Biuletyn WAT
Vol. LXVIII, Nr 1, 2019

Metody deanonimizacji użytkowników wybranych kryptowalut na przykładzie bitcoina

PRZEMYSŁAW RODWALD[1], WITOLD SOBOLEWSKI[2],
MAJA RODWALD[3]

[1] Akademia Marynarki Wojennej, Wydział Nawigacji i Uzbrojenia Okrętowego,
Instytut Uzbrojenia Okrętowego i Informatyki, ul. Śmidowicza 69, 81-103 Gdynia,
[2] VS DATA Laboratorium Śledcze, ul. Świętojańska 55/16, 81-391 Gdynia,
[3] Prokuratura Rejonowa w Gdyni, ul. 10 Lutego 39, 81-364 Gdynia,
p.rodwald@amw.gdynia.pl, w.sobolewski@vsdata.pl, mrodwald@wp.pl

Streszczenie. Celem artykułu jest przedstawienie metod umożliwiających deanonimizację użytkowników kryptowalut na przykładzie najpopularniejszej z nich — bitcoina. Na wstępie przedstawiono podstawowe pojęcia oraz zasadę działania tej kryptowaluty, po czym dokonano autorskiej systematyzacji typów transakcji wzbogaconej o wykresy ukazujące ich ilościowe występowanie w łańcuchu bloków. W głównej części pracy przedstawiono heurystyki wykorzystywane przy deanonimizacji użytkowników. Następnie skupiono się na praktycznych wskazówkach ułatwiających implementację omówionych heurystyk w rzeczywistym systemie deanonimizacyjnym. Pokazane zostały także rzeczywiste scenariusze wykorzystania heurystyk wzbogacone o komentarze będące wynikiem doświadczeń płynących z przeprowadzonych przez autorów ekspertyz. W ostatniej części wskazano uwarunkowania prawne oraz istniejące narzędzia wspomagające przeprowadzanie czynności deanonimizacyjnych.
Słowa kluczowe: deanonimizacja, kryptowaluta, bitcoin
DOI: 10.5604/01.3001.0013.1466

1. Wstęp

W ostatnim czasie coraz więcej słyszy się o walutach wirtualnych, głównie z powodu spektakularnych wzrostów ich notowań (przykładowo bitcoin pod koniec 2017 roku osiągnął wartość 20 000 dol.). Bitcoin, będący ciągle najpopularniejszą kryptowalutą, powstał jako system mający charakteryzować się anonimowością.

Jedną z intencji autora/-ów[1] bitcoina było stworzenie systemu, w którym „każdy może zobaczyć, że ktoś przesłał środki do kogoś innego, ale bez możliwości powiązania tej transakcji z konkretną osobą"[2]. Warunkiem wystarczającym do przekazania środków w sieci Bitcoin jest znajomość adresu[3] odbiorcy, nie istnieje natomiast, w przeciwieństwie do tradycyjnych systemów bankowych, nadrzędna instytucja (np. bank) zarządzająca adresami i przypisująca im rzeczywiste dane personalne użytkowników. Transakcje przeprowadzane w sieci Bitcoin są pseudoanonimowe (ang. *psudoanonymous*), to znaczy, że szczegóły dotyczące transakcji (adresy wejściowe, wyjściowe, przekazywane kwoty, daty) są publicznie dostępne w łańcuchu bloków (ang. *blockchain*), ale nie istnieje w nim przyporządkowanie adresów do konkretnych podmiotów czy osób. Jednak ta publicznie dostępna historia wszystkich transakcji powoduje, że w połączeniu z innymi źródłami danych można skutecznie próbować deanonimizować użytkowników będących właścicielami adresów BTC.

2. Podstawowe pojęcia

W celu ujednolicenia terminologii stosowanej w niniejszym artykule zostaną najpierw zdefiniowane podstawowe pojęcia. **Bitcoin**[4] (oznaczany skrótem BTC) to zdecentralizowana kryptowaluta, opierająca zasady swojego funkcjonowania na kryptografii. Została przedstawiona w 2008 roku przez osobę (bądź grupę osób) o pseudonimie Satoshi Nakamoto [1], jednak do dziś nie udało się ustalić rzeczywistej tożsamości twórcy. Bitcoin dzieli się na mniejsze jednostki zwane satoshi (odpowiednik groszy w polskim systemie płatniczym), przy czym jeden bitcoin jest równy 100 000 000 satoshi. Bitcoiny przechowywane są na adresach. **Adres** (np. 15gZhgbX1f1JC1ZUxwWBqedXGuTaLaYYdb) to unikalny ciąg alfanumeryczny składający się z 26-34 znaków pozwalający na wysyłanie i odbieranie bitcoinów. Pierwotnie adresy rozpoczynały się zawsze od liczby 1[5] lub 3[6], zawierały wielkie i małe litery oraz cyfry alfabetu łacińskiego z wykluczeniem cyfry 0, wielkiej litery

1 Do dzisiaj nie wiadomo, kto ukrywa się pod pseudonimem Satoshi Nakamoto.

2 „The public can see that someone is sending an amount to someone else, but without information linking the transaction to anyone." [1].

3 Pojęcie „adresu" zostało zdefiniowane w dalszej części artykułu.

4 W literaturze przyjęło się oznaczać słowem bitcoin (pisanym małą literą) kryptowalutę, natomiast słowem Bitcoin (pisanym wielką literą) określa się całą infrastrukturę związaną z nim.

5 Adresy rozpoczynające się od cyfry 1 (ang. *Pay to PubkeyHash*) wymagają znajomości jednego klucza prywatnego powiązanego z danym adresem w celu wydania środków znajdujących się pod danym adresem.

6 Adresy rozpoczynające się od cyfry 3 (ang. *Pay to ScriptHash*) wymagają znajomości n z m kluczy prywatnych powiązanych z danym adresem w celu wydania środków znajdujących się pod danym adresem. Zostały utworzone w trakcie rozwoju sieci Bitcoin w celu wprowadzenia możliwości wykonywania transakcji kontrolowanych przez grupę użytkowników.

O, wielkiej litery I i małej litery l[7]. Adresy nie zawierają żadnej informacji na temat ich właściciela. Użytkownicy sieci Bitcoin mogą posiadać wiele adresów. Adres tworzony jest na podstawie klucza publicznego za pomocą kilkukrokowego algorytmu, wykorzystującego w swojej strukturze dwie kryptograficzne funkcje skrótu: RIPEMD-160 oraz SHA-256. Wszystkie transakcje w sieci Bitcoin zawierane są pomiędzy adresami i przechowywane są w publicznym rejestrze zwanym **łańcuchem bloków** (ang. *blockchain*). Rejestr ten składa się z kolejnych bloków (ang. *blocks*), dodawanych do łańcucha średnio co 10 minut, po potwierdzeniu transakcji znajdujących się wewnątrz danego bloku. **Transakcja** natomiast potocznie oznacza płatność. Polega na przenoszeniu bitcoinów z jednego adresu na inny adres. Każda transakcja, poza pierwszą w każdym bloku, posiada co najmniej jeden adres wejściowy oraz co najmniej jeden adres wyjściowy. W celu wykonania transakcji, czyli przekazania pewnej liczby bitcoinów, konieczna jest znajomość adresu nadawcy wraz ze skorelowanym z nim kluczem prywatnym nadawcy oraz adresu odbiorcy. **Portfel** (ang. *wallet*) to miejsce przechowywania par kuczy (prywatny, publiczny), przy czym nie ma potrzeby przechowywania samych adresów, gdyż są one generowane na podstawie klucza publicznego. Wyróżnia się kilka typów portfeli: portfel papierowy stanowiący po prostu wydruk pary kluczy (prywatny i publiczny), zwany też portfelem offline; portfel sprzętowy będący wyspecjalizowanym urządzeniem najczęściej w postaci klucza USB[8], portfel w postaci aplikacji będący oprogramowaniem przeznaczonym na konkretną platformę (komputer, smartfon), portfel „przeglądarkowy" stanowiący zewnętrzny serwis udostępniany przez pewien podmiot, do obsługi którego wystarczy przeglądarka internetowa. Bezpieczeństwo środków przechowywanych na poszczególnych typach portfeli maleje zgodnie z kolejnością, w której zostały przedstawione: największe jest dla portfeli papierowych, gdzie klucz prywatny znajduje się tylko na wydrukowanej kartce i znany jest tylko jej właścicielowi, a najmniejsze dla portfeli „przeglądarkowych", gdzie klucze prywatne użytkowników przechowywane są na zewnętrznych serwerach. Proces potwierdzania transakcji, czyli umieszczania ich w łańcuchu bloków, realizowany jest za pośrednictwem **górników** (ang. *miners*), którzy wykorzystują do tego moc obliczeniową swoich urządzeń[9]. Górnicy rywalizują ze sobą o to, który nowy blok

[7] Pierwotnie zostało to wprowadzone w celu zminimalizowania pomyłek przy prezentowaniu adresów. Od lipca 2017 roku, po zmianach wprowadzonych w pierwotnym protokole (ang. *soft fork*), sieć Bitcoin dopuszcza także adresy rozpoczynające się od znaków „bc1" zawierające wcześniej niedopuszczalne znaki (0, O, I, l).

[8] Do najpopularniejszych portfeli sprzętowych zaliczyć można: Ledger Nano S (ledgerwallet.com), Trezor (trezor.io), KeepKey (keepkey.com) i Ellipal (ellipal.com).

[9] Pierwotnie do potwierdzania transakcji w sieci Bitcoin wystarczała moc obliczeniowa procesorów, później moc obliczeniowa kart graficznych. Aktualnie proces wydobywania (ang. *miningu*) realizowany jest głównie za pomocą przeznaczonych do tego urządzeń opartych na układach ASIC, na przykład Antminer S9 (https://shop.bitmain.com/antminer_s9_asic_bitcoin_miner.htm) czy też Dragonmint T1 (https://halongmining.com/shop/dragonmint-16t-miner/).

złożony z transakcji zostanie umieszczony w łańcuchu bloków. Górnik, któremu uda się najszybciej rozwiązać „zagadkę kryptograficzną”[10], otrzymuje nagrodę w postaci nowych bitcoinów. Aktualnie górnicy raczej nie wydobywają nowych bitcoinów w pojedynkę, lecz łączą się w **spółdzielnie wydobywcze** (ang. *mining pools*), dzięki którym poprzez zwiększenie mocy obliczeniowej wzrasta prawdopodobieństwo wydobycia nowego bloku. Każda transakcja występująca w bloku może zawierać dodatkową **opłatę transakcyjną** (ang. *fee*). Jest to kwota oferowana przez nadawcę transakcji górnikom. Pierwsze transakcje w łańcuchu bloków były najczęściej pozbawione opłat transakcyjnych, aktualnie jednak praktycznie wszystkie transakcje zawierają te opłaty.

