SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Tomislav Babić

Detekcija *peeling chainova* u sustavu kriptovalute *Bitcoin*

Zagreb, 2022.

Ovaj rad izrađen je na Zavodu za elektroniku, mikroelektroniku, računalne i inteligentne sustave pod vodstvom doc. dr. sc. Ante Đereka i predan je na natječaj za dodjelu Rektorove nagrade u akademskoj godini 2021./2022.

Sadržaj

1. Uvod

Kriptovalute se u općoj populaciji smatraju anonimnim načinom razmjene sredstava. Zbog toga, uz naravno posve legalno korištenje, ljudi ih koriste i za ilegalne aktivnosti. *Bitcoin* je najstarija i najpoznatija kriptovaluta. Kako je *Bitcoinov* lanac blokova javan, moguće je pratiti sve *Bitcoin* transakcije. Ipak, kako *Bitcoin* adrese nisu povezane sa stvarnim imenima to korisnicima pruža određenu razinu anonimnosti, točnije to zovemo pseudonimnost. Prednost pseudonimnosti *Bitcoina* je dvojaka. Korisnicima to omogućava korištenje *Bitcoina* bez odavanja stvarnog identiteta. Ali, u slučaju krađe, i zlonamjerni korisnici profitiraju od pseudonimnosti. Postoje razne heuristike [1] za grupiranje *Bitcoin* adresa koje pripadaju istim osobama, ali postoje i razne tehnike kojima se to grupiranje otežava. Jedna od tehnika za skrivanje traga je obrazac ponašanja nazvan *peeling chain*. Korištenjem *peeling chaina* otežava se otkrivanje povezanosti između različitih adresa. Time zlonamjerni korisnici, primjerice kada netko ukrade *bitcoine*, mogu sakriti vezu između adresa za koje se zna da su izravno povezane s krađom i drugih adresa na koje dalje raspodjeljuju ukradene *bitcoine*. To im omogućava da raspolažu s ukradenim *bitcoinima*.

Postojanje algoritma za detekciju *peeling chainova* bi zlonamjernim korisnicima otežalo skrivanje traga. Burze bi tada mogle odbiti zamjenu tih *bitcoina* za valute iz stvarnog svijeta. Time bi bila smanjena motivacija zlonamjernih korisnika za krađom.

Naravno, samo postojanje *peeling chaina* ne znači da se radi o ilegalnim aktivnostima. Ali, kako će biti pojašnjeno kasnije, upitna je motivacija dobronamjernih korisnika za korištenje *peeling chaina*.

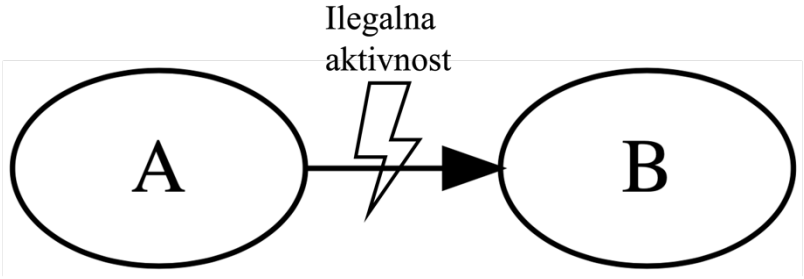
U ovom radu opisan je algoritam za detekciju *peeling chainova*. Algoritam je pokrenut na skupu svih *bitcoin* transakcija iz svibnja 2020. Algoritam pronalazi poznati slučaj korištenja *peeling chain* nakon krađe s burze Bitfinex koji se odvio u tom razdoblju. Također, algoritam pronalazi nove, do sada neotkrivene, *peeling chainove*.

2. Motivacija

2.1. Zašto skrivati trag na *Bitcoinu*?

Podatci o svim *Bitcoin* transakcijama su javno dostupni, ali ne pod stvarnim imenima nego pod *Bitcoin* adresama [2]. Zbog toga *Bitcoin* nije potpuno anoniman, on je pseudoniman [3]. To znači da je na *Bitcoinu* moguće pratiti kretanje sredstava s jedne adrese na drugu, kao što je moguće i kod računa u bankama. Zbog toga neki korisnici, bilo iz dobronamjernih ili zlonamjernih razloga, imaju potrebu za skrivanjem traga.

Pretpostavimo da je poznato da se na adresi B nalaze sredstva do kojih je zlonamjerni korisnik došao ilegalnim putem. Postoje razne situacije u kojima se to može dogoditi, zlonamjerni korisnik može nekome ukrasti *bitcoine* s adrese A i prebaciti ih na svoju adresu B [4], zlonamjerni korisnik je do *bitcoina* došao iznudom (gdje mu osoba od koje se iznuđuje s adrese A prebacuje sredstva) [5], korisnik je do *bitcoina* došao prodajom ilegalne robe poput droge [6] i mnoge druge situacije. Ono što je zajedničko svim tim situacijama je da su s adrese A zbog neke ilegalne aktivnosti sredstva prebačena na adresu B. Grafički prikaz transakcije gdje su sredstva prebačena s adrese A na adresu B zlonamjernog korisnika uslijed ilegalne aktivnosti nalazi se na Sl. 2.1.



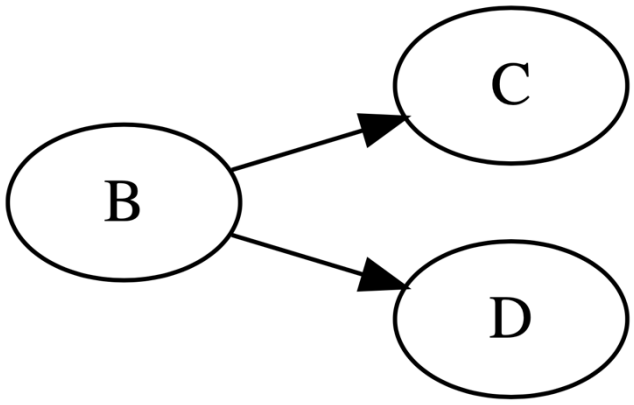
Sl. 2.1 Transakcija s adrese A na adresu B za ilegalnu aktivnost

U svim tim situacijama, bilo krađa, iznuda ili nešto treće, zlonamjerni korisnik mora negdje javno obznaniti svoju adresu B na koju će mu biti izvršena uplata (ili u slučaju krađe na javnom *blockchainu* vidi se na koju adresu su prebačena ukradena sredstva). Kad bi zlonamjerni korisnik sredstva direktno s adrese B pokušao prebaciti na burzu za isplatu

stvarne valute, burza bi vjerojatno odbila napraviti isplatu. Također, policija bi mogla od burze zatražiti informacije o korisniku koji je pokušao prebaciti sredstva s adrese B na burzu i tako mu lakše ući u trag. Zbog toga se javlja potreba za skrivanjem traga na *Bitcoinu*.

