JoinMarket

An in depth analysis of a distributed and trustless mixing system

powered by
NEUTRINO

Bitcoin and privacy

Bitcoin is a digital currency that utilizes cryptographic identities (bitcoin addresses) to receive and send money. While this guarantees its users a certain level of privacy, an analysis of the blockchain, the technology which bitcoin is based upon, permits the activity of every single address to be tracked. The blockchain is a public ledger containing all of the effected transactions using this currency and thus can be consulted by anyone.

For example: Bob credits Alice her paycheck on a bitcoin address which she has provided; by monitoring the activity of that address on the blockchain, Bob will be able to know how and when Alice spends the money. Other elements such as *address reuse*[1] and *clustering*[2] algorithms can further permit for this type of monitoring to be extended from a single address to Alice's entire *wallet*[3].

The demand for tools which increase the level of privacy offered by Bitcoin has led to the creation of the first centralized mixers. A user sends a certain amount of bitcoins to the mixing service where they are then grouped together with funds from other users. The total sum is then returned to more than one destination along a series of addresses indicated by the user, in this way the returned bitcoins will no longer be linked directly to their source. Many of the mixing services are active on the *darknet* in order to further increase the level of anonymity. Among the most noted of these sites is BitcoinFog, which was used, along with others, for money laundering from the first generation of CryptoLocker and for stealing from Sheep Marketplace[4].

CoinJoin and its implementations

The main problem in using these systems is that the user must place complete trust in those that manage the service, not only because they will be able to keep track of all the phases of mixing, but most of all because they will manage the users bitcoins directly, amounting in the risk of potential theft. This can represent a certain level of criticality given the scarce accountability of these services, especially for those active on the darknet. It is with the aim of overcoming this obstacle that *trustless mixing services* were born. In a

[1] https://en.bitcoin.it/wiki/Address_reuse
[2] Evaluating User Privacy in Bitcoin - https://eprint.iacr.org/2012/596.pdf
[3] All the addresses belonging to the same spending entity
[4] https://en.wikipedia.org/wiki/Sheep_Marketplace

Forum in 2013, G. Maxwell proposed a new way of carrying out Bitcoin transactions which he called CoinJoin[5]. The idea behind it is simple: *N* users agree to perform a transaction with *N* output (separate and of the same value), they then provide a certain amount of input for the established sum. At this point each of the *N* users must individually place their own signature on the transaction in order to unblock their portion of the input, eliminating the risk that other participants steal their portion from them. When all of the participants have signed the transaction, the transaction can then be propagated on the network and inserted into the blockchain.

Given an output address in this this type of transaction, it is not easy to establish the real source of the bitcoins sent among the input given by the *N* users. Obviously, the more users who participate in every single transaction, the higher the level of privacy will be and consequently, the lower the fee per user.

It is nonetheless important to note how the original CoinJoin idea does not define in and of itself an entirely distributed system, given that it does not specify how the various actors must communicate amongst themselves. The *SharedCoin* service, offered by *blockchain.info*[6], serves as an example of one of the first to implement the functionalities of CoinJoin.

Over time the original idea underwent a series of refinements aimed at increasing its reliability and level of privacy, resulting in a series of new implementations. JoinMarket[7], one of the versions of CoinJoin, is the topic of the analyses reported in this document. JoinMarket is a trustless and distributed mixing system (*peer-to-peer*), which represents, as of today, perhaps the most reliable[8] implementation of G. Maxwell's original idea.

This is a relatively new system; the idea was first described on a post in January 2015[9], and is officially functional on the Bitcoin main-net since May 2015.

The improvement introduced by JoinMarket calls for the definition of two distinct roles among the users who agree to effect a CoinJoin transaction: the *maker* and the *taker*. On one side there are users with time and bitcoins at their disposal (maker), which, for a fee, will make these resources available to users who have an immediate need to carry out a mixing (taker). The negotiation is automatic and is initially done on an IRC channel and then finalized via an exchange of enciphered peer-to-peer messages. In order to further

[5] https://bitcointalk.org/index.php?topic=279249.0
[6] https://twitter.com/blockchain/status/402224010492006400
[7] https://github.com/joinmarket-org/joinmarket/wiki
[8] https://en.bitcoin.it/wiki/User:Gmaxwell/state_of_coinjoin
[9] https://bitcointalk.org/index.php?topic=919116.0

increase the level of privacy, every taker will typically link more than one mixing transaction consecutively, resulting in the creation of mixing chains.

