカリフォルニア州サンディエゴに位置するカリフォルニア州立大学にて   
コンピューターサイエンスの教授を務めています。彼はコンピューターシステム、ネットワーク、セキュリティを扱っています。[voelker@cs.ucsd.edu](mailto:voelker@cs.ucsd.edu)

Stefan Savageはカリフォルニア州サンディエゴに位置するカリフォルニア州立大学にて   
コンピューターサイエンスと工学の教授を務めています。彼はワシントン大学にてコンピューターサイエンスと工学の博士号を取得し、カーネギーメロン大学にて応用歴史学の学位を取得しています。Savageはスローン大学院とACMのフェローでありながら非常に堅実な人物であり、自分について書く時にも第三者の目線から捉えています。[savage@cs.ucsd.edu](mailto:savage@cs.ucsd.edu)

Sarah Meiklejohnはサンディエゴに位置するカリフォルニア州立大学にて   
コンピューターサイエンスと工学の博士号を取得中です。彼女はブラウン大学にて、コンピューターサイエンスの修士号と数学の学位を修めています。UCSDにて、彼女はMihir BellareとStefan Savageの両者から指導を受け、暗号化とセキュリティの分野で幅広く研究に情熱を傾けています。[smeiklej@cs.ucsd.edu](mailto:smeiklej@cs.ucsd.edu)

Marjori Pomaroleはサンディエゴに位置するカリフォルニア州立大学にて   
コンピューターサイエンスの学位を取得中です。彼女はグーグルとフェースブックにてインターンの経験があり、インフラのモニタリングについて学んでいます。[marjoripomarole@gmail.com](mailto:marjoripomarole@gmail.com)

Grant Jordanはサンディエゴに位置するカリフォルニア州立大学にて   
コンピューターセキュリティの修士課程に所属しています。彼はオンラインスパム配布とボットネットに関する研究を行っており、さらに空軍の研究所のために無人飛行機の開発に携わっています。[gejordan@cs.ucsd.edu](mailto:gejordan@cs.ucsd.edu)

Kirill Levchenkoはサンディエゴに位置するカリフォルニア州立大学にて上級研究員を務めています。彼はコンピューターネットワークとセキュリティに注力しています。[klevchen@cs.ucsd.edu](mailto:klevchen@cs.ucsd.edu)

Damon McCoyはジョージ・メイソン大学のコンピューターサイエンス学部において助教授を務めています。彼はコロラド州ボールダーに位置するコロラド大学から博士号を取得しており、匿名通信システム、サイバー／物理セキュリティ、ネット犯罪、ワイヤレスプライバシーなどを研究対象としています。[damon.mccoy@gmail.com](mailto:damon.mccoy@gmail.com)

ビットコインは権力分散型の仮想通貨であり、2009年1月に創設されて以来、利用が急速に広まっています。現金と同様に、ビットコインは匿名で保有が可能であり、既に知れ渡っている実名ではなく、偽名を用いて取引を行うことが可能です。本記事においては、ビットコインの匿名性が持つ限界を探ります。所有権の共有に関する経験則に基づき、偽名をクラスタ化することで、ビットコイン経済圏の中でも活発な領域が、かなりの割合で特定可能（実態またはユーザーを特定）なことを示します。並行して、ビットコインの仕組みについても詳細に解説します。

ビットコインはサトシ・ナカモト（偽名）によって2008年に創設された電子通貨です。名前の通り、ビットコインは現金によく似ています。取引は不可逆であり、取引の参加者は身元を明確にする必要がありません。お金の送り手と受け手の両方は偽名のみを通じてお互いを確認でき、特に手間をかけずに異なる偽名をいくつも使って取引に参加できます。しかし、現金と異なる特性が2つ、ビットコインにはあります：（１）完全な権力分散型です。つまり、中央に位置する単一の権威ではなくグローバルなピアツーピア型のネットワークがビットコインを管理して生成する役割を担います。また、（２）取引台帳が公開されています。取引は実名ではなく偽名を通して行えるのは確かですが、代わりに取引記録は全て世界中から参照できます。