3. Rodzaje transakcji

W celu zrozumienia typów transakcji zostaną najpierw przedstawione podstawowe reguły nimi rządzące:

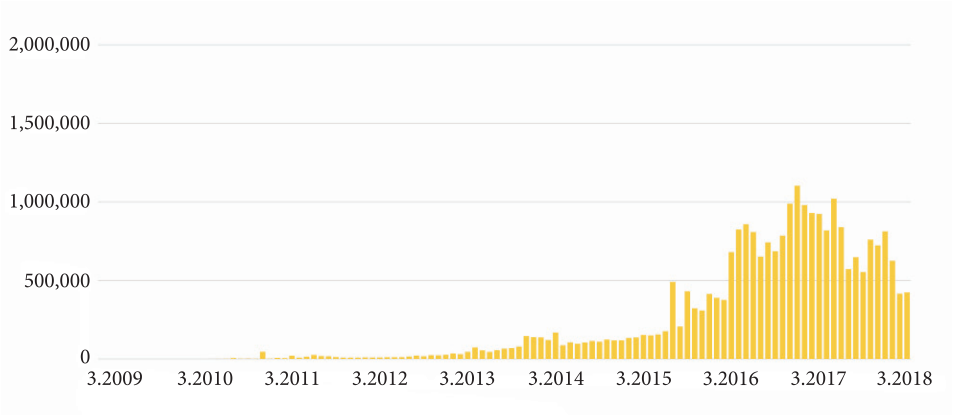
- Każda osoba może mieć wiele portfeli, a każdy portfel może składać się z wielu adresów.
- Tylko niezerowa liczba bitcoinów może być przesyłana z jednego adresu na drugi.
- Każda transakcja składa się z dwóch stron: wejściowej i wyjściowej. Strona wejściowa wskazuje, skąd pochodzą bitcoiny wchodzące w skład danej transakcji, natomiast strona wyjściowa pokazuje, dokąd są one wysyłane.
- Adresy znajdujące się po stronie wejściowej, zwane dalej w pracy adresami wejściowymi, muszą posiadać wystarczającą liczbę bitcoinów dla danej transakcji. Nie ma możliwości przesłania z danego adresu większej liczby bitcoinów, niż jest w jego posiadaniu.
- Cała kwota znajdująca się na adresach wejściowych musi zostać wydana w pojedynczej transakcji. Przy czym po stronie wyjściowej mogą znajdować się te same adresy co po stronie wejściowej.
- Transakcje znajdujące się w łańcuchu bloków mogą zawierać opłatę transakcyjną.
- Liczba bitcoinów znajdujących się po stronie wejściowej transakcji musi być równa liczbie bitcoinów znajdujących się po stronie wyjściowej transakcji powiększonej o ewentualną opłatę transakcyjną.

[10] Rozwiązanie „zagadki kryptograficznej” dla danego bloku polega na znalezieniu takiej wartości losowej, która w połączeniu ze skrótem wszystkich transakcji wchodzących w skład danego bloku da skrót (wynik działania kryptograficznej funkcji skrótu SHA-256) zaczynający się pewną liczbą zer, np. 000000000… Liczba zer dostosowywana jest do aktualnych możliwości obliczeniowych sieci Bitcoin w taki sposób, aby średni czas potwierdzenia bloku wynosił 10 minut.

Należy pamiętać, że przeniesienie bitcoinów z jednego adresu na inny nie musi oznaczać przekazania środków do innej osoby. Może mieć miejsce sytuacja, gdy ten sam właściciel przekazuje środki w ramach swojego portfela (z jednego adresu na drugi) lub na swój inny portfel.

3.1. Transakcja 1 → 1

Transakcja posiadająca jeden adres po stronie wejściowej i jeden adres po stronie wyjściowej. Występuje ona stosunkowo rzadko w łańcuchu bloków, a liczba transakcji tego typu w ujęciu miesięcznym została przedstawiona na wykresie 1. Jej procentowy udział dla wszystkich transakcji istniejących w łańcuchu bloków do końca marca 2018 roku wynosi zaledwie 8,35%.



Wykres 1. Liczba transakcji 1 → 1 w ujęciu miesięcznym

Polega na przekazaniu wszystkich środków znajdujących się pod jednym adresem na inny z uwzględnieniem ewentualnej opłaty transakcyjnej. Przykład takiej transakcji został zaprezentowany na rysunku i w tabeli 1.

Rys. 1. Schemat transakcji 1 → 1

Przed transakcją adres A (kończący się ciągiem …KTzy) dysponuje kwotą 0.01098 BTC. Podczas transakcji całość środków, pomniejszona o opłatę transakcyjną w wysokości 0.00000691 BTC, zostaje przekazana na adres B (...JhyS). Po transakcji adres B dysponuje kwotą 0.01097309 BTC.

Tabela 1

Przykład transakcji 1 → 1 (źródło: blochchain.info)

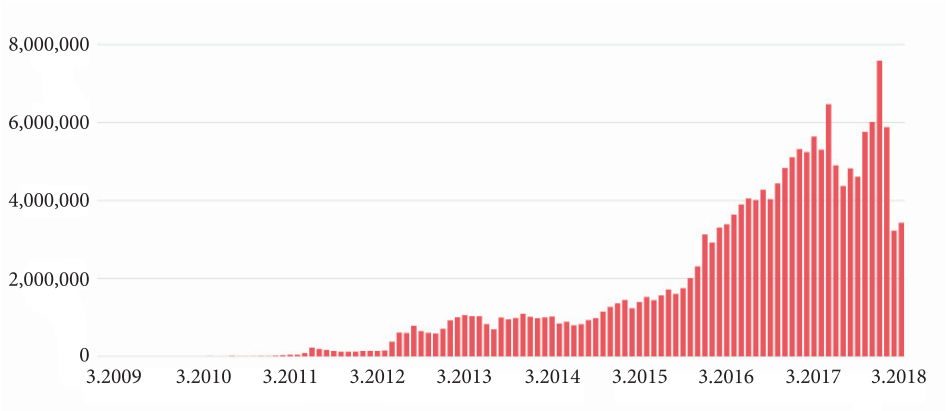
281b0e73b30f2bbb4f1859ed432cceb9e1431418794b27247f72ea834da73268

16Nj2vwbcidN1miG5qFvY2iUSmeZ1cKTzy	➡	1KejHZHms744ua6rcFAfo211c3HS71JhyS	0.01097309 BTC
			0.01097309 BTC

Podsumowanie		Przychody i wyjścia	
Rozmiar	223 (Bajtów)	Razem przychodów	0.01098 BTC
Waga	892	Razem wychodzących	0.01097309 BTC
Czas otrzymania	2016-05-15 20:16:39	Opłaty	0.00000691 BTC

3.2. Transakcja 1 → 2

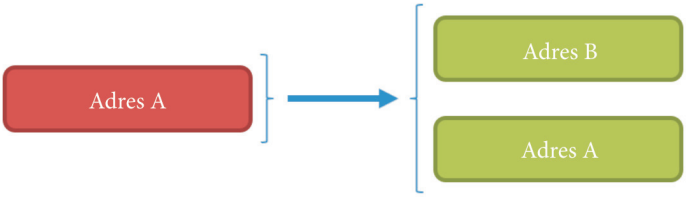
Transakcja posiadająca jeden adres po stronie wejściowej i dwa adresy po stronie wyjściowej. Jest to transakcja najczęściej występująca w łańcuchu bloków, co zostało przedstawione na wykresie 2. Jej procentowy udział dla wszystkich transakcji istniejących w łańcuchu bloków do końca marca 2018 roku wynosi aż 57,93%.



Wykres 2. Liczba transakcji 1 → 2 w ujęciu miesięcznym

Polega na przekazaniu wszystkich środków znajdujących się pod jednym adresem na dwa inne adresy z uwzględnieniem ewentualnej opłaty transakcyjnej.

Najczęściej jeden z adresów wyjściowych jest adresem, na który następuje pewna płatność, drugi adres stanowi natomiast miejsce, na które zwracana jest reszta (ang. *change*). W zależności od tego, czy dany portfel umożliwia przesłanie reszty na adres wejściowy, czy też za każdym razem generowany jest nowy adres, na który przesyłana jest reszta, mamy do czynienia z dwoma scenariuszami dla tego typu transakcji. Scenariusz z przekazaniem reszty na adres wejściowy został zaprezentowany na rysunku i w tabeli 2.



Rys. 2. Schemat transakcji 1 → 2 (jeden adres wyjściowy taki sam jak adres wejściowy)

Przed transakcją adres A (..V99F) dysponuje kwotą 0.49299667 BTC. Podczas transakcji część środków w wysokości 0.0416 BTC zostaje przekazana na adres B (…Luo2), natomiast pozostała część w wysokości 0.45089667 BTC powraca na adres wejściowy A. Powstała różnica w wysokości 0.0005 BTC została wydana na opłatę transakcyjną.

Tabela 2

Przykład transakcji 1 → 2, gdy jeden z adresów wyjściowych jest taki sam jak adres wejściowy (źródło: blochchain.info)

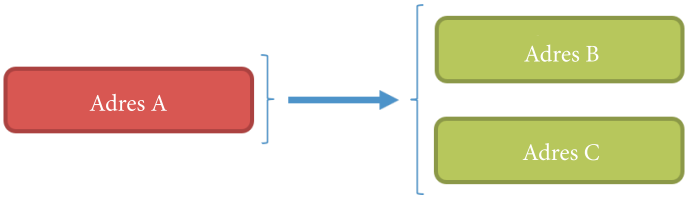
f157b7335528164a31710798133b7477c4c366f643f045f69d7fcc28f9448e87

39RwB8D6fg8mA1m7VGAGobKRtZM1vHV99F	→	3H6NEkRy4ukmdhTjS85hERXYRMrkb9Luo2	0.0416 BTC
		39RwB8D6fg8mA1m7VGAGobKRtZM1vHV99F	0.45089667 BTC
			0.49249667 BTC

Podsumowanie		Przychody i wyjścia	
Rozmiar	333 (Bajtów)	Razem przychodów	0.49299667 BTC
Waga	1332	Razem wychodzących	0.49249667 BTC
Czas otrzymania	2016-06-26 08:19:11	Opłaty	0.0005 BTC

Natomiast scenariusz z przekazaniem reszty na nowy adres został przedstawiony na rysunku i w tabeli 3.

