A.J. Varshneya
Sugat Poudel

Problem Set 3: CSI, The Blockchain

Problem 1

a. 1BaQak1XvV4L6FAP741mWuFMh7yoyW2jjn
b. The address was in use from 8/6/2014 at 6:27 pm to 8/16/2014 at 10:24 am, first as the output address then as the input and the change address.
c. There was only one victim that paid into this address. However, this transaction has two input addresses.
d. This victim paid around 1.72 BTC or $1004.41 into this address.
e. There is one outgoing transaction from this address, 1.71 BTC (848.16) or 99.37% of the total funds was sent to the address 1JdM2X8S98f6sm8MAZuaqodhzuKmQDK7WB and the rest (minus the tx fee) was sent back to this ransom address as change, where it still remains unspent. There is no definite pattern of fees, as it seems this address was mainly a singleton intermediary for a larger pool for it seems to have been abandoned since the funds left (0.0099 BTC or $4.91) are inconsequential and have remained there for a year now.
f. This address sent most of its funds (1.71 BTC or 99.37%) to the address 1JdM2X8S98f6sm8MAZuaqodhzuKmQDK7WB. This output address was combines with several others to distribute the 1.71 coins over 33 addresses. Most of the addresses received payments of around 0.26 BTC or $130 while 3 addresses received a payment of exactly 1.4176735 BTC or $701.56. There are several possibilities to this scheme. First, it could be a payout to people involved in the ransom organization with the 0.26 BTC sent to lower level employees and the 1.4176735 sent to people higher up in the chain. The number of output address alludes to the size of this organization either in terms of people involved or the addresses and funds at their disposal. However, the following transaction are of the same structure with multiple inputs and outputs and this scheme leaves too few a degree of separation between the ransom act and the people involved. Second, the payment to this address could have been a withdrawal from an exchange therefore, there are several addresses involved in the inputs and outputs of transactions. However, this is unlikely because there would only be one degree of separation from the point of ransom to the cash payout. Third and the most likely given

the nature of this crime, it is potentially a feed into a mixing service, that is dividing the coins into other addresses essentially, mixing the original input and sending a fraction of the original payment to the requester from a completely different address thus preserving anonymity of the ransomist. I am more inclined to believe this case because all transactions following the ransom transactions are multi input and multi output transactions with discrete output quantities, making it divide coins into multiple accounts and sum them later to send to a common account. Most of these connected addresses have one incoming and one outgoing transaction and they all occur in a span of a few hours strongly suggesting the work of an automated mixing service.

g. Most of the other addresses seemed to have other victims paying into them whereas mine only had one. The victim on average had to pay around $500 to the ransomists while some had double that amount, my victim had to pay double the "standard" amount. Some of the other addresses did not send all of their available funds to one address but rather distributed fractions of it to multiple addresses possibly as fees or to different mixers. In comparison, my address sent to a single location. Also for some addresses, the subsequent "mixers" have many pages of one to one transactions with a large quantity payout in each associated with one address whereas the "mixers" for my address had a series of addresses with around 2 many to many transactions with the payout divided to smaller quantity payouts. Most of the addresses, the chain of funds could not be easily followed to one source, the motive was mainly that of obfuscation thus the reasoning for many confusing transactions.

A. 13BeAzA4mhwDYJEwhqNd2LsUnuhuVqKvw8
B. The latest transaction involving this address occurred on 8/30/14.
C. There appears to have been 38 victims, assuming the parties sending coin in each transaction are unique.
D. Most transactions were ~1 BTC (about $380), with just a few being 2 BTC (about $760).
E. The proceeds (39 BTC, $15125) were split primarily between two addresses. The first address received virtually all of the funds, giving about 5 percent to the second address. Going a little bit deeper into the transaction history, the funds received by the next address are split with about 15-20 percent going to two addresses (for a total of 30-40 percent) and the remainder going to a single address.

F. It seems like a lot of the money is being exchanged after 2-3 transactions from the ransomer's original address. The money goes to an address which has received in total about $250 million dollars which leads me to suspect that it's an exchange. I am interested to know how they manage to exchange the funds without getting arrested.