2.2. Kako skrivati trag na *Bitcoinu*?

U situaciji opisanoj u poglavlju 2.1 javno je poznato da se na adresi B nalaze sredstva do kojih je zlonamjerni korisnik došao ilegalnim putem. Ono što zlonamjerni korisnik tada može napraviti je stvoriti dvije nove adrese C i D i na njih raspodijeliti sredstva s adrese B kao što je prikazano na Sl. 2.2.



Sl. 2.2 Raspodjela sredstava s adrese B na adrese C i D

Sada se više ne može sa sigurnošću znati da zlonamjerni korisnik upravlja adresama C i D. Moguće je da je zlonamjerni korisnik trećoj osobi u zamjenu za neku robu ili uslugu uplati *bitcoine* na adresu C, a da je D njegova *change address* (adresa korisnika koji je platitelj na koju sebi isplaćuje ostatak prilikom transakcije [3]). Treća osoba moža nije znala da se radi o sredstvima koja su stekunta ilegalnim putem, a možda je i znala ali ju to nije smetalo. Moguće je i da je obrnuto, da je C *change address* zlonamjernog korsnika, a da adresa D pripada trećoj osobi. Također, moguće je i da obje adrese pripadaju ili trećoj osobi ili zlonamjernom korisniku. Dakle, više se ne zna sa sigurnošću da adresama C i D upravlja zlonamjerni korisnik i da su na njima ilegalna sredstva. Kada bi burze odbile razmjenu sredstava s adresa C i D za stvarne valute riskirale bi da odbiju isplatu dobronamjernom

korisniku koji legitimno raspolaže s tim sredstvima. Ali, s obzirom na malu udaljenost adresa C i D od adrese B (udaljenost je 1 transakcija) burze mogu odbijati isplate iz predostrožnosti. Ono što zlonamjerni korisnici tada rade je daljnji nastavak grananja s adresa C i D. Sa svakim grananje povećava se udaljenost (u broju transakcija) između ilegalno stečenih sredstava i krajnjih adresa. Taj obraza uporabe naziva se *peeling chain* i detaljnije je opisan u poglavlju 5.

3. Peeling chain

Peeling chain je obrazac uporabe prisutan kod kriptovalute *Bitcoin*. *Peeling chain* se koristi u različite svrhe, ne samo za kriminalne aktivnosti. Početak *peeling chaina* je jedna adresa na kojoj se nalaz veći iznos *bitcoina*. Taj iznos se tada dijeli na dvije nove adrese na dva načina, ovisno o vrsti *peeling chaina*. Navedeni postupa se ponavlja veliki broj puta čime nastaje *peeling chain*. Za svaku transakciju u *peeling chainu* koriste se nove, do sada neiskorištene adrese. Svaka od tih adresa imati će ukupno dvije transakcije: jednu u kojoj prima sredstva i jednu za daljnje grananje sredstava [7].

Peeling chain skriva trag time što povećava udaljenost između početne adrese i krajnjih adresa na koje su raspodijeljeni *bitcoini*. Što je veća udaljenost između dvije adrese to je teže pronaći poveznicu između njih u *Bitcoin* grafu transakcija. Također, sa svakim grananjem je manje vjerojatno da sredstva pripadaju istom korisniku.

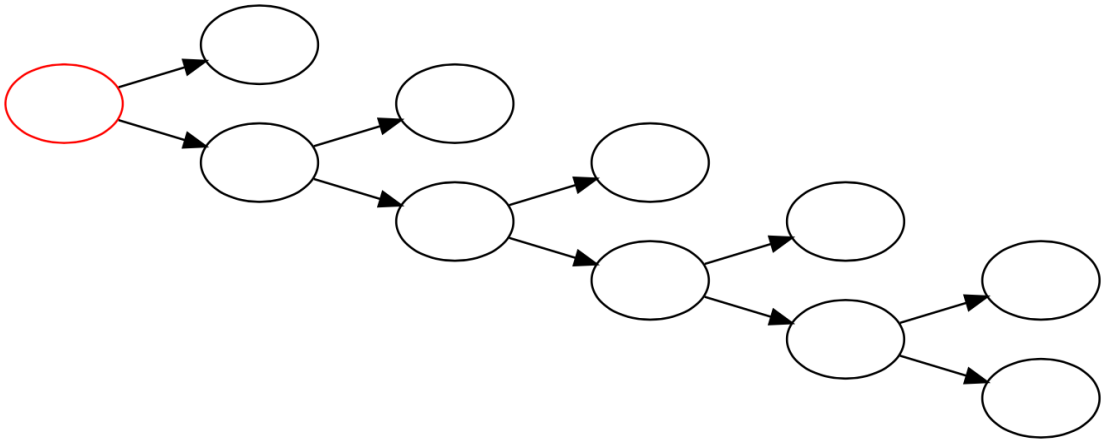
Naravno, samo postojanje *peeling chaina* ne znači da se radi o ilegalnim aktivnostima. Svatko može odlučiti svoje *bitcoine* razdijeliti na više različitih adresa korištenjem *peeling chaina*. Međutim, za svaku transakciju korisnik mora platiti naknadu. Zbog toga je nakon provođenja *peeling chaina* suma iznosa na krajnjim adresama manja od početnog iznosa *peeling chaina*. Zlonamjernim korisnicima to ne predstavlja problem, ako ne sakriju trag vjerojatno neće ni moći unovčiti svoje *bitcoine* pa će im oni zapravo biti bezvrijedni. Ali dobronamjernim korisnicima skrivanje traga korištenjem *peeling chaina* ne donosi direktnu korist, a smanjuje im se ukupni iznos kojim raspolažu. Zbog toga je upitna motivacija korisnika koji koriste *peeling chain* i može se pretpostaviti da, pogotovo kada se radi o većim *peeling chainovima*, se radi o zlonamjernim korisnicima.

3.1. Vrste *peeling chainova*

Postoje dvije vrste *peeling chainova*: linearni *peeling chain* i *peeling chain* stabla. Moguće je i kombinirati te dvije vrste.