創設以来、ビットコインへの注目は高まるばかりです。メディアのみならず、ビットコインを管理する方法を探る政府からも注目されています。注目が高まる理由のうち大きなものはビットコインの本質に根ざしています。政府組織は、資金洗浄や犯罪活動に利用される懸念を表明しており、また脆弱さや通貨としての成長力についても疑問が浴びせられています。2012年の終わりにビットコインの為替レートは異常な伸びを見せており、最終的に2013年4月に235ドルを記録し、その後100ドル前後に落ち着きました（2013年9月）。

ビットコインについての関心が高まっているにも関わらず、偽名で利用されるおかげで、ビットコインをどんな目的に、どのように使用されているかについてはあまり正しく理解されていません。抽象的なビットコインのプロトコルは、限界まで利用しきれば、かなり強固に匿名性を確保できるからです。しかしながら、現在に至っては、多くのユーザーが個人でデスクトップソフトウェアを利用するよりも、両替や財布などについては、第三者（例：銀行）によるサービスを通じてビットコインを貯めています。この文脈を念頭に置き、私達の目的は、こうしたサービスを利用するユーザーの行動傾向を通じて、匿名性を暴くことです。それを通じて、個人の匿名性を暴くのみならず、ネットワーク内のビットコイン取引の流れをも晒すことを試みます。

私達は2種類の手法を用います。まず、さまざまな種類のビットコイン取引を行うことで、確固たるデータを得ます。例えば、ビットコインを最大のビットコイン両替所であるMt. Gox内の口座に入れ、あるアドレスを確実にMt. Goxサービスに連携させます。後にそのビットコインを引き出す際に、別のアドレスを特定することができます。ここで得られた最小限の確固データを拡大するために、私達は次にビットコインアドレスを2つの経験則に従い*クラスタ化*しました。一つはビットコインが本質的に持つ特性を追求し、もう一つはビットコインネットワーク内における現在の専門用語の使われ方を追求します。クラスタ解析の層を確固データの集合の上に追加することで、一時的に全てのアドレスクラスタが特定のユーザーやサービスに所属しているとみなし、マーキングします。例えば、以前タグ付けしたアドレスが、あるクラスタ内のMt. Goxに所属していると判明した場合は、そのクラスタ内のアドレス全てを、Mt. Goxに所属しているとみなし、安心してタグ付けできます。

ビットコインの仕組み

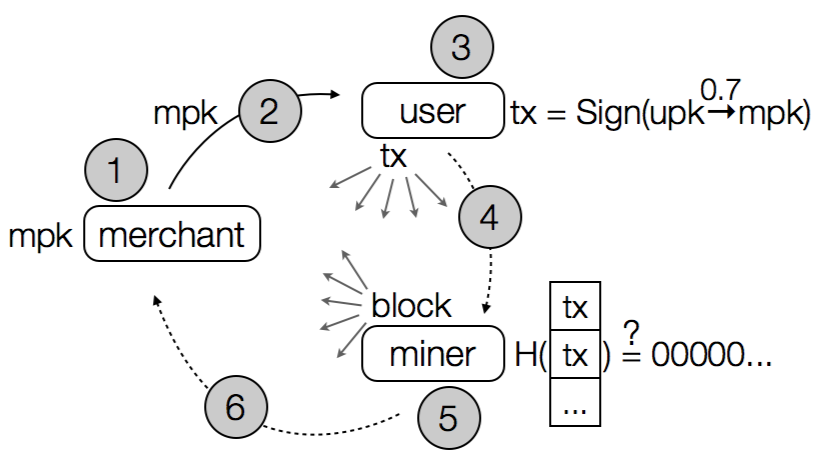
解析手順を解説する前に、ビットコインのプロトコルについて理解しておくことが必要です。暗号化の面では、ビットコインは2つの基本原理から出来上がっています：デジタル署名技術（実際には、ECDSA）と単方向ハッシュ機能（実際には、SHA-256）です。ユーザーの偽名は署名技術における公開鍵となります。ユーザーは署名キーのペアを生成して、任意に多くの偽名を作り出すことができます。この段落と以下の段落を通じて、「ビットコイン」とはピアツーピアネットワークと抽象的な特性を指し、「BTC」とは通貨単位を指します。公開鍵、アドレス、偽名は同じ意味で用いられます。