G. It seems that many of these ransomer's addresses were active from mid to late 2014. Most (probably all) addresses have a final balance of $0.00. From reading other comments, it seems that most ransomers had fewer than 30 victims for a particular address and had ransoms of around $500.

Problem 2

a. In the process of buying or opening an account with these vendors, it was possible to directly determine their addresses and public keys, and tag them in the block chain making it easier to trace their flow of funds in the public ledger and group addresses that they determined to be under the control of these vendors rather than gathering such data from forums or other external source, which might prove to be illegitimate. This strategy gives them an idea of what addresses each of these services uses and how they may interact with other services.

b. We probably want to consider whether transacting with the questionable merchant will promote illicit activity and to what degree. For instance, transacting with an exchange which openly does business with ransomers might be ethically questionable because we are supporting a service that promotes criminal activity. If we aren't bothered by ethics, we still should be mindful of legal considerations. For instance, using a mixing service might somehow tie our funds to criminal enterprises. By conducting transactions with groups like this, you may be supporting illicit services including the sale of drugs, firearms or even murder, which most would argue is ethically concerning.

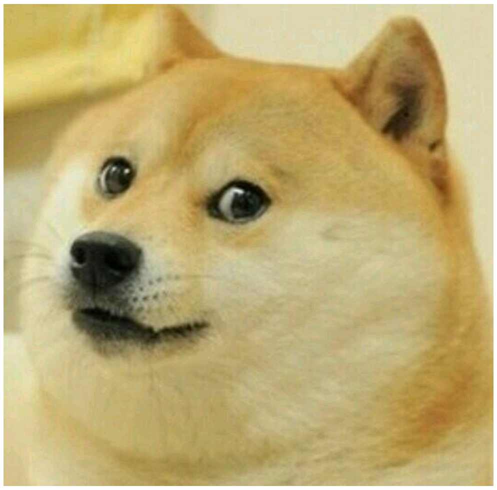
Problem 3

Explain why this is not true?

Multiple inputs can be from different users and each user does not have to know the private key of the others because of how transactions are signed in the unlocking script. Generally, unlocking scripts contain OP_DATA sig and OP_DATA pub_key, pushing the signature for the transaction and the public key onto the stack. The signature involves the transaction and the private key of the user however nowhere is the private key explicitly stated. The signature for each user is produced independently and then included as part of the transaction. Multiple users can sign a transaction using their private key and include the signature as part of the unlocking script without alerting the other users of their credentials and thus one can’t definitively say that multiple input addresses belong to the same user.

Problem 4
How does the widespread use of HD wallets today impact the effectiveness of their Heuristic 2? Heuristic 2 states that the one time change address is controlled by the same user as the input address; this can be further used to cluster related address into a single entity. Since users are better able to protect their anonymity by using addresses with a single input, they derive many such addresses with their original seed using an HD wallet. The paper assumes that change addresses have only one input, but with the widespread adoption of HD wallets, there are many more non-change addresses with single inputs. Hence, if this analysis was performed again, there would be more opportunity for false positives in detecting change addresses, which are difficult to determine in the first place.

Problem 5



We found the anonymity set to contain 4 elements. Using the range given in the assignment we were only able to find one transaction with a matching output however the associated address was the change address thus it could not possibly be the forward address. We increased the range up to 1% of the input value and determined the following anonymity set:

```
{('1KoYHhftuWMMmQu3gUQgjfpJ5jaibV1FmL', 1462960),
('1EUAJ4ei5Qi3Z9ybEh2BCJjDnUyBjDMyMM', 1463971),
('1K5LWuc6R9LAK5stnnzn6nxbrotV1vYDBo', 1461034),
('1Jztg1YqyHda4bkcPD36E786Wfxm7NVHdw', 1462430)}
```