3.1.1. Linearni *peeling chain*

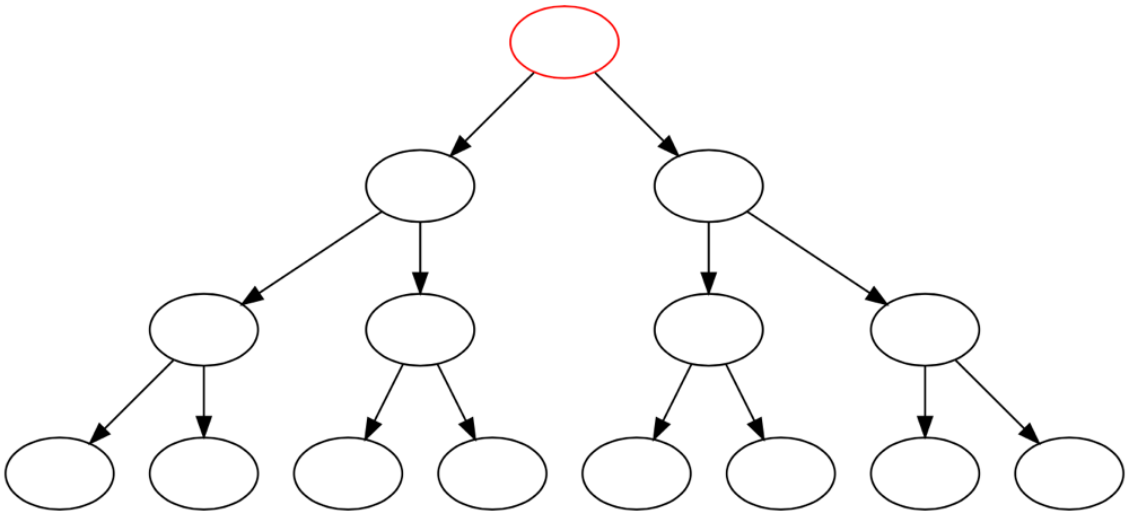
Kod linearnog *peeling chaina* s početne adrese se manji iznos prebacuje na jednu novu adresu, a ostatak se prebacuje na drugu adresu. Grananje se dalje nastavlja s adrese koja je primila većinu sredstava. Na Sl. 3.1 prikazan je jedan linearni *peeling chain*. Čvorovi grafa su adrese, a bridovi transakcije (jedna transakcija može biti prikazana s više bridova, za svaku kombinaciju ulaza i izlaza transakcije prikazan je jedan brid).



Sl. 3.1 Linearni *peeling chain*, crveni čvor je početak *peeling chain*

3.1.2. *Peeling chain* stablo

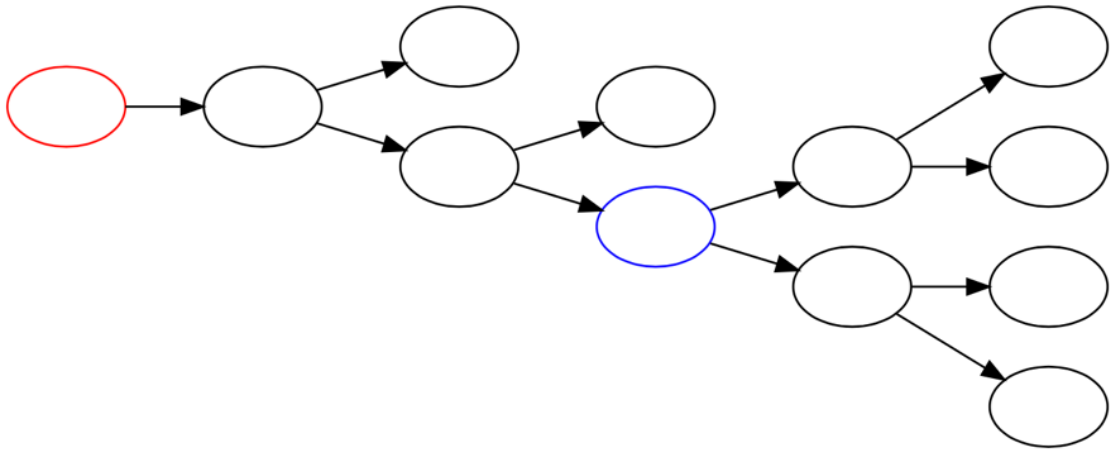
Kod *peeling chain* stabla iznosi se s početne adrese također granaju na dvije adrese. Razlika u odnosu na linearni *peeling chain* je što se daljnje grananje nastavlja iz obje nove adrese. Kako se grananje dalje nastavlja iz obje adrese, iznosi koji se s početne adrese prebacuju na dvije nove adrese su približno jednaki. Na Sl. 3.2 prikazano je *peeling chain* stablo.



Sl. 3.2 *Peeling chain* stablo, crveni čvor je početak *peeling chain*

3.1.3. Kombinacija stabla i linearnog *peeling chain*

Moguće je kombinirati linearni *peeling chain* i *peeling chain* stablo. U bilo kojem čvoru *peeling chaina* moguće je prijeći iz jedne vrste u drugu. Na Sl. 3.3 prikazan je *peeling chain* koji počinje kao linearni te nakon nekoliko grananja prelazi u *peeling chain* stablo.



Sl. 3.3 Kombinacija linearnog *peeling chaina* i *peeling chain* stabla, crveni čvor je početak linearnog *peeling chaina*, *plavi čvor* je adresa u kojoj se prelazi iz linearnog u *peeling chain* stablo

3.2. *Peeling chain* u praksi

Jedan poznati slučaj u kojemu su kradljivci iskoristili *peeling chain* za skrivanje traga je krađa *bitcoina* s burze Bitfinex [8]. U kolovozu 2016. burzi Bitfinex ukradeno je 119 754 *bitcoina* [4] koji su u to vrijeme vrijedili približno 70 milijuna $ [9]. Ukradeni *bitcoini* su prebačeni na više adresa. Neke od njih su bile početak *peeling chaina* [8].

3.2.1. Bitfinex krađa - *peeling chain* 1

Počevši od adrese `3CA1UDYQy47Z46HKCVqRV8b1XVduocWAcW` proveden je *peeling chain*. Taj *peeling chain* započeo je s 30,6675418 BTC

3.2.2. Bitfinex krađa – *peeling chain* 2

Počevši od adrese `32Ev5wv1hwcM1xiYYfQ8P8kFGVXg9jmfQ3` proveden je *peeling chain*. Taj *peeling chain* započeo je s 271,22718210 BTC.