Przed transakcją adres A (…TwUg) dysponuje kwotą 89.5676286 BTC. Podczas transakcji część środków w wysokości 89.38738 BTC zostaje przekazana na adres B (…RsQ8), a druga część w wysokości 0.18 BTC zostaje przekazana na adres C (…jq6j). Kwota w wysokości 0.0002486 BTC została przeznaczona na opłatę transakcyjną.



Rys. 3. Schemat transakcji 1 → 2 (adresy wyjściowe różne od adresu wejściowego)

Tabela 3

Przykład transakcji 1 → 2, gdy adresy wyjściowe są różne od adresu wejściowego (źródło: blochchain.info)

3c6409ecb159df64d59a55b9368a40787b4dd7eed721a11593c1b04e6e5328a5

1CuBgumLwcR2gSac6nW23uWuhMqVs6TwUg	→	1KLpmhwuiF4JhPJ8e7p6nFFbeChCzirSq8 1KYMJ5T5YJkLN7LYgs9Rrzj7FCvqWPjq6j	89.38738 BTC 0.18 BTC
			89.56738 BTC

Podsumowanie		Przychody i wyjścia	
Rozmiar	226 (Bajtów)	Razem przychodów	89.5676286 BTC
Waga	904	Razem wychodzących	89.56738 BTC
Czas otrzymania	2017-01-27 13:30:45	Opłaty	0.0002486 BTC

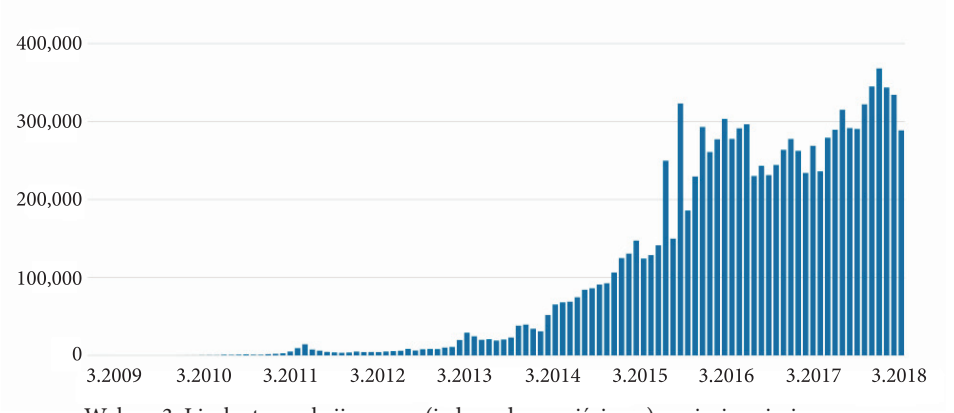
3.3. Transakcja n → m

Transakcja posiadająca kilka adresów po stronie wejściowej i kilka adresów po stronie wyjściowej. Polega na przekazaniu wszystkich środków znajdujących się pod adresami wejściowymi na adresy wyjściowe z uwzględnieniem ewentualnej opłaty transakcyjnej.

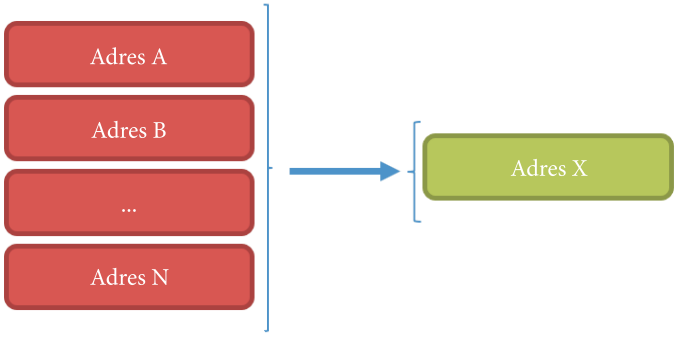
Pierwszym przykładem transakcji tego typu jest transakcja posiadająca kilka adresów po stronie wejściowej i tylko jeden adres po stronie wyjściowej. Liczba wystąpień transakcji tego typu została przedstawiona na wykresie 3. Jej procentowy udział dla wszystkich transakcji istniejących w łańcuchu bloków do końca marca 2018 roku wynosi 3,63%.

Transakcja ta polega na przekazaniu środków z różnych adresów wejściowych na jeden nowy adres wyjściowy. Jest to transakcja konsolidująca i została przedstawiona na rysunku i w tabeli 4.

Przed transakcją adresy wejściowe (…jq6j, …Awzn, …, …3VJf) dysponują sumaryczną kwotą 4.6317288 BTC. Podczas transakcji wszystkie środki z tych adresów zostają przekazane na adres wyjściowy (…R7kR), przy opłacie transakcyjnej w wysokości 0.0017288 BTC.



Wykres 3. Liczba transakcji n → m (jeden adres wyjściowy) w ujęciu miesięcznym



Rys. 4. Schemat transakcji n → m (jeden adres wyjściowy)

Tabela 4

Przykład transakcji n → m, jeden adres wyjściowy (źródło: blochchain.info)

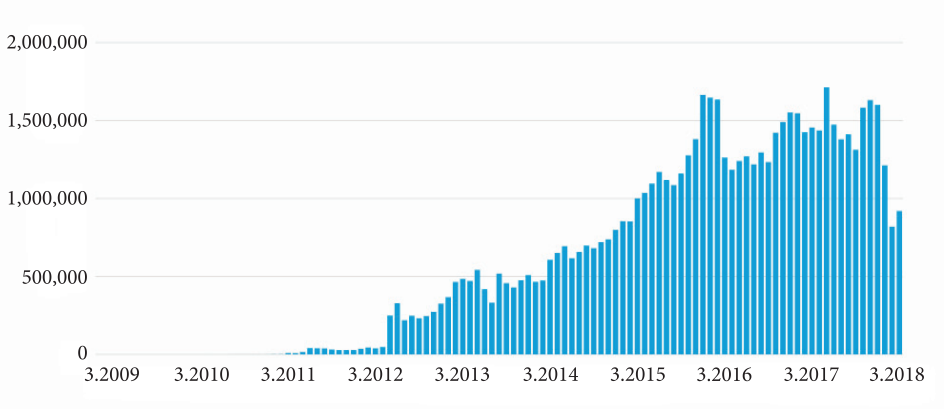
3b978e586018d51de04e1a2624081bf0f9679b40af5aada49a903607fc3a052b

1KYMJ5T5YJkLN7LYgs9Rrzj7FCvqWPjq6j 1MUqFqUZsB6BKppMPRkah5cnQYp77ZAwzn 1Lfkzuf6Hr86wRWCtWBsa2vefqWKKZmQJ5 1HgEKpzKzQ98EfGALcUF32Ngn4aAcivG1A 1LhCu8jh8AVLEm9qETAQBX3PHqJNhDnmvQ 1MRsGAVUJnm1PkxhuwVkZqmpKG2ec6EFLa 1D8k5b4b4ZCyBf1WpaRdSswQxNnXgu1MCM 1Je2XMNhnxz5W3q8cuYE52Cx7GHzwG3VJf	➡	1Pm5mYudt5jWYP29AvM6g9fDcagidGR7kR	4.63 BTC

4.63 BTC

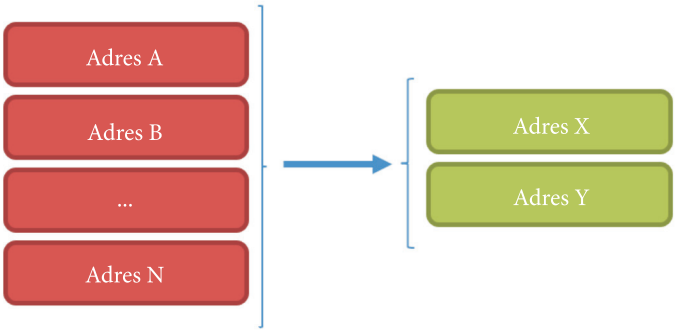
Podsumowanie		Przychody i wyjścia	
Rozmiar	1668 (Bajtów)	Razem przychodów	4.6317288 BTC
Waga	6672	Razem wychodzących	4.63 BTC
Czas otrzymania	2017-01-31 11:42:24	Opłaty	0.0017288 BTC

Innym przykładem transakcji tego typu jest transakcja posiadająca dwa adresy wyjściowe. Jest to drugi najczęściej występujący typ transakcji, a jego częstotliwość występowania została przedstawiona na wykresie 4. Jej procentowy udział dla wszystkich transakcji istniejących w łańcuchu bloków do końca marca 2018 roku wynosi 21,74%.



Wykres 4. Liczba transakcji n → m (dwa adresy wyjściowe) w ujęciu miesięcznym

Sytuacja ta odpowiada najczęściej scenariuszowi, gdy użytkownik chce dokonać płatności z kilku swoich adresów na adres odbiorcy, a reszta trafia na nowy adres wyjściowy użytkownika. Na rysunku i w tabeli 5 została pokazana transakcja z kilkoma adresami wejściowymi i dwoma adresami wyjściowymi.



Rys. 5. Schemat transakcji n → m (dwa adresy wyjściowe)

Przed transakcją adresy wejściowe (...mU7o, …8tbn, …, …LUbG) dysponują sumaryczną kwotą 1.41353384 BTC. Podczas transakcji część środków w wysokości 1.40111321 BTC zostaje przekazana na adres wyjściowy X (…nNYG), a część 0.01000463 BTC na adres wyjściowy Y (...kvdg) i stanowi resztę transakcji, przy opłacie transakcyjnej w wysokości 0.002416 BTC. Za tym, że w podanym przykładzie to właśnie adres Y stanowi resztę transakcji, przemawiają kwoty znajdujące, którymi dysponują adresy wejściowe. Jeśli płatność miałaby być przekazana na adres X, wówczas po stronie wejściowej transakcji wystarczyłby dowolny z zaprezentowanych adresów, gdyż każdy z nich dysponuje kwotą powyżej 0.012 BTC, a więc wystarczającą, aby przekazać 0.01000463 BTC[11].