The following was the function used to find the anonymity set:

```
def get_anon_set(API_KEY):
    amount = 0.01523
    target_blk = 379818
    min_val = (amount - (amount*0.01 + 0.0005)) * 100000000
    max_val = (amount - (amount*0.005 + 0.0005)) * 100000000

    b = blockcypher.get_block_overview(target_blk,
                                       txn_limit=3000, api_key=API_KEY)
    receiving_addresses = set()

    count = 0
    for x in b['txids']:
        transaction = blockcypher.get_transaction_details(x, api_key=API_KEY)
        print('looking at tx ' + str(count) + ': ' + str(x))
        for output in transaction['outputs']:
            if min_val <= output['value'] <= max_val:
                print('tx matched')
                for adr in output['addresses']:
                    receiving_addresses.add((adr, output['value']))
        count += 1

    return receiving_addresses
```

Problem 6

How much does it enhance anonymity to split your output across two forward addresses?
With two forwards addresses, for every output value less than the mixed amount, we will need to add it to every other value less than the mixed amount to see if their sum is within the amount that could be returned by the mixing service. At worst case, this would have a runtime of $O(n^2)$ assuming n is the number of outputs, effectively increases work by a power of 2. With one forward address, we have a specific interval of mixed values however with 2 forwarding address we will have to look at every output value less than the max mixed amount, which would increase the range of candidate output addresses drastically increasing our anonymity set.

Problem 7

To observe any interesting transactions from the suspect addresses, we iterated through the addresses in the file 'suspects.txt' and found the set of receiving addresses from each suspect addresses. We then looked at any labels associated with each of these addresses. We found several of the addresses with the label 'BTC-e.com' and a few with the label 'MtGoxAndOthers'. We were also able to find a couple of addresses that wallet explorer identified as 'BTC-e.com-old'.

The following was the function used to find the labels:

```
def track_outputs(api_key):
    file = open('suspects.txt', 'r')
    count = 0
    for adr in file:
        adr = adr.strip()
        try:
            receiving_addresses = get_receivers(adr, api_key)
        except KeyError:
            receiving_addresses = []
        addresses = set()
        for r in receiving_addresses:
            addresses.add(r[0])
        for x in addresses:
            disp = address_display(x)
            if disp != 'No label':
                print('%s: %s' % (count, disp))
            count += 1
    print(count)
```