4. Algoritam za detekciju *peeling chaina*

Svaki *peeling chain* sastoji se od više *peeling chain* transakcija. Na temelju opisa *peeling chaina* u poglavlju 3 dolazimo do algoritma za detekciju *peeling chain* transakcija u početnom skupu transakcija (Algoritam 4.1).

```
Ulaz: T – početni skup transakcija
Izlaz: transakcije iz T koje su dio peeling chainova
T_peel := ∅
for each transakcija t ∈ T do
  if len(t.inputs) = 1 and len(t.outputs) = 2 then
    if t.inputs[0].address.n_tx ≤ 2 and t.outputs[0].address.n_tx ≤ 2 and t.outputs[1].address.n_tx ≤ 2 then
      T_peel := T_peel ∪ t
    end if
  end if
end for
return T_peel
```

Algoritam 4.1 Detekcija *peeling chain*

Ulaz u algoritam je početni skup transakcija u kojemu se traže *peeling chain* transakcije. Jedan od mogućih odabira početnog skupa transakcija je odabir svih *Bitcoin* transakcija iz određenog vremenskog perioda. U početnom skupu transakcija traže se *peeling chain* transakcije na temelju njihovih svojstava opisanih u poglavlju 3:

- Svaka *peeling chain* transakcija ima točno jedan ulaz i točno dva izlaza
    - Ulaz je primanje sredstava
    - Dva izlaza su daljnje grananje sredstava
- Svaka adresa koja je dio *peeling chaina* sudjeluje u ukupno najviše dvije transakcije
    - Prva transakcija je primanje sredstava
    - Druga transakcija je daljnje grananje sredstava (ako je nema to znači da je tu kraj *peeling chaina*)

Izlaz iz algoritma su transakcije iz početnog skupa transakcija koje su identificirane da pripadaju *peeling chainu*. Te *peeling chain* transakcije opisuju usmjereni graf G:

- Čvor je *Bitcoin* adresa
- Usmjereni brid označava dio transakcije gdje je početak brida ulaz u transakciju, a kraj brida je jedan od dva izlaza iz transakcije

Grupiranje *peeling chain* transakcija u *peeling chainove* ekvivalentno je pronalasku slabo povezanih komponenti grafa G.

4.1. Implementacija

Prikupljanje početnog skupa transakcija i pronalaženje *peeling chain* transakcija među njima implementirano je u Google BigQueryju. Google BigQuery je skladište podataka u oblaku. Na Google BigQueryju javno i besplatno je dostupan skup podataka s *Bitcoin* transakcijama koji se osvježava u stvarnom vremenu [10].

Kako se u Algoritam 4.1 koristi informacija o ukupnom broju transakcija u kojemu sudjeluje pojedina adresa, prvo je bilo potrebno u Google BigQueryju napraviti tablicu s tim podatcima. SQL upit za pronalazak broja transakcija u kojem sudjeluje svaka adresa prikazan je u Kod 4.1.

```
1  WITH double_entry_book AS (
2      -- debits
3      SELECT
4          ARRAY_TO_STRING(inputs.addresses, ",") AS address,
5          inputs.type,
6          -inputs.value AS value,
7          inputs.transaction_hash
8      FROM
9          `bigquery-public-data.crypto_bitcoin.inputs` AS inputs
10     UNION ALL
11     -- credits
12     SELECT
13         ARRAY_TO_STRING(outputs.addresses, ",") AS address,
14         outputs.type,
15         outputs.value AS value,
16         outputs.transaction_hash
17     FROM
18         `bigquery-public-data.crypto_bitcoin.outputs` AS outputs
19 ), addr_tx AS (
20     SELECT DISTINCT address, transaction_hash
21     FROM double_entry_book)
22 SELECT address, COUNT(address) AS cnt
23 FROM addr_tx
24 GROUP BY address
```

Kod 4.1 SQL upit za kreiranje tablice s brojem transakcija u kojima je sudjelovala svaka adresa

Linije 1 do 18 za stvaranje dvojnog knjigovodstva *Bitcoin* transakcija preuzete su s [11] uz izmjenu da se odabire i stupac s *hashem* transakcije. U linijama 19 do 21 iz dvojnog knjigovodstva odabiru se jedinstvene kombinacije *Bitcoin* adresa i *hasheva* transakcija u kojima je ta adresa sudjelovala. Potrebno je gledati jedinstvene kombinacije jer adresa može u isto vrijeme biti i ulaz i izlaz iz transakcije. U linijama 22 do 24 prethodno odabrane jedinstvene kombinacije adresa i *hasheva* transakcija grupiraju se po adres i broji se u koliko transakcija je korištena svaka adresa. Izvođenje tog SQL upita na Google BigQuerju u regijama US i EU prema cijeni važećoj u lipnju 2022. košta 0,90 $.

U Kod 4.2 prikazan je SQL upit za pronalazak *peeling chain* transakcija u skupu svih transakcija iz svibnja 2020.

```
1  WITH candidate_transactions AS (
2      SELECT *
3      FROM `bigquery-public-data.crypto_bitcoin.transactions` AS transactions
4      WHERE transactions.block_timestamp >= '2020-05-01' AND transactions.block_timestamp < '
           2020-06-01' AND transactions.input_count = 1 AND transactions.output_count = 2
5  ), candidate_addresses AS (
6      SELECT address
7      FROM `diplomski2-351920.bitcoin.address_count` address_count
8      WHERE cnt <= 2 AND cnt > 0
9  ), peeling_transactions AS (
10     SELECT *
11     FROM candidate_transactions
12     WHERE
13         ARRAY_TO_STRING(candidate_transactions.inputs[ORDINAL(1)].addresses, ",") IN (SELECT
               address FROM candidate_addresses) AND
14         ARRAY_TO_STRING(candidate_transactions.outputs[ORDINAL(1)].addresses, ",") IN (SELECT
               address FROM candidate_addresses) AND
15         ARRAY_TO_STRING(candidate_transactions.outputs[ORDINAL(2)].addresses, ",") IN (SELECT
               address FROM candidate_addresses)
16 )
17 SELECT peeling_transactions.hash, ARRAY_TO_STRING(flattened_inputs.addresses, ",") AS source,
       ARRAY_TO_STRING(flattened_outputs.addresses, ",") AS target
18 FROM peeling_transactions
19 CROSS JOIN UNNEST (peeling_transactions.inputs) AS flattened_inputs
20 CROSS JOIN UNNEST (peeling_transactions.outputs) AS flattened_outputs
```

Kod 4.2 SQL upit za pronalazak *peeling chain* transakcija u svibnju 2020.