Tabela 5

Przykład transakcji n → m, dwa adresy wyjściowe (źródło: blochchain.info)

4410c0c086eaa45365b7e6ca35953539304949c90e01acf8a96d92dff35cb7d5

19knN1QLjBytwr5FV5eVCgSAfUwuRqmU7o 1PVMtEbbR5twrXBPtRtmZEZ7VrzkYA8tbn 1GkpT5sLP4DyQki8ZfBqzLYK2U9pQCbpvr 1JLicegbLWiyyJiDbhQfUFwQKfQPxs4dkD 14fMCHMNV53wkak1jv3dEb9erCWtSRiByx 1PE51qiV8m3KWZuDheBvzcWCQNfzxpenuv 1D4o8ECXFckiFkSDxvFi2NcQX7kWpSLUbG	→	1PxEEisNJDc3uAe75vbAsb4jkUyV8QhNYG 19zJb3sg78aHv1Jg9B5aVjUwSQeWQ6kvdg	1.40111321 BTC 0.01000463 BTC

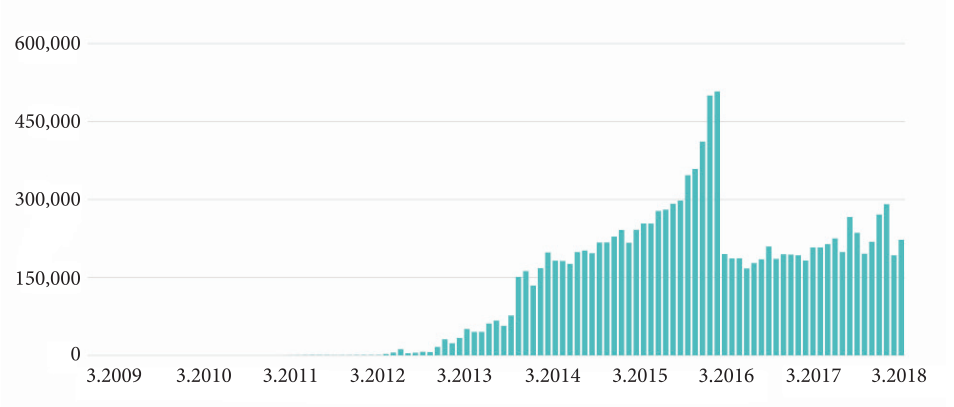
1.41111784 BTC

Podsumowanie		Przychody i wyjścia	
Rozmiar	1175 (Bajtów)	Razem przychodów	1.41353384 BTC
Waga	4700	Razem wychodzących	1.41111784 BTC
Czas otrzymania	2016-07-09 09:28:09	Opłaty	0.002416 BTC

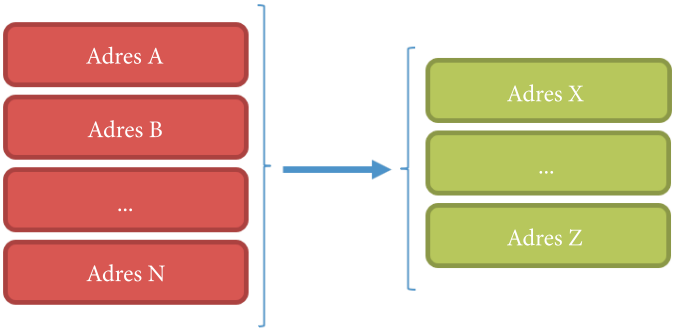
Jeszcze innym przykładem transakcji tego samego typu jest transakcja posiadająca wiele adresów po stronie wejściowej i wiele adresów po stronie wyjściowej. Częstotliwość występowania transakcji tego typu została przedstawiona na wykresie 5. Jej procentowy udział dla wszystkich transakcji istniejących w łańcuchu bloków do końca marca 2018 roku wynosi 4,18%.

Jest to transakcja najczęściej dokonywana nie przez pojedynczych użytkowników, lecz przez serwisy obsługujące znaczne liczby transakcji na przykład giełdy, ale również tak zwane miksery (ang. *mixers*). Głównym celem przeprowadzania takich transakcji jest odpowiednio: oszczędność na opłatach transakcyjnych lub zwiększenie anonimowości. Przykład takiej transakcji został zaprezentowany na rysunku i w tabeli 6.

[11] Nie zostało to przedstawione w tabeli, lecz można dokonać sprawdzenia na przykład pod adresem https://www.sydeus.pl/pl/search/4410c0c086eaa45365b7e6ca35953539304949c90e01acf8a96d92dff35cb7d5



Wykres 5. Liczba transakcji n → m (wiele wejść i wiele wyjść) w ujęciu miesięcznym



Rys. 6. Schemat transakcji n → m (wiele wejść i wiele wyjść)

Przed transakcją adresy wejściowe (…WB6o, ...oggM, ..., ...317X) dysponują sumaryczną kwotą 35.56430611 BTC. Podczas transakcji środki zostają rozdysponowane na adresy wyjściowe (...Hv9d, ...CTCE, ..., ...Z99a), przy opłacie transakcyjnej w wysokości 0.0005 BTC.

Jeszcze innym przykładem transakcji typu n → m jest transakcja posiadająca wiele adresów po stronie wyjściowej i tylko jeden adres po stronie wejściowej. Częstotliwość występowania transakcji tego typu została przedstawiona na wykresie 6. Jej procentowy udział dla wszystkich transakcji istniejących w łańcuchu bloków do końca marca 2018 roku wynosi 4,00%.

Jest to transakcja najczęściej dokonywana przez pojedynczych użytkowników, chcących oszczędzić na opłatach transakcyjnych (wówczas liczba adresów wyjściowych jest relatywnie mała — kilka), lub duże podmioty takie jak giełdy czy kantory

przekazujące środki z jednego ze swoich adresów na adresy swoich użytkowników (wówczas liczba adresów wyjściowych jest stosunkowo duża — kilkanaście, kilkadziesiąt). Przykład takiej transakcji został zaprezentowany na rysunku i w tabeli 7.

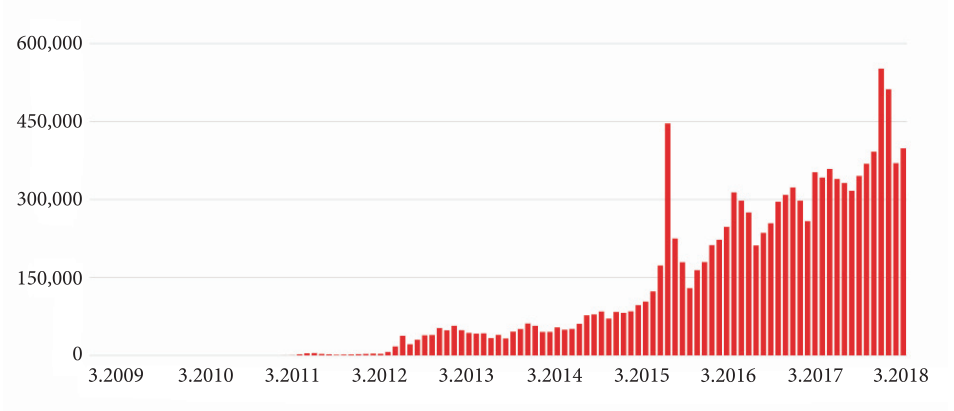
Tabela 6

Przykład transakcji n → m, wiele wejść i wiele wyjść (źródło: blochchain.info)

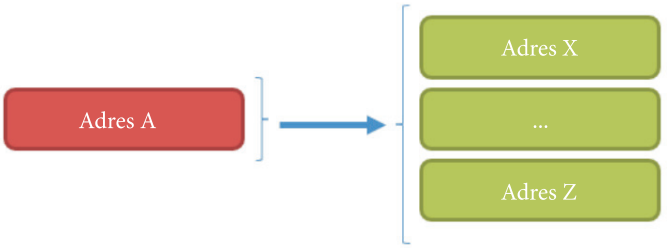
4dab9ccd4c42043080ffaf1711fe4430dad854b966c79d890059ca9d1f9e7091

19NaxUoUWEJLZXzwnahi9wjf16HoUjWB6o	→	1pKNa39SaGzjv2B3PhWTA62Pku1XUHv9d	1.683572 BTC
1BZoJ5A8DZY2wk2dkrFCYcNp3m4wHpoggM		1EWkz13EfEHFcEmu52MMiMJhcV5GfECTCE	8.364422 BTC
13phPJ4a8yjo7uYqBEqNpPCw3YF4yYRvjr		19f2WKJ2ibfG1Rzp9RNaURDiciL6PKq7xF	1.091 BTC
1CPMLny76P1Myk9B78hJBHa1t8YSWFYsGM		1MaKPepS4vPxn5n4TGpkcTwosMG5oPcwhZ	1.10550838 BTC
112ch8mPU9r4R7gtTvPUU7UgR1GR2bEUoh		12NW2eWA95a6bjuDmpbJj4LhbcwpyDScav	7.021222 BTC
12u3DDXzR4Q9m5uUrPYxpucJ4eoYq5rdDi		15BEVbxtYT8AQTZe6gbWyxeQDoEHZnyghZ	1.11068017 BTC
1pgJ4AoUrgroiqFgL8zoSiX8kuZQ4Vhsq		1JGNAwQRqmngzu2pfr2t3K3pY82s5CbTwq	1.019 BTC
1C66sXCXGa7uA2PYkiTfAZf5aG8C5CyejP		15cXikA1U6gWnoXsQGKVh98sTvQbYRcuAZ	1.144 BTC
1JYqeUrjcP3R2PiWq2cazDfjNK6JifF9zG		1MvYxmpag6xwrPKZoraKAJjvcahfnhx44	1.0884 BTC
19emG7L4Cu53229H7S44SpevtHtGhWyseT		1CoN7sh1gyCqkH9KJzVX6oFQm7wa6L6kXB	0.993 BTC
1DYjKcvEWbuSsMKVuopYvpembYcnVcstoB		1JbwR7Nu9B1rTcdppqkgum9JVSmLiakuWc	1.1471 BTC
1HmZ9HQX38edAbpkT98TdG2jVmkzEtMmpE		1JYqeUrjcP3R2PiWq2cazDfjNK6JifF9zG	0.15417011 BTC
1K8K6Agh5TX3y6wrf1g1i3WL8UAbebuEEP		1Eyj7mh4wicN3TddTvyoHH5C2CkWsRM4uc	1.07665831 BTC
1Mp5RFhhWCwjKapnQfgJreSgJ4CChHx4Xa		1KdvDgrK9rjNsWnTmSFkC5oXgAn58rGvmn	1.04840443 BTC
1GbFwTPqcdfzDku5431UWze4RKjC77NVJ7		12uK2YyVG4urGmarZEfLVYSnsFXr1Ujtgf	1.11200831 BTC
1KjJtuN4znGkd77bwcBTuWzHCZpY5WPzYx		19xUAF4u5SafZrncHMC9FBFKzMyXSehaxp	1.119581 BTC
15wsVKy11BScmrBhgNFYWwDmir7XdALfkc		1BKirpHM9xwMLrgjiUTox5BSzKujxUanFF	1.040242 BTC
1KBPLnSaLhLgSxjmDwEspjiL7ZHjFDajNH		15DMH9GpNFxbSdLxdMHngcY2xf3qYn6dch	1.0477404 BTC
1Nm7995VV54mzUp7K7VN2enNP5yVAF4BWq		1C34GFcC4Xe1rU7Q6298zP8FL3WBD3MQgU	0.9835 BTC
13SGQp1MR1otMRJi4yGMRRpQRz1mtpxUQA		13eRtad7LBwHbTJZ11XMa7sit3ZHwpRT1U	1.083597 BTC
1GYYTHDjyG8C7Dx4r9PuxEMTyTmxCd317X		1PXrwjGa5XNrFzZYhyJWkXU54AAtvjZ99a	1.13 BTC
			35.56380611 BTC