The fact that the makers have an economic incentive to share their bitcoins should, in the intention of the author, attract "clean" money with which the other users (takers) can mix their own bitcoins. Moreover, with this implementation, there is no central entity present to keep track of the various phases of mixing[10] [11].

JoinMarket and the blockchain analysis

Basing upon public analyses of the privacy level offered by JoinMarket, and more generically by CoinJoin[12] [13], we have developed new and more refined algorithms included in the P-Flow[14] solution which allow us to carry out an extensive analysis of the blockchain. Thanks to the use of P-Flow it was possible to have a thorough picture of all of JoinMarket's transactions and consequently, to evaluate the use of this service[15]. Starting off with the sum of all of JoinMarket's transactions, we were able to isolate the single chains and pinpoint the transactions that had input bitcoins into the mixing system (funding transactions[16]), as well as the transactions where the mixed funds were spent (outflow transactions). Based on the results gathered it is possible to confirm that in a little more than 18 months, and starting with nearly 16,000 funding transactions, over **65,000 BTC**[17] transited in the JoinMarket system; the equivalent of nearly **31,000,000 USD**[18]**.**

[10] Specific attack aimed at monitoring the market of *makers* have been implemented, but in all the cases *workaround* for mitigate their impact have been found - https://github.com/JoinMarket-Org/joinmarket/issues/156

[11] What Logs are kept - https://en.bitcoin.it/wiki/Shared_coin

[12] https://github.com/AdamISZ/JMPrivacyAnalysis/blob/master/tumbler_privacy.md

[13] http://www.coinjoinsudoku.com/advisory/

[14] P-Flow is a tool, developed by Neutrino srl, for analyze blockchain and bitcoin transactions

[15]Please take into consideration that the algorithm we have used may spot – as false positive – some transactions performed with other mixing services similar to CoinJoin.

[16] They are standard bitcoin transactions used for moving funds into addresses to be used as first hop of the mixing chain.

[17] Of which about 10.000 BTC are still *unspent* after the mixing.

[18] Based on the value of BTC/USD exchange rate at the date of the funding transactions. It is important to note that the same bitcoin could have been mixed in different moments from different entities: it is not precise to consider this amount as a percentage of the total existing bitcoin

Identification of the sources and the recipients

Once the total outflow from the mixing system and the recipient bitcoin addresses have been identified, the problem which arises is that of back tracking and tracing the respective wallets: by taking advantage of the smart-clustering and categorization functionalities present in P-Flow, this is now possible. In this way we were able to have a clear picture of which "*well –known-actors*" have received bitcoins directly from the mixing chains without intermediate steps.

The results of the analysis show that the following services received the highest number of BTC:

Service wallet	BTC received	TX count	Know Your Customer
BTC-e	827	193	No
LocalBitcoins	525	170	No
Bitstamp	254	9	Yes

These services are among the most common money exchangers[19], it is therefore plausible to suggest that the funds were then converted into USD.

The next step of the analysis was to identify the geographical origin of the funding and outflow transactions. This was possible thanks to P-Flow's tracking engine[20] which uses machine learning algorithms on Bitcoin network traffic in order to analyze the propagation of the transactions in real time[21].