U linijama 1 do 4 odabire se početni skup transakcija, u ovom slučaju sve transakcije iz svibnja 2020. Također, filtriraju se transakcije tako da ostanu samo one koje imaju točno jedan ulaz i dva izlaza. U linijama 5 do 8 iz tablice kreirane Kod 4.1 odabiru se sve adrese koje su sudjelovale u jednoj ili dvije transakcije. U linijama 9 do 16 odabiru se transakcije čiji su ulaz i izlazi korišteni samo u jednoj ili dvije transakcije. Te transakcije su *peeling chain* transakcije. U linijama 17 do 20 kreira se tablica koja opisuje izgled grafa transakcija *peeling chainova* pronađenih u početnom skupu transakcija. Ta tablica ima tri stupca: *hash* transakcije, izvor i odredište. Kako sve *peeling chain* transakcije imaju jedan ulaz i dva izlaza, za svaku *peeling chain* transakciju u toj tablici postojati će dva retka. Iz Google BigQueryja preuzeta je CSV datoteka te tablice. Izvođenje SQL upita prikazanog u Kod 4.2 na Google BigQueryju u regijama US i EU prema cijeni važećoj u lipnju 2022. košta 1,19 $.

Datoteka preuzeta iz Google BigQueryja dalje se obrađuje Python skriptom. Grupiranje *peeling chain* transakcija u pojedine *peeling chainove* prkazano je u Kod 4.3.

```
1  # Load BigQuery .csv to pandas dataframe
2  df = pd.read_csv("./big_query_5_2020/big_query_5_2020.csv")
3
4  # Replace nonstandard* addresses (OP_RETURN) with OP_RETURN_* to ensure that they are unique
5  new_target = []
6  i = 0
7  for addr in df["target"]:
8      if addr.startswith("nonstandard"):
9          new_target.append(f"OP_RETURN_{i}")
10         i = i + 1
11     else:
12         new_target.append(addr)
13 df["target"] = new_target
14
15 # Create networkx graph from pandas dataframe
16 G = nx.from_pandas_edgelist(df, source="source", target="target", create_using=nx.DiGraph())
17 print(f"Created␣graph␣with␣{len(G.nodes)}␣nodes␣and␣{len(G.edges)}␣edges.")
18
19 # Find peeling chains
20 peeling_chains = list(nx.weakly_connected_components(G))
```

Kod 4.3 Grupiranje *peeling chain* transakcija u *peeling chainove*

U liniji 2 .csv datoteka spremljena iz Google BigQueryja učitava se u pandas [12] tablicu. U linijama 5 do 13 mijenja se BigQueryjev način imenovanja OP_RETURN izlaza transakcija. U liniji 16 stvara se NetworkX [13] graf G. U liniji 20 korištenjem NetworkX biblioteke traže se slabo povezane komponente grafa G – *peeling chainovi*.

Traženje adrese koja je početak *peeling chaina* i vrijednosti s kojom započinje *peeling chain* prikazano je u Kod 4.4.

```
1  def find_source(peeling_chain, df):
2      sources = []
3
4      target_values = df["target"].values
5
6      for addr in peeling_chain:
7          if addr.startswith("OP_RETURN"):
8              continue
9
10         if addr not in target_values:
11             sources.append(addr)
12
13     print(f"Peeling chain source:")
14     for source in sources:
15         value = blockexplorer.get_address(address=source, session=session).total_sent
             /100000000
16         print(f"     {source} {value} BTC")
17
18     return sources
```

Kod 4.4 Traženje početka i početne vrijednosti *peeling chaina*

Funkcija prikazana u Kod 4.4 prima jedan *peeling chain* za koji se traži početak i početni iznos i skup svih *peeling chain* transakcija. Početna adresa *peeling chaina* je ona za koju u popisu svih *peeling chain* transakcija ne postoji ni jedan zapis u kojemu je ta adresa odredište. U linijama 6 do 11 traži se ta adresa. Za slučaj da je došlo do pogreške u procesiranju ova funkcija je napravljena tako da traži sve adrese koje zadovoljavaju kriterij da su izvor. Na taj način prilikom korištenja bi uočili da se za jedan *peeling chain* pojavljuje više izvora što bi ukazalo na pogrešku. U linijama 13 do 16 ispisuju se podatci o početno adresi i početnom iznosu zadanog *peeling chaina*. U liniji 15 korištenjem izmijenjene verzije blockchain.com Python klijenta [14] dohvaća se početna vrijednost *peeling chaina* (vrijednost *bitcoina* s početne adrese) s besplatne javne usluge za dohvat podataka o *Bitcoin*

transakcijama blockchain.com. U blockchain.com Python klijent dodan je *cache* korištenjem requests-cache [15] biblioteke kako bi se izbjeglo slanje prevelikog broja zahtjeva na blockchain.com uslugu.

U Kod 4.5 prikazano je kreiranje vizualizacije *peeling chaina* implementirano korištenjem pyvis biblioteke [16].