ビットコインは以下の要領で使用されます。あるユーザーが偽名を用いてビットコインを蓄えているとします。話を単純にするために、取引は入力と出力（のアドレス）が一対一で行われるとします。一般的に、取引の入力および出力アドレス数は自由に設定できます。このビットコインを送金するために、ユーザーは公開鍵を通じて特定されたビットコインの受け手と、彼の偽名がビットコインを受け取る取引の情報を含んだメッセージを作成します。（メッセージには他の情報も含まれます。）送金者は彼の偽名に連携している秘密鍵を使用してこのメッセージに署名し、署名を発行します。送金者は署名とメッセージを配信し、この両者が一体となって取引が形成されます。配信は彼の周辺にあるピアへ行き渡り、そのピアがさらに周辺のピアへ配信を行います（図1を参照してください）。

取引を配信する前に、各ピアは取引が有効であることを2つの方法を用いて確認します。まず、署名が認証され（署名技術は偽造不可能なため）、ビットコインの所持者によって正直に作成されたと確認できます。次に、同じ取引内容が既に使用されていないかどうかを確認します。後半の部分はビットコインの二重使用を防ぐ上で非常に重要な点です。これがあるがゆえに、全てのピアは全取引記録にアクセスできなければなりません。（あるいは、少なくとも受領されるビットコインがまだ使用されていない取引記録全て、になります。）ビットコインとは単一の物体ではなく、こうした取引記録の連鎖のことを指します。

こうした取引はネットワークに配信された後、ブロックへ吸収されます。ブロックは取引のタイムスタンプを記録し、有効性を裏付けます。ブロック作成のプロセスは*採掘（mining）*と呼ばれ、ビットコインを作成するプロセスでもあります。採掘者（ブロック作成を試みるユーザー）はまず、ブロックにまだ組み込まれていない取引情報を、知る限り全て取引プールに貯めます。手数料が含まれている取引がしばしば優先されます。ただし、現時点ではほとんどの取引は手数料を取りません。（例外は、入力や出力（アドレス）を多数持つ取引か、ビットコインの金額が大きい取引です。）採掘者は次にプールにて特殊な*ビットコイン生成用*の取引を作成し、この取引プールのハッシュ値を計算します。

図1:ビットコイン取引の仕組み：この例においては、ユーザーは業者として0.7ビットコインを支払代金として送金を試みます。（１）業者は公開鍵を新規作成するか、既存の公開鍵（mpk）を用いて（２）この公開鍵をユーザーへ送ります。デジタル署名を作成することで（３）、ユーザーは取引（tx）を作成し、彼の公開鍵（upk）から業者のアドレス（mpk）へ0.7ビットコインを送金します。（４）ユーザーは取引を彼のピアへ配信します。もし取引が有効であれば、ネットワーク全体へ配信されます。このようにして、採掘者と呼ばれる人々は自身の取引情報を得ることができます。（５）採掘者は、この取引および他の取引について、ハッシュ値が指定範囲内に収まっているかどうかを確認して、両者をブロックへ組み込みます。（6）ユーザーは取引を彼のピアへ配信します。もし取引が有効であれば、ネットワーク全体へ配信されます。このようにして、取引が認証されて全体のブロックチェーンへ追加され、ユーザーの支払を受け取った、と業者は確認できます。



採掘者の目的は、取引プールのハッシュ値が指定された数のゼロで始まることです。（取引プールは取引内容だけでなく、最も新しく生成されたブロックへの参照も含んでいます。）ここまでと、ここからの説明は採掘プロセスをある程度簡略化しています。実際には、採掘者の目的はある特定のハッシュ値より小さなハッシュ値を作成することです。要求されている最初のゼロ列の長さはネットワークの困難さに比例しており、現在のハッシュレートに依存しています。ネットワーク全体で10分毎にブロックを新規作成することが目的なので、困難さは都度調整されます。（例：ハッシュレートが上がると、困難さも同時に増えます。）