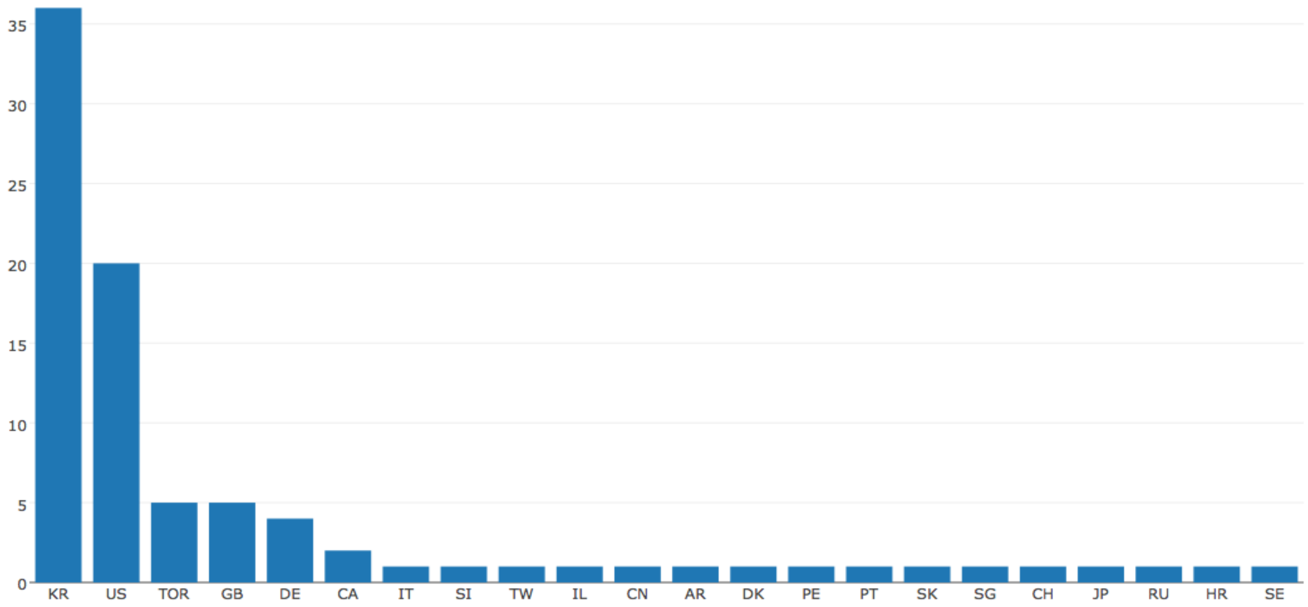
The following graphs demonstrate the geographic distribution[22] per percentage of the funding and outflow transactions.

[19] Know your customer is the process of a business, identifying and verifying the identity of its clients.

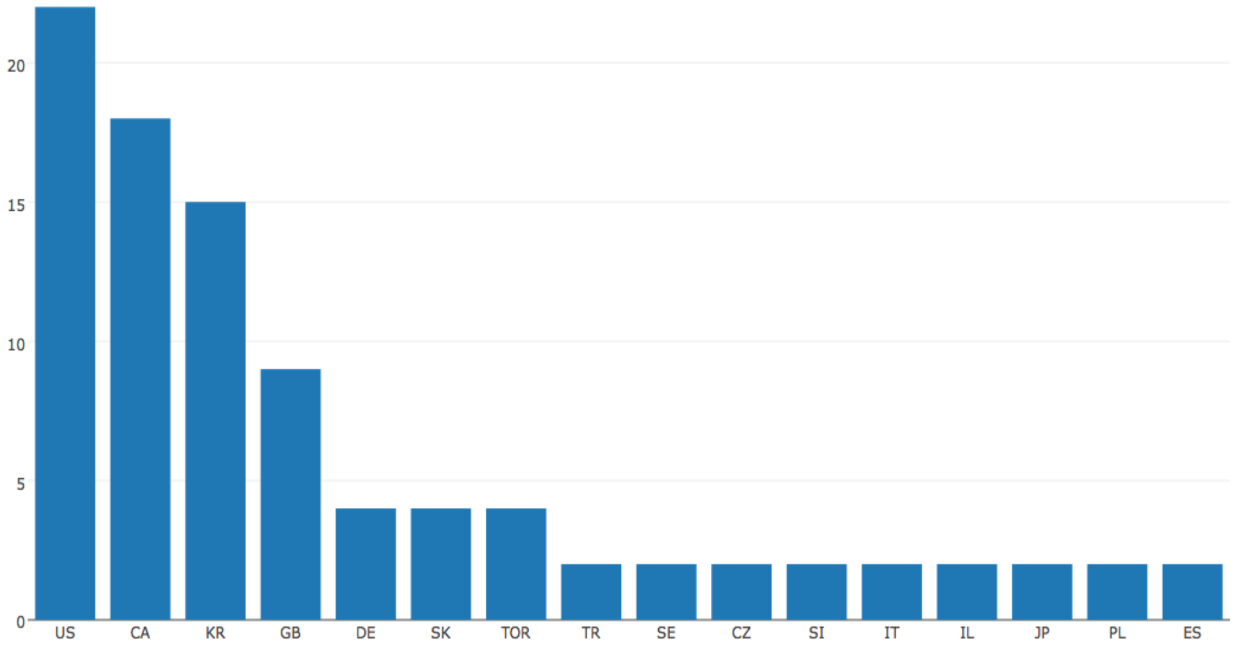
[20] This features extends and improves what is already provided by other services, such as blcokchai.info, As blockchain.info clearly states on its website, this is not a highly accurate.