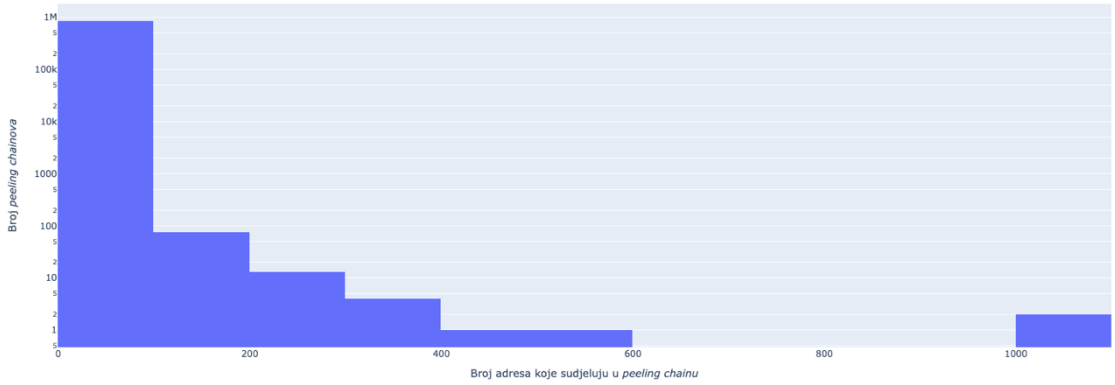
```
1  def visualize(peeling_chain, df):
2      tmp = df[np.logical_or(df["source"].isin(peeling_chain), df["target"].isin(peeling_chain)
           )]
3
4      sources = find_source(peeling_chain, df)
5
6      nx_graph = nx.from_pandas_edgelist(tmp, source="source", target="target", create_using=nx
           .DiGraph())
7
8      for source in sources:
9          nx_graph.nodes[source]["color"] = "red"
10
11     graph = Network(height='80%', width='100%',heading='', notebook=True, directed=True)
12     graph.from_nx(nx_graph)
13     graph.set_edge_smooth('dynamic')
14     graph.show_buttons(filter_=["physics"])
15     graph.toggle_physics(True)
16     graph.show(f"visualizations/peeling_chain_{len(peeling_chain)}.html")
```

Kod 4.5 Vizualizacija *peeling chaina*

Funkcija prikazana u Kod 4.5 prima jedan *peeling chain* za koji kreira i skup svih *peeling chain* transakcija. U liniji 1 iz skupa svih *peeling chain* transakcija odabiru se samo one koje su dio zadanog *peeling chaina*. U liniji 4 poziva se funkcija prikazana u Kod 4.4 za traženje početka i početne vrijednosti *peeling chaina* kako bi početna adresa mogla biti označena crvenom bojom na grafu. U liniji 6 kreira se NetworkX graf iz transakcija filtriranih u liniji 2. U linijama 8 i 9 postavlja se crvena boja čvora za izvor. U linijama 11 do 16 kreira se pyvis vizualizacija *peeling chaina* i sprema u HTML datoteku.

5. Rezultati

Algoritam za detekciju *peeling chainova* opisan u poglavlju 4 pokrenut je na svim *Bitocin* transakcijama iz svibnja 2020. Taj skup podataka sadrži ukupno 9 092 682 transakcija. Algoritam je detektirao 851 728 *peeling chainova*. Distribucija broja adresa koje sudjeljuju u pojedinim *peeling chainovima* prikazana je na Sl. 5.1.



Sl. 5.1 Distribucija broja adresa koje sudjeluju u pojedinim *peeling chainovima* u svibnju 2020.

5.1. Bitfinex krađa

Dio jednog od *peeling chainova* provedenih s ciljem skrivanja traga nakon krađe *bitcoina* s burze Bitfinex (poglavlja 3.2.1) odvio se tijekom svibnja 2020. Algoritam je uspješno detektirao dio tog *peeling chaina* koji se odvio tijekom svibnja 2020 (Sl. 5.2).

Sl. 5.2 *Peeling chain* proveden nakon krađe *bitcoina* s burze Bitfinex (dio koji se odvio u svibnju 2020.)

Time je pokazana uspješnost algoritma opisanog u poglavlju 4 u pronalaženju *peeling chainova* u skupu *Bitcoin* transakcija.

5.2. Najveći *peeling chainovi* u svibnju 2020.

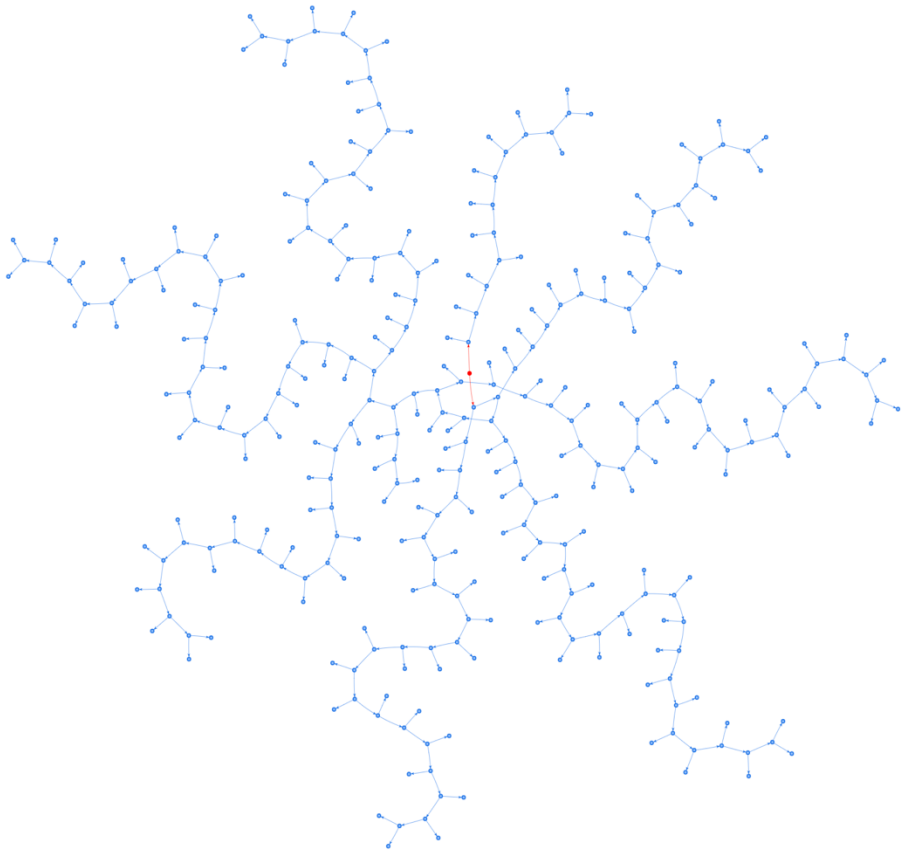
Od svih *peeling chainova* koje je algoritam detektirao u svibnju 2020. njih deset s najvećim brojem adresa koje sudjeluju u njima sadrže 1083, 1065, 553, 441, 369, 355, 341, 309, 267 i 261 adresu. Važno je napomenuti da je moguće da su ti *peeling chainovi* i veći, ali kako je u ovom radu detekcija rađena samo na skupu svih transakcija iz svibnja 2020. onda su i detektirani samo njihovi dijelovi koji su se dogodili u tom vremenskom razdoblju. U daljnjoj analizi ti *peeling chainovi* biti će nazivani po broju adresa koje sudjeluju u njima, primjerice *peeling chain* u kojemu sudjeluje 1065 adresa bit će nazivan *peeling chain* 1065. U Tablica 5.1 prikazana je početna adresa i početni iznos deset najvećih *peeling chainova* iz svibnja 2020.