Podsumowanie		Przychody i wyjścia	
Rozmiar	4219 (Bajtów)	Razem przychodów	35.56430611 BTC
Waga	16876	Razem wychodzących	35.56380611 BTC
Czas otrzymania	2014-10-05 05:28:51	Opłaty	0.0005 BTC



Wykres 6. Liczba transakcji n → m (jedno wejście i wiele wyjść) w ujęciu miesięcznym



Rys. 7. Schemat transakcji n → m (jedno wejście i wiele wyjść)

Przed transakcją adres wejściowy A (…Vzna) dysponuje kwotą 0.77839253 BTC. Podczas transakcji środki zostają przekazane na adresy wyjściowe (...QJwa, ...Dntn, ...R4Yz, ...x9Y2), a resztę stanowi opłata transakcyjna w wysokości 0.00044276 BTC.

Tabela 7

Przykład transakcji n → m, jedno wejście i wiele wyjść (źródło: blochchain.info)

7957151ab2ac8ad51a6117c0e56baf4e8a769f93a603b9ea8e2b2f66c5b37536

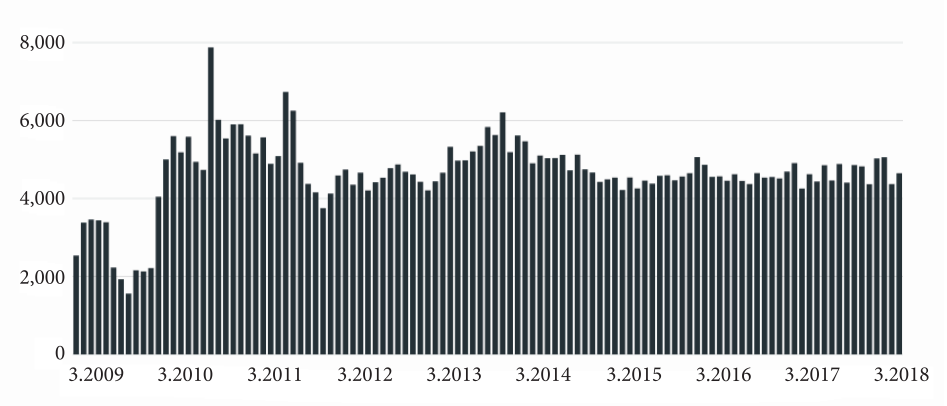
1GXerksFNbckEmQBt7fT4Vu97hRZNQVzna	1EHRiH14C7ojKAX7H3AWxmXHQxGrizQJwa	0.01 BTC
	1KuuHVAbJSnfkYf93wmUf1WJUQhFx2Dntn	0.74794977 BTC
	1Bs4F9sBNeTnV6EE9jjq3DoYn6Vq1LR4Yz	0.01 BTC
	1BBnMCFjf5KijJwRGzZdVdvrx8mxMKx9Y2	0.01 BTC
		0.77794977 BTC

Podsumowanie		Przychody i wyjścia	
Rozmiar	294 (Bajtów)	Razem przychodów	0.77839253 BTC
Waga	1176	Razem wychodzących	0.77794977 BTC
Czas otrzymania	2016-07-09 23:26:37	Opłaty	0.00044276 BTC

3.4. Transakcja 0 → 1

Transakcja posiadająca zero adresów po stronie wejściowej i jeden adres po stronie wyjściowej. Jest związana z wydobyciem nowego bloku przez górnika. Liczba występowania transakcji tego typu została przedstawiona na wykresie 7. Jej procentowy udział dla wszystkich transakcji istniejących w łańcuchu bloków do końca marca 2018 roku wynosi 0,17%.

Jest to zawsze pierwsza transakcja w nowym bloku, a na jej wysokość składają się dwie składowe. Pierwsza jest nagrodą za wydobycie danego bloku, jej kwota uzależniona jest od aktualne obowiązującego wynagrodzenia za wydobyty blok: pierwotnie było to 50 BTC, od listopada 2012 roku 25 BTC, od lipca 2016 roku 12,5 BTC itd. Drugą składową stanowi natomiast suma opłat transakcyjnych transakcji zawartych w danym bloku. Przykład takiej transakcji został zaprezentowany na rysunku 8.



Wykres 7. Liczba transakcji 0 → 1

Rys. 8. Schemat transakcji 0 → 1

Jako pierwszy przykład wybrano pierwszy blok, jaki został kiedykolwiek wydobyty. Operacja ta miała miejsce w styczniu 2009 roku i została zrealizowana przez Satoshi Nakamoto. Na adres A (1A1zP1eP5QGefi2DMPTfTL5SLmv7DivfNa) została przekazana kwota 50 BTC, co zostało zaprezentowane w tabeli 8.

Tabela 8

Przykład transakcji 0 → 1, pierwszy wydobyty blok (źródło: blochchain.info)

4a5e1e4baab89f3a32518a88c31bc87f618f76673e2cc77ab2127b7afdeda33b			
Brak wejść (nowe monety wygenerowane)	→	1A1zP1eP5QGefi... (Genesis of Bitcoin)	50 BTC
			50 BTC

Innym przykładem będzie bardziej współczesny blok, który został zaprezentowany w tabeli 9, pokazujący udział opłat transakcyjnych. Ze względu na fakt, że aktualnie większość górników wydobywa bitcoiny poprzez spółdzielnie wydobywcze, łącząc swoje moce obliczeniowe, większość transakcji tego typu trafia nie na adresy pojedynczych użytkowników, lecz właśnie na adresy spółdzielni wydobywczych. Na adres A (…59rB) została przekazana kwota 12.77401803 BTC, na którą składają się dwa czynniki: 12.5 BTC — nagroda za wydobyty blok oraz 0.27401803 BTC — suma wszystkich opłat transakcyjnych należących do danego bloku.

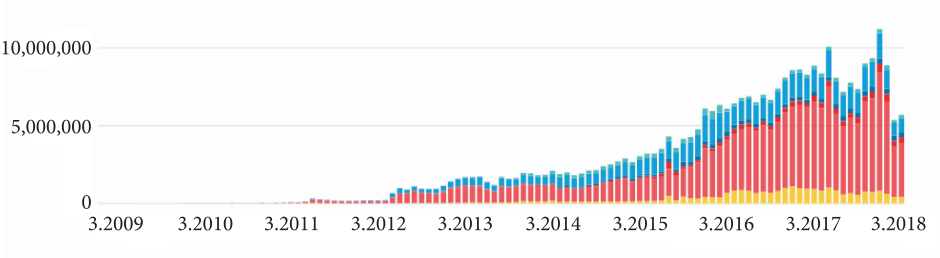
Tabela 9

Przykład transakcji 0 → 1, współczesna transakcja (źródło: blochchain.info)

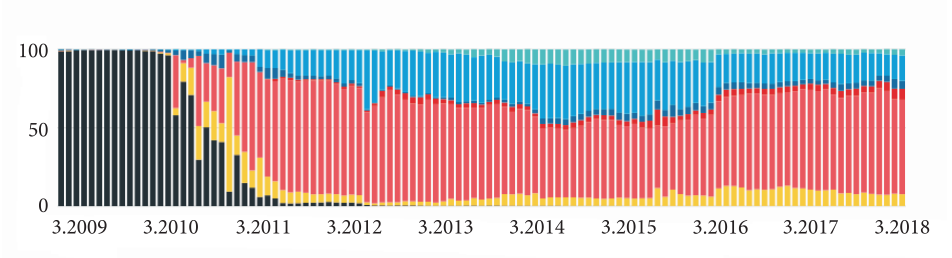
4624ef0a9fd63c40ca6f789ad34eee5f1ac92a23490596192f7a7286c51a2cd7		
Brak wejść (nowe monety wygenerowane)	1Nh7uHdvY6fNwtQtM1G5EZAFPLC33B59rB Nie można dekodować adresu wyjściowego	12.77401803 BTC 0 BTC
	9 Potwierdzenia	12.77401803 BTC

3.5. Porównanie częstości występowania transakcji

Dla uwidocznienia zmian zachodzących w czasie dotyczących typów transakcji znajdujących się w łańcuchu bloków dodatkowo zostały zaprezentowane wykresy wszystkich transakcji zarówno w ujęciu bezwzględnym (wykres 8), jak i w ujęciu procentowym (wykres 9). Kolory na wykresach odpowiadają kolorom dla poszczególnych typów transakcji zaprezentowanych na poprzednich wykresach.