[21] Due to bitcoin network peculiarities, data can not be guaranteed with 100% accuracy.

[22] Transaction associated with anonymization services are tagged as “Tor”.



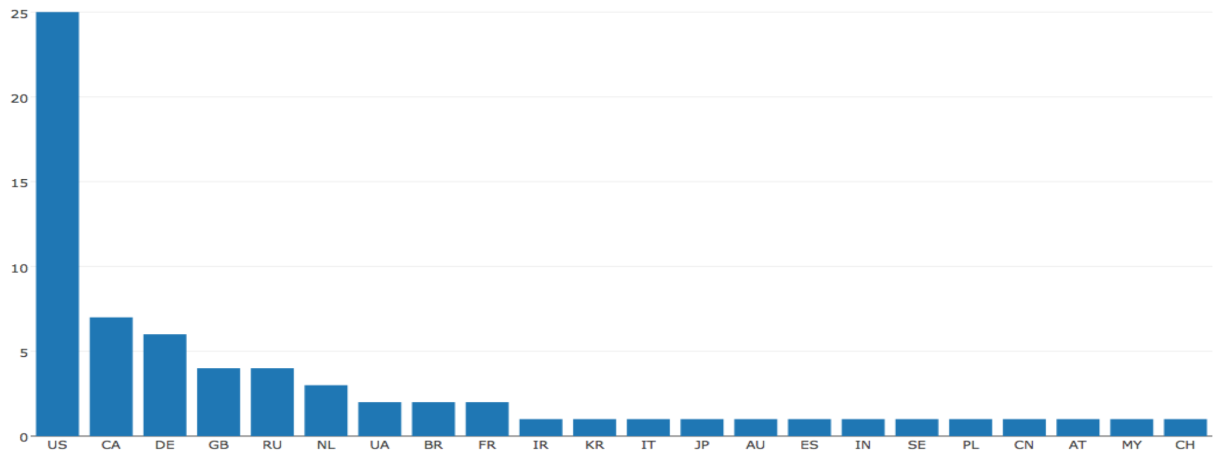
Distribution of funding transaction by country



Distribution of outflow transactions by country

In order to verify the consistency of these discoveries we evaluated the geographic distribution for all of the active bitcoin clients. The analysis was done by monitoring the announcements sent by clients[23] to active nodes on the network and by geo-referencing their IP addresses. If the distribution of Joinmarket was the same as the distribution of the bitcoin nodes, the two graphs should be very similar.

Bitcoin client distribution by country[24]



Excluding those countries with minor relevancy, where only a few errors could majorly influence results, we find that both graphs are consistent. Noteworthy however is the discrepancy in data relative to Korea where, percentage wise, JoinMarket seems to be much more widespread among Bitcoin users in respect to other countries.

Notable as well, the percentage of users utilizing TOR: this percentage, which makes up only a small portion of the total Bitcoin users, is however the much higher among JoinMarket users. One reason for this contrast could be that JoinMarket users may be more attentive to privacy.

Considerations regarding the use of JoinMarket

As suggested on the project's website by JoinMarket creator, we can identify three common uses for this service:

[23] https://en.bitcoin.it/wiki/Protocol_documentation

[24] Analyzing the graph data related to China are unexpected. Chinese mining pool are in control of most of the ashpower available. The number clients seems not to reflect this proportion.