同じ取引プールを維持しながら、目標のハッシュ値を生成するために、採掘者は*使い捨てデータ*を入れ込みます。採掘プロセスは取引記録の収集から始まり、使い捨てデータは1に設定されます。これで目標範囲内のハッシュ値が生成されれば、採掘者は有効なブロックを作成したことになります。もしそうでなければ、使い捨てデータの値を増やし、再度試みることになります。

一旦、採掘者が有効なブロックを手に入れると、取引の配信と同様な形式でネットワークを通じて配信します。ハッシュ値が目標範囲内に収まっているかどうか、ピアが確認し、ブロックの有効性を検証します。新しいブロックは、別のブロックに参照された後、全体の取引台帳に組み込まれます。各ブロックは前回のブロックを参照するため、取引と同様にブロックもまた連鎖チェーンを構成します。ゆえに、この取引台帳は*ブロックチェーン*とも呼ばれます。

ハッシュ関数の単方向性は計算処理を非常に複雑にする特性を持つため、ブロックを作成する報酬として、採掘者は自身のビットコイン生成取引に結び付けられた公開鍵内にて一定額のビットコインを受け取ります。ビットコインの額はブロックチェーンの*高さ*により決まります。初期の状態では、報酬は50ビットコインでした。しかし、高さが210,000（つまり、2012年11月28日において、210,000ブロックが生成された後）、報酬は半額になりました。2100万ビットコインが生成されるまで半減が繰り返され、ゼロに達します。採掘者は、取引手数料でのみ報酬を受け取ります。よって、手数料は結果として増加するはずです。

結論としては、ビットコインネットワークに参加するピアが全てダウンロードする台帳とはブロックチェーンのことであり、チェーンは前回のブロックを参照しあうブロックの連鎖で構成されています。あるブロックがブロックチェーンに受け入れられるかどうかは、合議制が採択されています。あるブロックが有効だと、充分な数のピアが同意すれば（つまり、目標範囲内に収まっており、適切な数のビットコインを生成するとみなされた場合）、各ピアは自身のブロックを作成する際に、そのブロックを選択して参照します。よって、ブロックの採掘（加えて、後に続くビットコイン生成）は、システム要求ではなく、合意をベースにしたルールに従って行われます。こうしたブロックは、ブロック自体と同様にネットワーク内のピアによる合意に基づいて認証された取引の数々を含んでいます。その取引内にて、偽名同士のビットコインのやり取りが特定できます。

ビットコインがどこで使用されているか

2013年4月時点において、ブロックチェーンは1200万の公開鍵同士で1600万以上の取引を含んでいます。1100万以上のビットコインが既に生成されており（これはビットコインの絶対数の半分を既に超えています）、さらに何倍もの勢いで消費されています。1兆以上のビットコインが既に取引されています。

この流れを見て、誰もが自然に考える疑問は、ビットコインがどこで消費されているのか、です。2010年以降、さまざまなビットコイン用サービスが急速に生まれ、数を増やしています。もっともよく行われる活動である「両替」を通じて、ユーザーはビットコインを他の通貨に交換できます。ドルのような不換紙幣だけでなく、セカンドライフ内のLindenのような仮想通貨との交換も可能です。こうした両替商のほとんどは銀行としても機能しており、ユーザーのビットコインを保管してくれます。ただし、保管専用にオンライン財布サービスも存在しています。こうしたサービスと共に、窃盗のリスクも増加しています。実際、頻繁に発生しています。

ビットコイン採掘用のASICが2013年2月に発表され、毎秒6400万のSHA-256計算を処理できます。つまり、一般的なCPUやGPUですら、ブロック生成を行う能力は到底ありません。計算処理が膨大なため、*共同採掘サービス(mining pools)*がビットコイン経済圏で新たに注目を集めています。採掘者は作業の一部を肩代わりし（例えば、使い捨てデータの一部を検証する役割など）、代わりに貢献度に応じてビットコインを部分的に報酬として受け取ります。

貯めたり生成するだけでなく、ビットコインを使いたいと希望するユーザーはさまざまな業者を通じて行えます。ワードプレスにはBitPayと呼ばれる支払処理用ゲートウェイがあり、ビットコインでの支払を受付つつも業者には好みの通貨で支払が行われます。（ゆえに、業者はビットコインにまつわるリスクを無くせます。）ユーザーはビットコインでギャンブルを行うこともできます。BitZinoでポーカーを行ったり、Satoshi Diceなどの人気のサイコロゲームで遊ぶことができます。