Wykres 8. Typy transakcji w ujęciu miesięcznym (wartości bezwzględne)



Wykres 9. Typy transakcji w ujęciu miesięcznym (wartości procentowe)

4. Heurystyki

Znając typy przeprowadzanych transakcji, można przedstawić heurystyki[12] pozwalające ustalać, które adresy należą do poszczególnych kopalni, kantorów, giełd, a nawet pojedynczych użytkowników. Autorzy w niniejszej pracy skupiają się wyłącznie na technikach pozwalających agregować adresy w grupy należące do tego samego „identyfikowalnego" podmiotu, którym może być na przykład giełda kryptowalut, sklep internetowy akceptujący bitcoiny czy też pojedynczy użytkownik. Mając wiedzę o przynależności danego adresu do konkretnego podmiotu, zainteresowane organy (na przykład organy ścigania) mogą zwrócić się do zidentyfikowanego podmiotu z żądaniem ujawnienia danych osobowych właściciela badanego adresu. Wszystkie metody grupujące adresy w klastry (ang. *clustering*) pozwalają identyfikować dany podmiot, gdy choć jeden adres należący do danego klastra zostanie poprawnie przyporządkowany do realnego podmiotu (giełdy, sklepu itp.). W łańcuchu bloków takiej informacji jednak nie znajdziemy. Na szczęście z pomocą przychodzą nam zasoby Internetu, gdzie w wielu miejscach (fora dyskusyjne związane z kryptowalutami, strony podmiotów ujawniające ich adresy BTC w celu dokonywania darowizn itp.) możemy znaleźć informacje jednoznacznie wiążącą dany adres z konkretnym podmiotem.

Tematyka agregacji adresów poruszana była w literaturze szczególnie intensywnie w 2013 roku. Wówczas to Ron i Shamir [2] przeanalizowali łańcuch transakcji, pokazując wyniki statystyczne dla typowych zachowań użytkowników sieci Bitcoin, bazując na budowanych grafach transakcyjnych. W tym samym roku w kilku pracach [3], [4], [5], [6] pokazano heurystyki pozwalające grupować adresy tak, że z dużą dozą prawdopodobieństwa można przyjąć, że należą one do tego samego podmiotu.

4.1. Heurystyka 1

Pierwsza heurystyka, prezentowana w wyżej cytowanych pracach i nazywana tam *idioms-of-use* lub *multi-input transactions*, polega na grupowaniu adresów wejściowych wchodzących w skład pojedynczej transakcji. Jeśli dwa lub większa liczba adresów wejściowych wchodzi w skład pojedynczej transakcji, to zakładamy, że są one kontrolowane przez ten sam podmiot. Heurystykę tę zawęzimy wyłącznie do adresów wejściowych wymagających tylko jednego klucza prywatnego i zdefiniujemy ją następująco:

Heurystyka 1. *Jeśli dwa lub większa liczba adresów wejściowych wymagających użycia pojedynczego klucza prywatnego wchodzi w skład pojedynczej transakcji, to są one kontrolowane przez ten sam podmiot.*

[12] Pod pojęciem heurystyki rozumie się tutaj metodę znajdowania rozwiązań, dla której nie ma gwarancji znalezienia rozwiązania prawidłowego.

Praktyczna implementacja heurystyki 1 jest stosunkowo łatwa. Dla nowego bloku pojawiającego się w sieci Bitcoin wystarczy sprawdzić transakcje wchodzące w jego skład, a następnie dla każdej transakcji dokonać grupowania adresów wejściowych w klastry zgodnie z powyższą heurystyką. Jeśli wśród adresów wejściowych będzie adres zidentyfikowany wcześniej w systemie i przyporządkowany konkretnemu podmiotowi, to nowe adresy wejściowe także zostaną automatycznie przypisane do danego podmiotu. Do przykładowych transakcji podlegających pod niniejszą heurystykę należą transakcje zaprezentowane na rysunkach 4 (transakcja konsolidacyjna), 5 (płatność z kilku adresów użytkownika z resztą), 6 (transakcja charakterystyczna dla dużych podmiotów takich jak giełdy czy kantory).

4.2. Heurystyka 2

Druga heurystyka polega na grupowaniu jednego z adresów wyjściowych (tak zwanego adresu reszty) z adresami wejściowymi wchodzącymi w skład pojedynczej transakcji. Jeśli transakcja zawiera dwa adresy wyjściowe, to zakładamy, że jeden z nich kontrolowany jest przez ten sam podmiot co adresy wejściowe. Heurystyka ta znana jest w literaturze jako *shadow adresses* lub *change closure*.

Heurystyka 2. *Jeśli transakcja składa się z dwóch adresów wyjściowych, to jeden z nich kontrolowany jest przez ten sam podmiot co adresy wejściowe.*

Heurystyka ta opiera się dodatkowo na założeniu, że użytkownicy rzadko przekazują środki do dwóch różnych odbiorców podczas jednej pojedynczej transakcji. Główną trudnością jest tutaj poprawne zidentyfikowanie adresu reszty spośród dwóch adresów wyjściowych. Jednym z podejść jest sprawdzenie, czy jeden z adresów nigdy wcześniej nie występował w łańcuchu bloków. Jeśli tak, to można z dużym prawdopodobieństwem założyć, że jest to właśnie adres reszty, gdyż wiele portfeli bitcoinowych działa właśnie w ten sposób — tworzy nowy adres dla kwoty stanowiącej resztę transakcji. Drugim obserwowanym scenariuszem jest przekazanie znacznej kwoty (np. 100 BTC) na pewien adres należący do dużego podmiotu (np. giełdy) i systematyczne przekazywanie z tego adresu pojedynczych wypłat (najczęściej małych kwot) na konta użytkowników i przesłanie reszty (pozostałej dużej kwoty) na nowo utworzony adres danego podmiotu. Charakterystyczna jest tutaj systematyczność w wypłatach na kolejnych nowo tworzonych adresach reszt. Kolejne operacje znajdują się w kolejnych blokach łańcucha bloków lub co kilka bloków. Trzecim scenariuszem jest transakcja zawierająca kilka adresów wejściowych. Oprogramowanie obsługujące transakcje użytkowników najczęściej działa w ten sposób, że do wykonania transakcji o określonej kwocie dobiera adresy wejściowe tak, aby suma środków znajdujących się na nich była równa lub nieznacznie większa od kwoty transakcji. Czwarty scenariusz umożliwiający jednoznaczne

identyfikowanie adresów reszty, wykorzystujący niniejszą heurystykę, jest związany z wszystkimi transakcjami, które miały miejsce przed 30 stycznia 2013 roku. Do tego czasu w oficjalnym oprogramowaniu bitcoina znajdował się błąd związany z generatorem liczb pseudolosowych [7]. Błąd ten skutkował tym, że adres reszty zawsze pojawiał się jako pierwszy wśród adresów wyjściowych.

4.3. Heurystyka 3

Ostatnia heurystyka polega na identyfikowaniu właścicieli pojedynczych adresów. Informacje te można pozyskiwać z wielu źródeł: z wycieków, z istniejących serwisów internetowych poświęconych deanonimizacji (zostaną one wymienione w dalszej części artykułu), z list dyskusyjnych poświęconych kryptowalutom (użytkownicy często publikują adresy), z różnorodnych stron internetowych (część autorów, programistów, a nawet firm zachęca do oferowania im darowizn właśnie poprzez publikację własnych adresów), pozyskanych w trakcie realizacji ekspertyz na potrzeby organów procesowych.

5. Implementacja heurystyk

W rozdziale zostaną przedstawione wskazówki implementacyjne pozwalające skutecznie wdrożyć zaprezentowane heurystyki w rzeczywistych systemach.

5.1. Heurystyka 1

System wykorzystujący heurystykę 1 powinien cyklicznie, na przykład za każdym razem, gdy nowy blok zostanie dodany do łańcucha bloków, czyli średnio co 10 minut, pobierać wszystkie transakcje znajdujące się w bloku. Dla każdej z transakcji powinien analizować, czy zawiera ona co najmniej dwa adresy wejściowe rozpoczynające się od cyfry 1 (czyli te wymagające jednego klucza prywatnego do przekazania środków), a jeśli tak, to sprawdzać, czy dowolny z tych adresów został już wcześniej zidentyfikowany (przypisany do konkretnego podmiotu). Jeśli dowolny z adresów wejściowych był już wcześniej zidentyfikowany, to pozostałe adresy wejściowe danej transakcji można przypisać do tego samego podmiotu. Wiele prac skupiało się na grupowaniu adresów z wykorzystaniem niniejszej heurystyki [3], [4], [5], [6]. W naszym podejściu, dopóki danego adresu nie da się przyporządkować do konkretnego podmiotu, to ich grupowanie jest bezcelowe i nie wnosi wartości dodanej w procesie deanonimizacji.

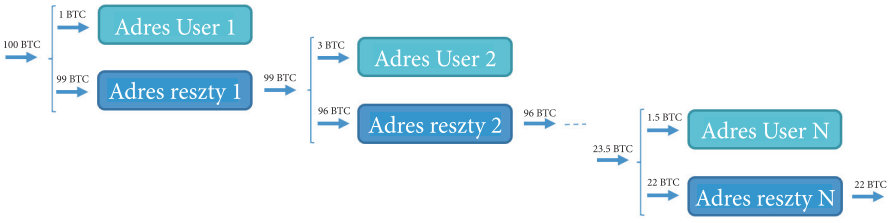
Drugie wykorzystanie tej heurystyki w systemie powinno następować w momencie, gdy uda się jednoznacznie przyporządkować nowy adres do konkretnego podmiotu, na przykład za pomocą heurystyki 3. Wówczas system powinien przeanalizować wszystkie transakcje występujące w łańcuchu bloków zawierające ten

adres po stronie wejściowej. W przypadku znalezienia transakcji z wieloma adresami wejściowymi przyporządkować je do konkretnego podmiotu, a następnie wykonać powyższą procedurę rekurencyjnie dla wszystkich nowo przyporządkowanych adresów. Implementacja heurystyki 1 może być realizowana w systemie w sposób całkowicie zautomatyzowany na przykład przy użyciu CRON-a[13].

5.2. Heurystyka 2

Implementacja heurystyki 2 nie jest trywialna do zautomatyzowania. Głównym problemem jest tutaj poprawne zidentyfikowanie adresu reszty. Wspomniane wcześniej podejście polegające na sprawdzeniu, czy jeden z adresów nigdy wcześniej nie występował w łańcuchu bloków, może być pomocne, ale nie jest gwarantem poprawności i należy do niego podchodzić ostrożnie.