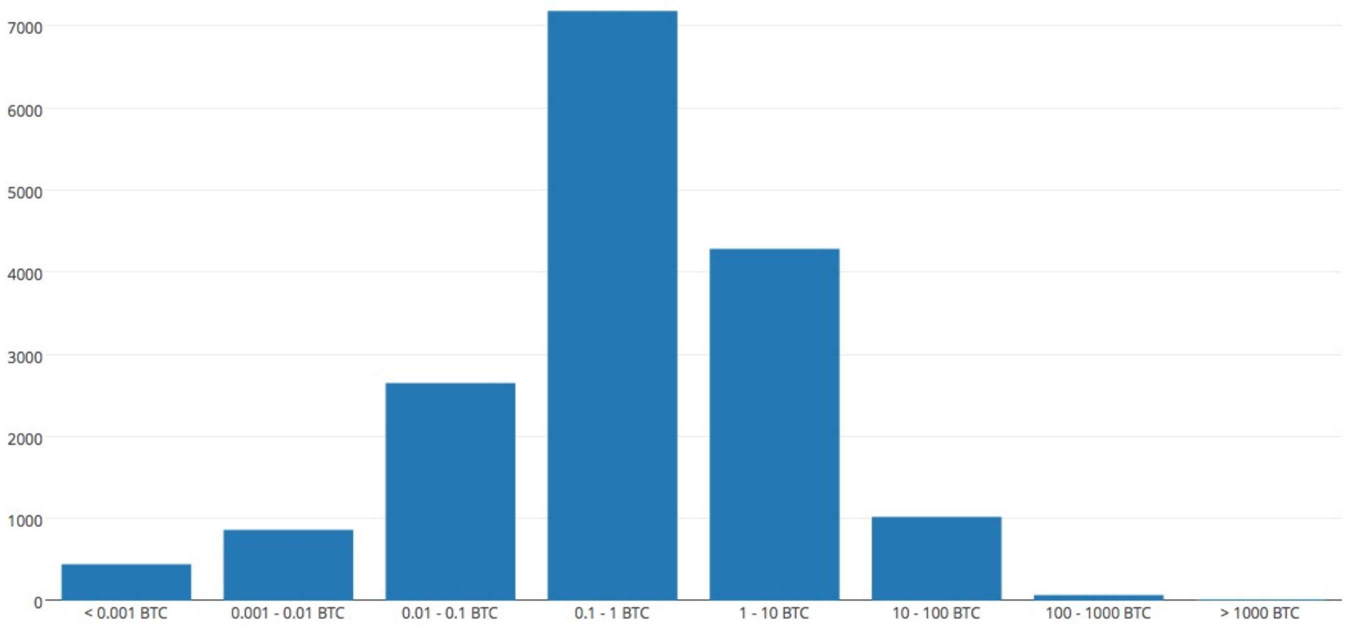
- Users who have received bitcoins from a service that requires some form of identification can use mixing service in order to spend anonymously[25].

- User who have bitcoins originated from suspicious activity can mix their bitcoins before seding money to a money exchanger who requires identification

- Makers can earn by making their bitcoins available to other users of the mixing services.

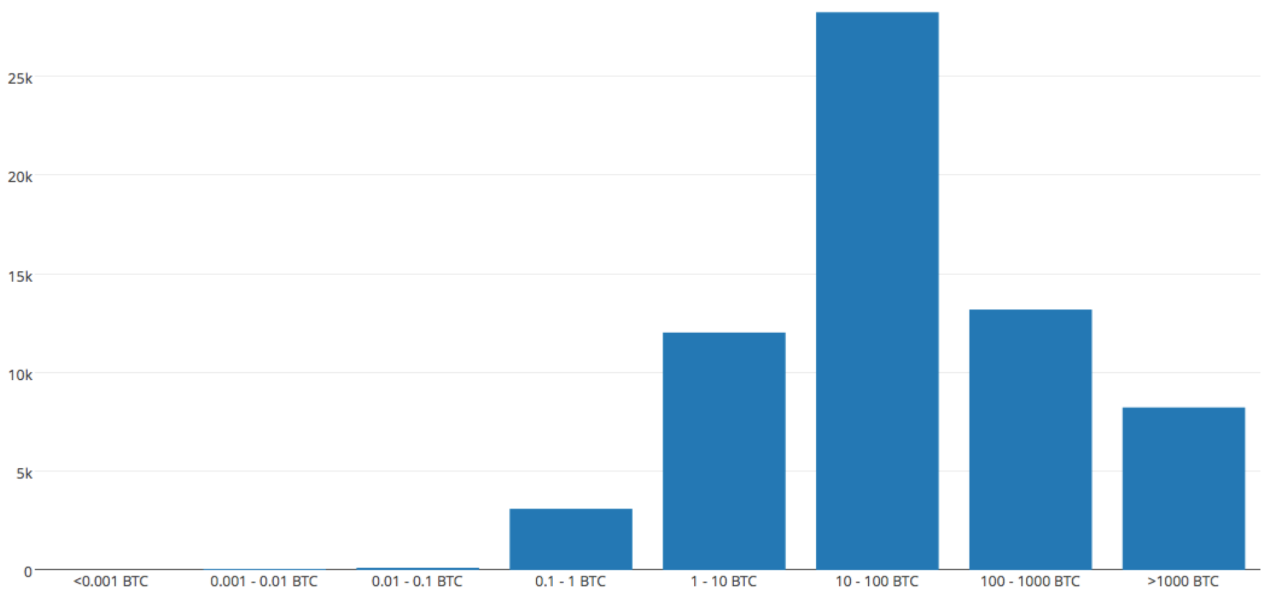
In order to understand how Joinmarket is used, we analyzed the funding transactions by dividing them first and foremost by amount sent. According to the following graph, it is evident that most of the transactions have a value between 0.1 and 1 BTC.