Results:

```
1868D1uFb37sxnDEcY8HqvrTuyi2f9erLF (BTC-e.com)
17CBexLgbV97DJNTMxCVWuDPeonsdVCfWn (MtGoxAndOthers)
1HTPr5Fm4B7SRQpDRFq9mEDi4CRsUZY5N1 (MtGoxAndOthers)
1EiVrdKTemTu4sUZvpvngxFVNv8XQRRwHT (BTC-e.com-old)
1G9yhYN4tbRjP9nAgZcaTM1GSyUzbVsqjs (BTC-e.com)
1PLXGbd5zn7z31bUSZKU75SgnnTbD9fEsa (BTC-e.com)
1Q3DKokAdXstCegFVJpVWVZpR9FpyKs9wJ (MtGoxAndOthers)
1GR7SzCBRUY7EseLGE32dnWMjkJ3hPVykx (MtGoxAndOthers)
183SAtEDMR8uRmtpE2Vm1HG1uLjrXWHrmi (MtGoxAndOthers)
1JirvMktsdZhLJQ44DqkptzJD7E2JoDpCW (BitPay.com-old)
1FCY9286DwkfgKvx7TjhtmyXTeTjM1w7WP (MtGoxAndOthers)
16KnVF5hQNrR3KZpnBqpZP96iCsWrUmEvP (MtGoxAndOthers)
1JXATE9G22VtHaXJrDfbjUKMcixt9BUCvy (BTC-e.com)
13hZnvvKTjcRvuSMys9K3yavyKo8KRsSHB (BTC-e.com-old)
1Csj5DVgMyn5Ui3g4TEAihcjLxcRLatE9 (BTC-e.com)
1D94Z2xpnfTMrbggWQy8aS75P8wX3FfJTf (MtGoxAndOthers)
1QEvsoFM3m3xzsMD9g3mAKTjpYhzYLgwtj (BTC-e.com)
1KZxuRq9XE3agSAnrDuaFNEMihS1sX8jjy (MtGoxAndOthers)
15b1FECZc4ULStRSQmmT1GF2n1wrqyTnma (BTC-e.com)
1HtdbJVs9EiFvaRqi8mdF6XSqpBcQh9czV (BTC-e.com-old)
16WNx8DbM2fGnKCHAif4Pqsm7oPwKyJEp1 (BTC-e.com)
17eFLK9mS2AmAUN51be5rtAXUUkPUgkTeH (MtGoxAndOthers)
19m6F6ULT2xNVEVEhNuLahe7na32wdJAzD (BTC-e.com)
1EuZ14muU2KHnx6gMcmvHiUMLrPX1nShWi (BTC-e.com-old)
1Ym1zJXPZ88MjUsT17ZRfr1TVfdkAagMn (BTC-e.com-old)
1JybiPnDipWvBsJMmM1eNiEXpeX5o5b5VD (MtGoxAndOthers)
1CmWF4mYoaBWyQ8u3vDknrPHEFbYQBmuUK (BTC-e.com)
```

```
1Q96MDPwVEAUDek4hKYtXgRyTxyqdESWbA (MtGoxAndOthers)
1FiBsrZJbntU2gG3YR9bJM7VsWis43UjaK (MtGoxAndOthers)
19p1VwVCbub7Jepc5tdD3Dxv2Sqq8fugkF (MtGoxAndOthers)
1Mw8BVwyBrtkiDxGK2Uvf1amXR4f4rppqZ (BTC-e.com)
1QDXckhUmGVWn1u4ZstR4G4oKuXMJREtzC (MtGoxAndOthers)
1KYP1AbWaz2Ayx4PYvcHLCNcKqedqSHjsi (BTC-e.com)
12Af1YDmCacZCAaU7QYmaa42uApQmqmizM (MtGoxAndOthers)
1NkjXACtWM358virDfhbE8rZsWWgrkVhsa (BTC-e.com)
1MEqWgRJagKD5Gghv92fxXAjLvePon44yF (MtGoxAndOthers)
1HHzawQkvdUZfKxafbcsag9P38fbBxJCMf (MtGoxAndOthers)
1K7PFLSDdCqDXUdjWvvmXPjuiD9776CgPD (MtGoxAndOthers)
1JdM2X8S98f6sm8MAZuaqodhzuKmQDK7WB (MtGoxAndOthers)
1FRtvFuwn736RmCfnroE3NmtWDR8jmjMZz (BTC-e.com-old)
1BjaJYsXWhFdB1NtAbcujwByUx3RYHHXSP (BitPay.com-old)
17pdZBqdPwYHfihSgCXmPHa7u3cXYjV7Ac (BTC-e.com)
1AJUVUSZ1Lx1hTwUK8PR6BLiBmxQPJic94 (MtGoxAndOthers)
19yDnvAzfhjoDwssWUBXbS2oErzptWmidY (BTC-e.com-old)
141BUEQmdFdYT88jqRKfwn4qakEFGvFwoD (BTC-e.com)
1Q7VApg98W8mtAFdhJc7EPFCbHqcskLKAB (MtGoxAndOthers)
12iYaS1psNH97xVsdYDwZXZAiwSouJtXT3 (MtGoxAndOthers)
1MAizeQjJLdVxiqMKWd7Q6qTEJheEzFZHQ (BTC-e.com)
1DTmn1TJXqjR3azLxm66Mvt8exyYpk2jis (BTC-e.com)
14JSxVrYsibrZDxrTE3nbpBEAbTnpsHPDV (BTC-e.com)
1K1Dr9vqyMwdsgG5QVAnt3g6tN7JTnceFn (MtGoxAndOthers)
1KAJjSbd5c9WSeNzTS5ZSnStmWhk5pr2Vk (MtGoxAndOthers)
1PSfefL7fzxagWtaH3apo6o6vHanpJVVE3 (BTC-e.com)
```