Próbując jednak zaimplementować tę heurystykę w rzeczywistym systemie, tak by była wykonywana automatycznie, możemy zastosować dwa podejścia. Pierwsze związane jest z dużymi podmiotami. Jeśli w danej transakcji adres/y wejściowe należą do jednego podmiotu przyporządkowanego na przykład do kategorii „giełda kryptowalut" oraz w łańcuchu bloków uda się zidentyfikować ciąg transakcji o strukturze podobnej do tej zaprezentowanej na rysunku 9, wówczas kolejne adresy reszt można przyporządkować do tego samego podmiotu[14].

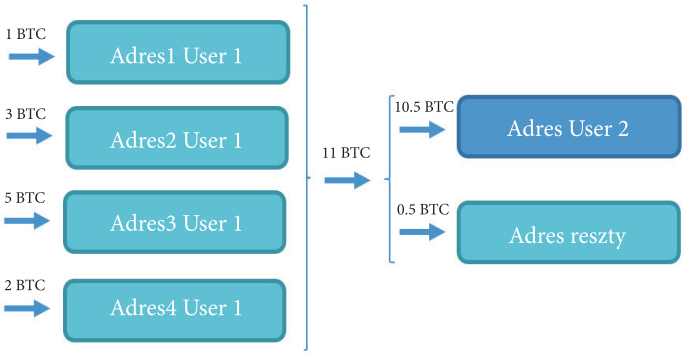


Rys. 9. Łańcuch transakcji charakterystyczny dla określonej grupy podmiotów umożliwiający automatyczne wykorzystanie heurystyki 2

[13] CRON — program do harmonogramowania zadań (programów, komend, skryptów), umożliwia ich uruchamianie cyklicznie (zgodnie z określonym interwałem, na przykład co 10 minut) lub o określonej porze.

[14] Zauważyć należy, że przedstawiony ciąg transakcji powinien charakteryzować się kilkoma cechami: a) pierwszy adres wejściowy powinien dysponować stosunkowo dużą liczbą BTC (jest to tak zwany *hot-wallet* służący danemu podmiotowi do realizacji kolejnych wypłat); b) kolejne transakcje powinny występować z dużą regularnością, na przykład co jeden blok w łańcuchu bloków; c) kwota przekazywana na adres reszty jest najczęściej znacznie większa od kwoty przekazywanej użytkownikom.

Drugie podejście bazuje na transakcjach z kilkoma adresami wejściowymi. Jeżeli dla wszystkich adresów wejściowych danej transakcji wysokość środków znajdujących się na każdym z nich jest większa od środków przekazanych na jeden z adresów wyjściowych, to adres ten stanowi adres reszty. Uzasadnienie tego opiera się na zasadzie działania wielu portfeli. Jeśli w danym portfelu znajduje się wiele adresów, na których znajdują się pewne środki, i użytkownik chce dokonać płatności w wysokości przekraczającej zasobność pojedynczego adresu, wówczas portfel tak dobiera adresy wejściowe do danej transakcji, aby suma środków znajdujących się na nich była równa kwocie samej transakcji (co się rzadko zdarza) lub niewiele ją przekraczała.



Rys. 10. Schemat transakcji zawierającej kilka adresów wejściowych umożliwiający automatyczne wykorzystanie heurystyki 2

Trzecie podejście do praktycznej implementacji bazuje na opisanym wcześniej błędzie w oficjalnym kliencie bitcoina. Dla wszystkich historycznych transakcji mających dwa adresy wyjściowe, odbywających się przed 3 lutego 2013 r. (wówczas została opublikowana poprawiona wersja klienta bitcoin), adres reszty znajduje się zawsze na pierwszej pozycji.

Dodać jednak należy, że implementacja tej heurystyki wymaga kompromisu pomiędzy jej skutecznością (rozumianą jako liczba zidentyfikowanych adresów reszty) a poprawnością (liczba poprawnie zidentyfikowanych adresów reszty). W systemach tworzonych na potrzeby ekspertyz procesowych poprawność ma priorytetowe znaczenie i dla wielu zidentyfikowanych łańcuchów transakcji przedstawionych na rysunku 9 dość szybko traci się pewność co do zidentyfikowania danego adresu jako adresu reszty i pozostała część łańcucha zostaje odrzucona.

5.3. Heurystyka 3

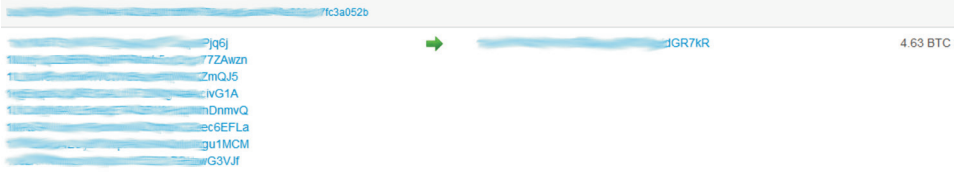
Ostatnia z prezentowanych heurystyk jest niemal niemożliwa do zautomatyzowania. Znalezienie w sieci Internet konkretnego adresu wymaga najczęściej „wczytania się" w treść danego komentarza, postu, strony internetowej. Zautomatyzowaniu może podlegać proces cyklicznego wyszukiwania nowo występujących adresów w łańcuchu bloków pod kątem ich istnienia w sieci Internet. Odpowiada to stworzeniu robota indeksującego (ang. *web-crawler*) zbierającego informacje o konkretnym adresie na stronach (najczęściej poświęconych kryptowalutom) znajdujących się w sieci Internet. Jednak samo przypisanie zidentyfikowanego adresu do rzeczywistego podmiotu wymaga już czynnika ludzkiego. Podobnie w przypadku adresów i podmiotów nimi zarządzających otrzymywanych od organów procesowych. Mimo że nie mamy tutaj problemu wiarygodności pozyskanych informacji, to jednak i tak należy „ręcznie" przyporządkowywać je do konkretnych podmiotów czy osób.

6. Studium przypadków

W tym punkcie przedstawionych zostanie kilka rzeczywistych scenariuszy, które zostały zidentyfikowane podczas przeprowadzanych przez autorów ekspertyz. Pierwszy z nich polega na wykorzystaniu heurystyki 1. Na podstawie zleconej ekspertyzy wiadomo było, że kilka adresów (m.in.: …jq6j, …Awzn, …mQJ5) należy do pewnego podmiotu (w tym przypadku konkretnej osoby). Dzięki znalezieniu transakcji konsolidacyjnej przedstawionej w tabeli 10 zidentyfikowano inne adresy należące do tego samego podmiotu (…EFLa, …3VJf), które następnie zostały poddane dalszej analizie śledczej.

Tabela 10

Przykład rzeczywistego użycia heurystyki 1 — transakcja konsolidacyjna
(źródło: blochchain.info)

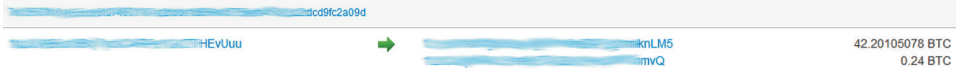


Drugi scenariusz pokazuje praktyczne wykorzystanie heurystyki 2. Na podstawie zleconej ekspertyzy wiadomo było, że adres …vUuu należy do pewnego dużego podmiotu. Na podstawie heurystyki 2 wywnioskowano, że adres …nLM5 także należy do tego samego podmiotu (stanowi adres reszty). Wniosek ten został

wyciągnięty na podstawie trzech faktów: wysokości środków przekazywanych pomiędzy adresami, tego, że adres reszty stanowi nowy adres — niewystępujący wcześniej w łańcuchu bloków, a kolejna transakcja dla adresu reszty jest podobna do tej zaprezentowanej w tabeli 11 i występuje w bloku o dwa numery większym (potwierdzonym jakieś 20 minut później).

Tabela 11

Przykład rzeczywistego użycia heurystyki 2 (źródło: blochchain.info)



Trzeci scenariusz pokazuje wykorzystanie wiadomości znajdujących się w sieci Internet umożliwiających przyporządkowanie konkretnego adresu do podmiotu. W wyniku działania robota indeksującego przeszukującego serwis bitcointalk.org na jednej z podstron [11] zidentyfikowano adres (…WG7q), który można przyporządkować do (nieistniejącej już) giełdy MtGox. Oczywiście znacznie bardziej wartościowe i wiarygodne są dane pozyskiwane od organów procesowych przy wykonywaniu ekspertyz sądowych, jednak ich liczba nie pozwala na wyeliminowanie zasobów Internetu jako źródła danych.

7. Uwarunkowania prawne

O ile przed 2018 r. w Polsce nie istniały mechanizmy zabraniające podmiotom zajmującym się obrotem kryptowalutami (giełdy, kantory) przeprowadzania transakcji pomiędzy anonimowymi użytkownikami, o tyle aktualnie każdy taki podmiot podlega pod regulacje stosownych ustaw [12], [13] i jego użytkownicy są identyfikowalni[15]. W szczególności ustawa o przeciwdziałaniu praniu pieniędzy oraz finansowaniu terroryzmu [13] definiuje waluty wirtualne i rozszerza zakres „instytucji obowiązanych" o podmioty świadczące usługi w zakresie wymiany pomiędzy walutami wirtualnymi i środkami płatniczymi czy też wymiany pomiędzy walutami wirtualnymi. Podobna sytuacja ma miejsce w wielu innych krajach. Na przykład w USA podmioty te podlegają pod regulacje AML[16]/KYC[17] nakazujące

[15] Aktualnie założenie konta w systemach zajmujących się obrotem kryptowalutami (giełdy, kantory) wymaga potwierdzenia zarówno tożsamości (np. dowód osobisty, paszport, prawo jazdy), jak i adresu (np. rachunek za prąd, wodę, gaz; pismo urzędowe z ZUS lub US).

[16] AML (ang. *anti-money laundering*).

[17] KYC (ang. *know your customer*).

podmiotom zajmującym się obrotem walutami wirtualnymi weryfikację i przechowywanie danych identyfikujących ich użytkowników. Umożliwia to jednoznaczną identyfikację właścicieli adresów.