Number of funding transactions divided by amount sent



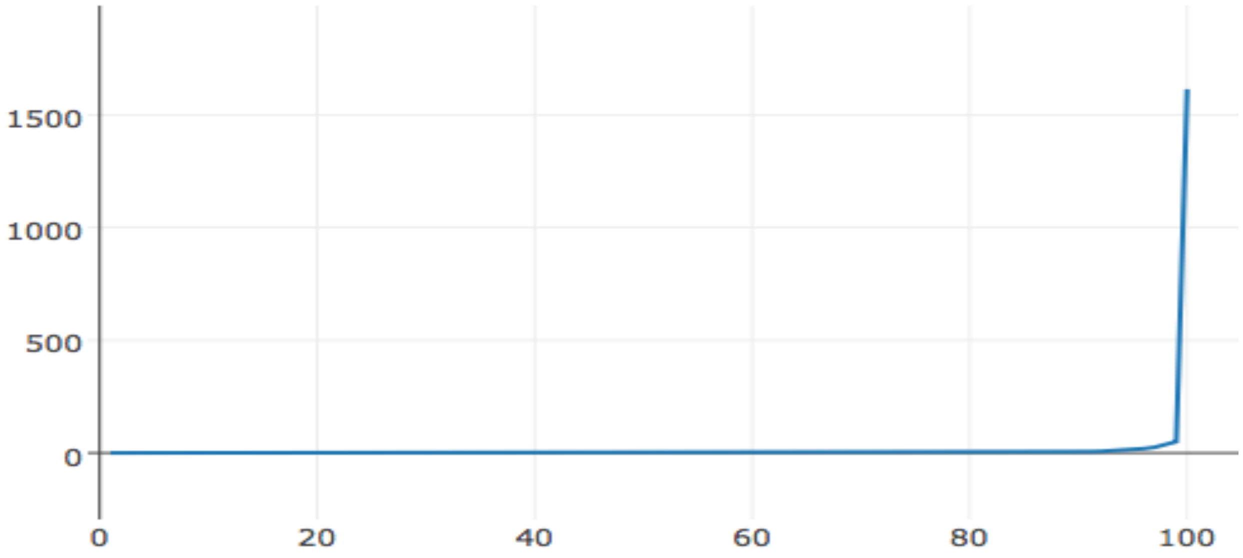
[25] The mixer can either return the bitcoins directly to the final recipient or have them temporarily passed on to a second anonymous wallet ,managed by the same user, to then be spent at a later time.

Number of mixed bitcoins divided by the value of funding transactions



Now let us analyze the percentile curve in regards to the value of the funding transactions. It is interesting to notice how a small group of transactions with a very high value (some above 1000 BTC), represent on their own a large part of the total sum of input. Just 1% of the most "expensive" transactions represent around 40% of the total bitcoins inserted in the system.

Percentile curve in regards to the value of the funding transactions



Using P-Flow's analysis functionalities on the outflow of these transactions, one can notice how this capital is used primarily as *makers* in the JoinMarket system: it is therefore possible to hypothesize that these cases are most similar to the third type of use identified above.

Now let us examine Korea's case by using P-Flow to analyze the sample of transactions attributed to the country.