Tablica 5.1 Deset najvećih *peeling chainova* detektiranih u svibnju 2020., sivo osjenčani redovi su *peeling chainovi* koji samo spaljuju *bitcoine*

Oznaka *peeling chaina*	**Početna adresa**	**Početni iznos (BTC)**
1083	bc1qwpt5689xp5eahszy7dm7vvu7ckjkf89cyhfgle	0,00704605
1065	bc1qe4epje4wknpdvgvad8apr62j89l0cjtpgejdjq	0,00705034
553	3ABz5em2nsSYE1rHkPPxnuJdds9SFTsGzt	0,0150049
441	bc1qv0mxswm8mjpetvfsyaj6y329ycjk4tj0urqydj	0,00898792
369	bc1qj3wfzehpujfxhmylp732h4tlk0u4as05tvca9m	0,5
355	bc1qe8jgferr6h629ep9ypuw356famywv5z6f4eaqs	0,00995241
341	1GpnAdUnZNcTbViXpdyrbvio9ApNd9cxTn	0,4996
309	3HaJGneCuf8R9ApLicdntnowXvjA8GKgXJ	0,39756156
267	133i6dbi4B5oG47uwAFJTbGU79YAgN6mvj	0,25
261	13HWakCV1vhZ8GUnWTVtvpBmEGJSi4r9qD	0,16694511

Prvo što je primijećeno u analizi tih *peeling chainova* je da neki od njih ne rade ništa korisno već samo spaljuju *bitcoine*. U svim transakcijama koje čine prvih pet najvećih *peeling chainova* (1083, 1065, 553, 441 i 369) jedan od izlaza je bio OP_RETURN što znači da su ti *bitcoini* spaljeni. Tih pet *peeling chainova* nije dalje analizirano, ali korištenje *peeling chainova* za spaljivanje *bitcoina* je svakako zanimljiva tema za buduće istraživanje. Postavlja se pitanje zašto se netko trudi provodit *peeling chain* za spaljivanje *bitcoina* kada je to isto mogao napraviti u jednoj transakciji. Kako se radi o spaljivanju ti *bitcoini* su zauvijek izgubljeni pa nema ni potrebe za skrivanjem ikakvog traga.

Peeling chain 355 počeo je 30. svibnja 2020. s iznosom od 0,00995241 BTC što je, prema zadnjem tečaju tog dana (1 BTC = 9 700,41 $, [8]), vrijedilo 96,54 $. Od *peeling* chaina 355 zanimljiviji je *peeling chain* 341. *Peeling chain* 341 sadrži samo 14 adresa manje od *peeling chaina* 355, ali njegova početna vrijednost u *bitcoinima* je preko 50 puta veća. *Peeling chain* 341 počeo je 19. svibnja 2020. s iznosom od 0,4996 BTC što je, prema zadnjem tečaju tog dana (1 BTC = 9 729,04 $, [8]), vrijedilo 4 860,63 $. Peeling chain 341 prikazan je na Sl. 5.3.



Sl. 5.3 *Peeling chain* koji se sastoji od 341 adrese u svibnju 2020., crveni čvor je početna adresa *peeling chaina*

Kao što se vidi na Sl. 5.3, *peeling chain* 341 počinje kao *peeling chain* stablo i nakon nekoliko transakcija prelazi u linearni *peeling chain*.

Prilikom pretraživanja literature za izradu ovog rada nije pronađena literatura koja spominje *peeling chainove* navedene u Tablica 5.1. Korištenjem internetskog pretraživača Google nije pronađeno spominjanje početnih adresa tih *peeling chainova*. BitcoinWhosWho (https://www.bitcoinwhoswho.com) i BitcoinAbuse (https://www.bitcoinabuse.com), sustavi za pretragu i označavanje *Bitcoin* adresa u svrhu sprječavanja ilegalnih aktivnosti, nemaju zapise o početnim adresama tih *peeling chainova*. Na temelju toga zaključujem kako su korištenjem algoritma za detekciju *peeling chainova* opisanog u poglavlju 4 detektirani do sada neotkriveni *peeling chainovi*.

6. Zaključak

U ovom radu pokazano je zašto postoji potreba za skrivanjem tragova na *Bitcoinu* iako je on pseudoniman. Predstavljen je obrazac uporabe *peeling chain* koji se koristi za skrivanje tragova. Korištenje *peeling chaina* ne znači samo po sebi da se radi o ilegalnim aktivnostima, ali zbog broja transakcija i naknada koje je potrebno platiti za provođenje *peeling chaina* postavlja se pitanje zašto bi ga dobronamjerni korisnik koristio.

Opisan je algoritam za detekciju *peeling chainova*. Algoritam je pokrenut nad skupom svih *Bitcoin* transakcija iz svibnja 2020. (9 092 682 transakcija). Algoritam se pokazao uspješnim u detekciji *peeling chainova*. U svibnju 2020. detektirano je ukupno 851 728 *peeling chainova*. Dio poznatog slučaja korištenja *peeling chaina* nakon krađe *bitcoina* s burze Bitfinex se odvio tijekom svibnja 2020. i njega algoritam uspješno detektira. Najdulji detektirani *peeling chain* sastoji se od 1083 adrese. Pet najduljih detektiranih *peeling chainova* ne rade ništa korisno, samo spaljuju *bitcoine*. Pretragom literature nije pronađeno spominjanje deset najvećih *peeling chainova* iz svibnja 2020. detektiranih korištenjem algoritma prikazanog u ovom radu. Na temelju toga može se zaključiti da su algoritmom opisanim u ovom radu pronađeni do sada neotkriveni *peeling chainovi*.

Iz ovog rada proizlazi nekoliko ideja za daljnje istraživanje *peeling chainova*. Moguće je proširiti algoritam tako da uzima u obzir dodatne parametre kod odlučivanje je li određeni niz transakcija *peeling chain*. Jedni od mogućih dodatnih parametara su ukupni broj adresa koje sudjeluju u *peeling chainu*, vremenski razmak između pojedinih transakcija i kome pripadaju (ukoliko je to moguće saznati iz nekog izvora) adrese na kojima na kraju *peeling chaina* završe sredstva (primjerice radi li se o burzi). Također, *peeling chainovi* koji samo spaljuju *bitcoine* su zanimljiva tema daljnjeg istraživanja.