Z punktu widzenia organów ścigania wyzwaniem jest właśnie identyfikacja podmiotu, do którego należy dany adres i do którego należy się zwrócić w celu otrzymania danych osobowych właściciela adresu. W przypadkach gdy znany jest numer telefonu, a organy chcą ustalić tożsamość abonenta polskich numerów telefonów komórkowych lub gdy organy dysponują numerem IMEI telefonu i chcą ustalić, z jakim numerem karty SIM współpracował on w danym okresie, mogą wysłać postanowienia o żądaniu udostępnienia danych telekomunikacyjnych do wszystkich operatorów funkcjonujących na krajowym rynku. Podobna sytuacja może mieć miejsce w przypadku poszukiwania krajowych kont bankowych konkretnego podejrzanego poprzez wysłanie postanowień o zwolnieniu z tajemnicy bankowej do wszystkich możliwych banków w Polsce. Sytuacje takie są powszechne i często stosowane w toku prowadzonych postępowań przygotowawczych. W przypadku walut wirtualnych sytuacja znacznie się komplikuje. Sama znajomość adresu (konkretnej waluty, na przykład bitcoina) nie przybliża bezpośrednio prowadzącego postępowanie do zidentyfikowania jego posiadacza. Adres może należeć do dużego podmiotu (giełdy, kantoru) mającego swoją siedzibę w dowolnej jurysdykcji (krajowej lub zagranicznej), może też być adresem wygenerowanym przez samego użytkownika. Wysłanie postanowień o zwolnieniu z tajemnicy do wszystkich możliwych podmiotów zajmujących się obrotem kryptowalutami, których są setki na światowym rynku, jest praktycznie niemal niemożliwe do zrealizowania. Stąd istnienie systemów informatycznych wspomagających proces deanonimizacji poprzez wskazanie, czy dany adres należy do jakiegoś podmiotu, w znaczący usprawniałoby i przyspieszało tok postępowania.

8. Przegląd narzędzi wspomagających deanonimizację

Realna potrzeba deanonimizacji użytkowników sieci Bitcoin, w połączeniu z wymienionymi wyżej heurystykami, spowodowała powstanie kilku narzędzi wspomagających ten proces. Do projektów opartych na licencji wolnego oprogramowania zaliczyć można: nierozwijany już *blockparser* [14], czy *bitcoin-deanonymization* [15] lub *address-reuse-tracker* [16]. Innym wartościowym projektem jest *walletexplorer* [17] dostarczający informację o zidentyfikowanych dużych podmiotach i ich adresach oraz *bitconeview* [18], [19] wizualizujący przepływy bitcoinów, zaczynając od zdefiniowanej przez użytkownika transakcji aż do określonego przez niego czasu. Do komercyjnych projektów, do których autorom udało się uzyskać tymczasowy dostęp i je przetestować, należą: *Chainanalysis* [20], [21], *Scorechain* [22] oraz *Qlue* [23]. Wspomnieć jeszcze należy o systemach *Elliptic* [24], *Skry* [25] czy *Numisight* [26], do których jednak autorom nie udało się uzyskać dostępu.

Autorzy, wychodząc naprzeciw potrzebom polskich organów wymiaru sprawiedliwości i organów ścigania, zaprojektowali prototyp autorskiego systemu [27], opierając się dodatkowo na najlepszych praktykach dostarczanych przez Europol [28]. Stworzony prototyp składa się koncepcyjnie z kilku modułów funkcjonalnych, do najważniejszych z nich należą: moduł przeszukujący zasoby sieci Internet w celu wyszukiwania adresów i przyporządkowywania ich do konkretnych podmiotów oraz moduł agregujący adresy za pomocą przedstawionych w niniejszej pracy heurystyk. Aktualnie, dzięki pozyskaniu rzeczywistych adresów wraz z odpowiadającymi im podmiotami z kilku źródeł (między innymi z: ekspertyz realizowanych przez autorów na zlecenia organów procesowych; licznych źródeł internetowych, takich jak wspomniane już [17] czy [29], [30]; oraz innych źródeł, które nie mogą być ujawnione), system umożliwia bezpośrednią deanonimizację, rozumianą jako przyporządkowanie adresu do konkretnego podmiotu, dla ponad 100 mln adresów[18].

9. Podsumowanie

Satoshi Nakamoto w swojej pracy [1] wskazał istnienie heurystyki (oznaczanej w niniejszej pracy jako heurystyka 1), która w połączeniu z ujawnieniem właściciela adresu powoduje deanonimizację także innych transakcji do niego należących. Pozostałe heurystyki zaprezentowane w niniejszej pracy, mimo że nie są idealne, to jednak przy poprawnej implementacji znacznie ułatwiają proces deanonimizacji. Istnienie skutecznych narzędzi na świecie i ich brak na krajowym rynku zachęcił autorów do podjęcia próby tworzenia autorskiego rozwiązania ukierunkowanego na potrzeby krajowych organów procesowych. Powyższe podejście deanonimizacyjne można zastosować także do wielu innych kryptowalut, zwanych potocznie altcoinami[19], charakteryzujących się pseudoanonimowością (np. ethereum). Natomiast w przypadku altcoinów gwarantujących anonimowość w swojej konstrukcji (np. Zcash, Monero) zaprezentowane w pracy heurystyki (1 i 2) nie znajdą już oczywiście zastosowania.

Źródło finansowania pracy — środki własne autorów.

Artykuł wpłynął do redakcji 18.04.2018 r. Zweryfikowaną wersję po recenzjach otrzymano 9.11.2018 r.

Przemysław Rodwald https://orcid.org/0000-0003-4261-8688
Witold Sobolewski https://orcid.org/0000-0003-4606-9604
Maja Rodwald https://orcid.org/0000-0002-7064-8879

[18] Sumaryczna liczba unikalnych adresów BTC może zostać oszacowana na przykład w serwisie https://blockchair.com/bitcoin/outputs i wynosi ponad 900 mln.

[19] Altcoin — potoczne określenie wszystkich pozostałych kryptowalut poza bitcoinem; nazwa pochodzi od angielskiego zwrotu alternative coins — waluty alternatywne.

LITERATURA

[1] Nakamoto S., *Bitcoin: A peer-to-peer electronic cash system*, 2008, https://bitcoin.org/bitcoin.pdf [dostęp 31.03.2018].

[2] Ron D., Shamir A., *Quantitative analysis of the full bitcoin transaction graph*, International Conference on Financial Cryptography and Data Security, Springer, Berlin, Heidelberg, 2013, s. 6-24.

[3] Androulaki E., Karame G.O., Roeschlin M., Scherer T., Capkun S., *Evaluating User Privacy in Bitcoin*, Financial Cryptography and Data Security, FC 2013, Lecture Notes in Computer Science, vol. 7859, Springer, Berlin, Heidelberg, 2013, s. 34-51.

[4] Meiklejohn S., Pomarole M., Jordan G., Levchenko K., McCoy D., Voelker G.M., Savage S., *A fistful of bitcoins: characterizing payments among men with no names*, Proceedings of the 2013 conference on Internet measurement conference, ACM, 2013, s. 127-140.

[5] Reid F., Harrigan M., *An analysis of anonymity in the bitcoin system*, Security and privacy in social networks, Springer, New York, 2013, s. 197-223.

[6] Ortega M.S., *The bitcoin transaction graph anonymity*. *Master's thesis*, Universitat Oberta de Catalunya, 2013.

[7] Spagnuolo M., Maggi F., Zanero S., *BitIodine: Extracting Intelligence from the Bitcoin Network*, Financial Cryptography and Data Security, FC 2014, Lecture Notes in Computer Science, vol. 8437, Springer, Berlin, Heidelberg, 2014, s. 457-468.

[8] Fleder M., Kester M.S., Pillai S., *Bitcoin transaction graph analysis*, arXiv preprint arXiv:1502.01657, 2015.

[9] Conti M., Lal C., Ruj S., *A Survey on Security and Privacy Issues of Bitcoin*, arXiv preprint arXiv:1706.00916, 2017.

[10] Akcora C.G., Gel Y.R., Kantarcioglu M., *Blockchain: A Graph Primer*, arXiv preprint arXiv:1708.08749, 2017.

[11] https://bitcointalk.org/index.php?topic=481949.msg5305138 [dostęp 31.03.2018].

[12] Ustawa z dnia 19 sierpnia 2011 r. o usługach płatniczych, Dz.U. 2011 nr 199, poz. 1175.

[13] Ustawa z dnia 1 marca 2018 r. o przeciwdziałaniu praniu pieniędzy oraz finansowaniu terroryzmu, Dz.U. 2018 poz. 723.

[14] https://github.com/znort987/blockparser [dostęp 31.03.2018].

[15] https://github.com/gfanti/bitcoin-deanonymization [dostęp 31.03.2018].

[16] https://github.com/kristovatlas/address-reuse-tracker [dostęp 31.03.2018].

[17] https://www.walletexplorer.com [dostęp 31.03.2018].

[18] http://www.bitconeview.info [dostęp 31.03.2018].

[19] Di Battista G., Di Donato V., Patrignani M., Pizzonia M., Roselli V., Tamassia R., *BitConeView: Visualization of Flows in the Bitcoin Transaction Graph*, IEEE Symposium on Visualization for Cyber Security (VizSec 2015), 2015.

[20] https://www.chainanalysis.com [dostęp 31.03.2018].

[21] Chainalysis Inc., *Chainalysis Reactor guide, version 3.5*, 2017.

[22] https://bitcoin.scorechain.com [dostęp 31.03.2018].

[23] https://qlue.io [dostęp 31.03.2018].

[24] https://www.elliptic.co [dostęp 31.03.2018].

[25] https://skry.tech [dostęp 31.03.2018].

[26] http://numisight.com [dostęp 31.03.2018].

[27] https://www.sydeus.pl/index.php [dostęp 31.03.2018].
[28] Europol, *A guide for bitcoin investigators,* version 1.06, 2017.
[29] https://www.blocktrail.com [dostęp 31.03.2018].
[30] http://blockchain.exposed [dostęp 31.03.2018].

P. RODWALD, W. SOBOLEWSKI, M. RODWALD

Deanonymization of bitcoin cryptocurrency users

Abstract. The aim of this article is to show how one can deanonymize users of cryptocurrencies. To this end the most popular of the cryptocurrencies, i.e. bitcoin is used as an example. At the beginning, the basic concepts about cryptocurrencies are presented. Afterwards, our approach to systematize the types of transactions existing in the blockchain is proposed. This part is enriched with the graphs showing their quantitative occurrence in the blockchain. The main part of this article presents the heuristics use to deanonymize users. A few practical pieces of advice for implementation of the presented heuristics in the real deanonymizing system are included. Then the real case studies are introduced. They are supported with comments based on the experience from court trials carried out by the authors. The final part contains legal regulations and existing tools supporting the deanonymizing process.
Keywords: deanonymization, cryptocurrency, bitcoin
DOI: 10.5604/01.3001.0013.1466