Considering that:

- Most of the funding transactions are used as *takers* in the Joinmarket system.
- Based on the number of individual wallets identified thanks to the spending pattern analysis[26], and in relation to the number of attributed funding transactions, it can be hypothesized that every user has carried out on average 2/3 mixing transactions.
- There is a large peak in funding transactions but not such a large peak in outflow transactions.

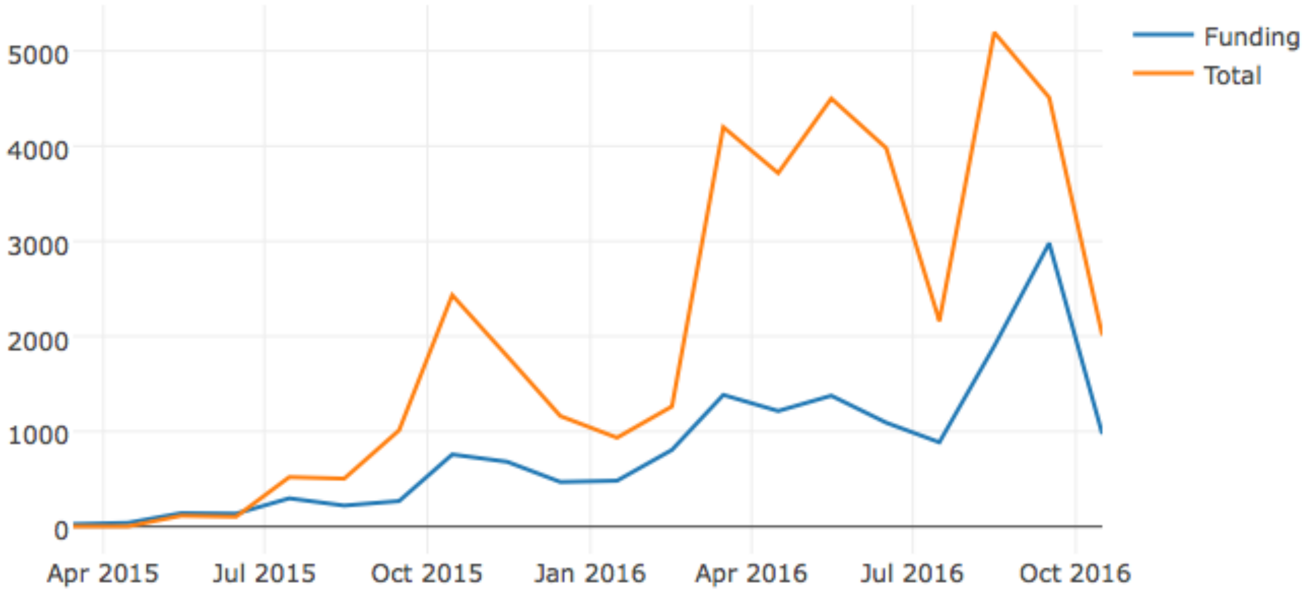
It can be speculated that Joinmarket is more widespread in Korea than in any other country for both impromptu cases and those associated with the second identified use case.

Conclusions

Based on the data collected it is possible to establish that, as of today, Joinmarket is not used systematically on a large scale for money laundering and the number of mixing transactions is still an insignificant part of the total sum of Bitcoin traffic. Nonetheless, the use of this system is visibly growing. This is evident in the following graph which illustrates the monthly evolution of JoinMarket transactions carried out.

[26] In this analysis we have considered the cluster of source addresses, specific pattern transaction of each client and the assumption that users are incline to prefer moving patterns with lower fees.

Number of monthly JoinMarket transactions (*funding* and total) carried out



In conclusion we can assume that the release of systems that will make JoinMarket more user friendly (such as a graphical interface and more wallets support, natively or via plugin) will make JoinMarket more popular and widespread.

This in turn will lead to a higher level of privacy for all of JoinMarket clients, including those interested in carrying out anonymous and sporadic money laundering activities which are difficult to trace without the right tools.

Neutrino S.r.l. An innovative startup founded in 2016 by a team with over ten years' experience in cyber security: a cyberlab dedicated to the research of innovative solutions to the new and complex challenges facing the security sector.

P-Flow, the first project developed by Neutrino, provides actionable insight on the blockchain and bitcoin network, offering to all of those interested in virtual currency, information which would otherwise not be accessible.

For more information: info@neutrino.nu