最後に、ビットコインを犯罪目的に利用しようとするユーザーもおり、Torネットワークからのみアクセス可能なシルクロードなどのサイトで麻薬や禁制品を買うことができます。ビットコインをBitfogなどのサービスで混入（＝洗浄）することもできます。ビットコインを渡せば、ユーザーが指定したアドレスへ元のビットコインとはなんのつながりもない新規ビットコインを送付してくれます。

私達の解析は、こうしたサービスや他の多くのサービスと接触することから始まりました。合計で、私達は26の両替所と10の財布サービスに口座を持ち、25の異なる業者から購入を行い、そのうち9つはBitPayゲートウェイを用いてました。表1に私達が関わったサービスの一覧が掲載されています。私達が行った購入物は図2に示されています。私達はこうしたサービスと344の取引を行い、832個のアドレスに対して確実にタグ付けを行いました。（取引には任意で複数の入力アドレスが付加できるため、取引ごとに複数のタグ付けが可能でした。）さらに、既に所有権が公表されているアドレスを削除しました。例えば、ビットコインフォーラムにおけるユーザー署名などです。ただ、私達が自らの手で調査を行えるタグのみを対象にしています。

図2:私達がビットコインを用いて購入した品物はBitPantryでのビーフジャーキーからBitmitでのボストンの中古CDに及びます。緑色の枠で囲まれた品物はCoinDL（ビットコインの「iTunes」）で購入し、青色枠はBitmit（ビットコインの「eBay」）にて、赤色はBitPayゲートウェイ経由で購入しています。



表1：私達は多くのサービスと取引を行いました。ここに掲げるのは大まかなグループ分けです。



ビットコインアドレスのクラスタ化

理論的には、偽名でビットコインを使うことでリンク不能（unlinkability）という特性が得られます。つまり、ある偽名を用いて行った取引は、そのユーザーが別の偽名を用いて行った取引に結び付けられることはない、ということです。実際には、ビットコインの用法によっては、この匿名性が失われます。

有効なビットコイン取引を行うためには、送金者はビットコインに結び付けられた公開鍵に対応する秘密鍵を持たなければならないことに注目してください。ここで、あるユーザーが業者に10BTCを送金しようとしている、と仮定します。しかし、このユーザーはあるアドレスに4BTC、別のアドレスに6BTCしか持っていません。業者に支払を行う方法の一つは、新しいアドレスを作成して、4BTCと6BTCをそのアドレスに送り、その次に合計の10BTCを新アドレスから業者に送ることです。（実際、この方法が最も匿名性を維持できます。）ビットコインのプロトコルはもっと簡単で、効率の良い方法を採用しています。一つの取引には任意で複数の入力を持たせられるため、4BTCと6BTCのアドレスは両方、ある取引に対する入力として使用できます。この場合、受け取り手は業者になります。

この観察結果により、最初のクラスタ化経験則が得られます：もし2つのアドレスが同じ取引に対する入力として使用されたなら、それらは同じユーザーによって管理されています。この経験則は極めて確固たるものです。なぜなら、有効に取引を行うためには、送金者は全ての入力アドレスに対応する秘密鍵を知らなければならないからです。この事実はビットコインに関連した資料の中であまりにも多く使われているため、この解析を行うための無料ツールがオンラインで入手できるほどです。

私達のクラスタ化経験則の第二法則は、第一法則を拡張して両替が行われる過程に着目しています。ビットコインプロトコルにおいて、あるアドレスがビットコインを受け取ると、必ずそのビットコインを一度に全て消費します。（以下の点を思い出してください：各取引は以前の取引結果を必ず参照します。また、取引結果を参照できるのは一度だけです。）この時点でビットコインの量が送金者が送金したい額を超えている場合（例：アドレスに4BTCあり、業者には3BTC送金したい場合）、送金者は出力を2つ持つ取引を作成します。一つは実際の受け手（つまり、3BTCを受け取る業者）であり、もう一つは送金者が管理しているアドレスであり、それを通じて送金者は差額を受け取ります（つまり、残りの1BTCです）。