Literatura

[1] Zhang, Y., Jun W., and Jie L. *Heuristic-based address clustering in bitcoin*, IEEE Access, (2020), 210582-210591.

[2] Fleder, M., Kester, M., Pillai, S. *Bitcoin Transaction Graph Analysis,* arXiv:1502.01657 (2015)

[3] Narayanan, A., Bonneau, J., Felten, E., Miller, A., Goldfeder, S. *Bitcoin and Cryptocurrency Technologies*. *Draft*. Princeton University Press, 2016.

[4] SlowMist, *Analysis of the $3.6 Billion Recovered by the U.S Government from the 2016 Bitifnex hack*, SlowMist Medium, (2022, veljača). Poveznica: https://slowmist.medium.com/analysis-of-the-3-6-billion-recovered-by-the-u-s-government-from-the-2016-bitifnex-hack-46abc296342d; pristupljeno 10. lipnja 2022.

[5] Collins K., *Watch as these bitcoin wallets receive ransomware payments from the ongoing global cyberattack*, Quartz, (2017, svibanj). Poveznica: https://qz.com/982993/watch-as-these-bitcoin-wallets-receive-ransomware-payments-from-the-ongoing-cyberattack/; pristupljeno 24. lipnja 2022.

[6] Ball J., *Silk Road: the online drug marketplace that officials seem powerless to stop*, The Guardian, (2013, ožujak). Poveznica: https://www.theguardian.com/world/2013/mar/22/silk-road-online-drug-marketplace; pristupljeno 19. lipnja 2022.

[7] 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 (IMC '13). Association for Computing Machinery, New York, (2013), str. 127-140.

[8] SlowMist, *Crypto Compliance Series | What is Peel Chain*, SlowMist Medium, (2022, veljača). Poveznica: https://slowmist.medium.com/crypto-compliance-series-what-is-peel-chain-7b5be0bb7214; pristupljeno 10. lipnja 2022.

[9] *Historical Data for Bitcoin*, CoinMarketCap. Poveznica: https://coinmarketcap.com/currencies/bitcoin/historical-data/; pristupljeno 25. lipnja 2022.

[10] Day A, Bookman C., *Bitcoin in BigQuery: blockchain analytics on public data*, Google Cloud Blog, (2018, veljača). Poveznica: https://cloud.google.com/blog/topics/public-datasets/bitcoin-in-bigquery-blockchain-analytics-on-public-data; pristupljeno 13. lipnja 2022.

[11] Day A., Medvedev E., AK N., Price W., I*ntroducing six new cryptocurrencies in BigQuery Public Datasets—and how to analyze them*, Google Cloud Blog, (2019, veljača). Poveznica: https://cloud.google.com/blog/products/data-analytics/introducing-six-new-cryptocurrencies-in-bigquery-public-datasets-and-how-to-analyze-them; pristupljeno 17. lipnja 2022.

[12] The pandas development team, *pandas-dev/pandas: Pandas*, Zenodo, (2020, veljača). DOI: 10.5281/zenodo.3509134

[13] Hagberg, A. A., Schult, D. A., Swart, P. J. *Exploring Network Structure, Dynamics, and Function using NetworkX*. Proceedings of the 7th Python in Science Conference (SciPy 2008). Pasadena, (2008), str. 11-15.

[14] blockchain 1.4.4. Poveznica: https://pypi.org/project/blockchain/; pristupljeno 24. lipnja 2022.

[15] Requests-Cache. Poveznica: https://requests-cache.readthedocs.io/en/stable/index.html; pristupljeno 24. lipnja 2022.

[16] Perrone, A., Unpingco, J., Lu, H. M. *Network visualizations with Pyvis and VisJS*, arXiv: 2006.04951 (2020)

Sažetak

Tomislav Babić

Detekcija *peeling chainova* u sustavu kriptovalute *Bitcoin*

Kriptovaluta *Bitcoin* je pseudonimna, korisnici se ne identificiraju stvarnim imenom nego *Bitcoin* adresom. Iz raznih razloga, bilo dobronamjernih ili zlonamjernih, korisnicima je u interesu da se njihove *Bitcoin* adrese međusobno ne povežu. U ovom radu analiziran je obrazac ponašanja *peeling chain* koji se koristi za zametanje traga (sprječavanje povezivanja različitih adresa kao adresa koje pripadaju istoj osobi) na *Bitcoinu*. *Peeling chain* može napraviti svaki korisnik *Bitcoin* sustava, ali upitna je motivacija dobronamjernih korisnika za to. Zlonamjernim korisnicima on omogućuje skrivanje veze između adresa direktno povezanih s ilegalnim aktivnostima i drugih adresa na koje raspodjeljuju svoje *bitcoine*. Zbog toga je prepoznavanje *peeling chainova* od velike važnosti za sprječavanje ilegalnih aktivnosti poput krađe i trgovine zabranjenom robom. U ovom radu opisan je algoritam za detekciju *peeling chainova* i demonstrirana je njegova uspješnost na različitim skupovima podataka. Algoritam je implementiran korištenjem Google BigQuery skladišta podataka u oblaku i Python skripte.

Ključne riječi: *Bitcoin*, anonimnost, *peeling chain*, Google BigQuery

Summary

Tomislav Babić

Detection of peeling chains in Bitcoin

Bitcoin is a pseudonymous cryptocurrency, users identify themselves using their Bitcoin addresses, not by using their real names. For a variety of reasons, users, being malicious or benevolent, don't want their multiple Bitcoin addresses being connected among themselves. In this paper I conducted an analysis of a Bitcoin idiom of use called peeling chain. It is used for hiding the trace between multiple Bitcoin addresses controlled by the same person. Every Bitcoin user can create a peeling chain, but the motivation of a benevolent user for creating one is questionable. Malicious users use peeling chains to hide the trace between their addresses that are directly connected with malicious activity and the addresses to which they spread their bitcoins. Because of that, detection of peeling chains is of the utmost importance in stopping illegal activities like stealing and receiving payments for the sale of illegal goods using Bitcoin. In this paper, I described an algorithm for detecting peeling chains and demonstrated its efficiency on a variety of datasets. Algorithm was implemented using Google BigQuery cloud data warehouse and a Python script.

Keywords: Bitcoin, anonymity, peeling chain, Google BigQuery