この行為により、第二のクラスタ化経験則が得られます：取引における差額アドレスは送金者が管理しています。差額アドレスは事前には他のアドレスと見分けがつかないので、特定する際には非常に注意が必要です。最初のステップとして、私達は標準のビットコインクライアントにおいては、差額アドレスは内部生成され、ユーザーですら知らないことに注目しました。（ビットコインを検査すれば、ユーザーが自分で特定することはいつでも可能です。）さらに、こうした差額アドレスは2回しか使用されません。一度目は取引における差額を受け取るため、二度目は別の取引において差額を全て使い切るためです。（この場合、クライアント側では差額を受け取るために新しくアドレスを生成します。）

取引内容を検査し、この使い捨てパターンに当てはまる出力を特定することで、私達は差額アドレスを特定しました。もし複数の出力がこのパターンに当てはまる場合、私達は慎重を期して、どのアドレスにもタグ付けを行いません。差額アドレスを特定するために1週間待つなど、いくつか予防策を講じながら、この方式を用いて、私達は350万の差額アドレスを特定しました。エラー率の最大幅は推定で0.17%です。エラー率の振り幅は推定しかできません。確固データが存在しないので、どれが本物の差額アドレスなのか、私達には真の意味では特定できません。この経験則に基づいてアドレスをクラスタ化した結果、1200万の公開鍵を330万のクラスタに圧縮しました。

まとめ

確固たるデータにクラスタ化解析の層を追加することで（すなわち、以前にタグ付けされたアドレスを含む全てのクラスタを一時的にタグ付けすることで）、私達は190万の公開鍵について、なんらかの実世界のサービスやIDを特定できました。ただし、多くの場合、IDは実名ではなく、例えばユーザー名やフォーラムでのニックネームです。これは公開鍵全体のわずか一部（16%ほど）に過ぎませんが、それでもネットワーク上に流れるビットコインのうち大きな割合で匿名性を解いています。

この成果を得るために、私達はまず、よく知られたビットコインサービスとのやり取りの内容を調査しました。さまざまなサービスのアドレスを大量に特定し（例：私はMt. Goxが管理しているアドレスのうち50万を特定し、シルクロードが管理しているアドレスのうち25万以上を特定しました）、私達はこうしたサービスとのやりとりを観察しています。例えば、両替所での預金や引き出しなどです。これでは取引に関わる個人の匿名を暴くことはできませんが（私達にはある個人があるサービスとやり取りすることはわかりますが、必ずしもその個人を特定できたわけではありません）、サービスへ入り、またサービスから出るビットコインの流れを明らかにする役には立ちます。

この手法を用いた解析が有用であることを証明するために、私達は犯罪活動に着目しました。ビットコイン経済においては、犯罪活動はさまざまな形式をとります。シルクロードで麻薬を取り扱ったり、単純に他人のビットコインを盗むものまで、あります。私達はシルクロードから出るビットコインの流れを追い（特に、ある悪名高いアドレスから）、また広く公開されている窃盗記録の数々から特定のサービスへのビットコインの流れを追及しました。窃盗者の中には洗練された混入テクニック（あるいは混入サービス）を用いてビットコインの流れをわかりにくくしようとした者もいました。しかし、ほとんどの場合、ビットコインの流れを追うのは非常に単純で、最終的には、私達は窃盗の現場（あるいはシルクロードからの引き出し）からさまざまな両替サービスへ大量のビットコインが流出するのを目撃できました。

上記でも確認されている通り、両替所に入れられるまで盗まれたビットコインを追跡しても、それだけでは泥棒を特定することはできません。しかし、それによってさらに匿名化を暴くことが可能で、そうすれば各種組織が盗まれたビットコインが振り込まれた口座の持ち主の特定を行うことは可能になります（たとえば、召喚礼状の発行など）。そうした両替所はビットコイン経済の入口、出口としてチェックポイントの役割を果たしているため（つまり、他にはビットコインを現金化する方法はほとんどありません）、私達の結論は、ビットコインを資金洗浄やほかの非合法目的に使用するのは（少なくとも現時点では）あまり魅力的な選択肢ではない、となります。