

# How Local Are Ripple Transactions

# Etienne Caquot

School of Computer and Communication Sciences

Decentralized and Distributed Systems lab

Semester Project

June 2019

Responsible
Prof. Bryan Ford
EPFL / DEDIS

Supervisor Cristina Basescu Kelong Cong EPFL / DEDIS

# Contents

1	Intr	roduction	1
2	Rip	ple	1
	2.1	The Ripple Network	1
	2.2	Transactions	1
	2.3	Main actors	2
	2.4	Main operations	3
	2.5	Nodes and Consensus	4
3	Estimating wallet geographical position		
	3.1	Using nodes	5
		3.1.1 Ripple crawler	5
		3.1.2 Connect to existing nodes	5
		3.1.3 Running our own node	6
	3.2	Running a gateway	6
	3.3	Nodes positions	7
	3.4	Using transactions data	8
		3.4.1 Looking at credit links	8
		3.4.2 Evolution of credit links	10
		3.4.3 Looking at currencies	14
4	Lim	itations	16
5	Con	aclusion	16
6	Further work		16
7	Rep	produce results	<b>17</b>
References			18

# 1 Introduction

We already know that economic activity is strongly correlated with distance: the closer two companies are, the higher the flow of transactions between them (all else equal). The question is, do blockchain transactions show the same pattern? We base our work on the Ripple blockchain to investigate the answer to this question.

# 2 Ripple

We live in an era we where we can have news from all around the globe nearly instantly. Despite being able to communicate overseas easily we still struggle to transfer money quickly. Ripple's goal is to enable money transfers as fast as we can exchange information. To achieve this, Ripple aims at connecting people through a network and to use this network so that money transfers from one user to another can be executed rapidly.

# 2.1 The Ripple Network

Based on an IOweYou credit network, Ripple network can be visualized as a directed graph where the vertices are wallets and the edges represent the credit link between those wallets. Wallets are used by entities (physical or moral) and attest the amounts this entity has in its possession. Non-negative weight on an edge from A to B indicates the amount that B owes to A. Ripple enable users to set a limit on the amounts that others owe to you. By default, the credit on a link is upper-bound by  $\infty$  but the wallet owner (A in our example) can customize it. There can be several links between two wallets and each link between A and B handle different currencies. We have two types of currencies in the Ripple network: XRP that is the native currency and all other currencies that are also called issued currencies. All currencies on credit links are issued currencies. Moreover, each wallet has a non-negative amount of XRP that is used to pay fees for transactions and can also be used as a currency.

Figure 1 depicts a possible fragment of the Ripple network. We have two gateways Bitstamp and Gatehub along with 6 users. Here, the credit link Gatehub  $\rightarrow$  Alice denotes that Gatehub owes Alice 1200 EUR, and such credit is bounded by 3000 EUR. The credit link Bitstamp  $\rightarrow$  Charlie tells us that Bitstamp owes Charlie 900 USD, and there is no upper limit on this credit link.

#### 2.2 Transactions

There are two types of transaction in the Ripple network: direct XRP-payments and path-based transactions.

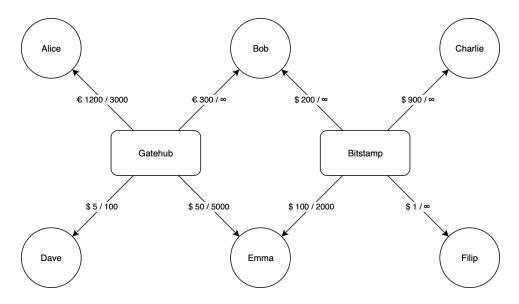


Figure 1: Ilustrative example of the Ripple network

A direct XRP payment enables two wallets to pay each other even if they are not connected in the Ripple network. Basically, the XRP balance of the wallets is debited (and respectively credited) with the amount to transfer. As the name suggests that this type of transaction only involves XRP.

All other currencies are exchanged over path-based transactions. Those transactions use a path of credit links that connects the sender to the receiver in order to make the transfer. Ripple has no way to enforce path based payments. It only records the amounts owed by one wallet to one other. A payment from A to B is only possible if there is a path that connects A to B. Moreover, A can only make a successful IOU payment to B if the payment value falls within the credit balance allocated on this path. For a given transaction, the total amount can be split into several paths. This enables to make payments even if there is not a single path that has enough credit capacity. In our example of Figure 1, let's assume that Dave wants to pay Filip 5 USD. A possible path would be Dave  $\leftarrow$  Gatehub  $\rightarrow$  Emma  $\leftarrow$  Bitstamp  $\rightarrow$  Filip. When the transactions succeed the credit links are updated. In our situation, the links Dave  $\leftarrow$  Gatehub and Emma  $\leftarrow$  Bitstamp are decreased by 5 USD whereas Gatehub  $\rightarrow$  Emma and Bitstamp  $\rightarrow$  Filip are increased by 5 USD.

The combination of direct XRP payments and path-based transactions is natively supported in the Ripple network and they are atomically executed as a whole thanks to the Ripple Consensus (see 1.5)

#### 2.3 Main actors

In addition to standard wallets, there are wallets with special characteristics.

Gateway A gateway is a well-known established business that aims to help new users to enter the ripple network in an authenticated manner. Gateways are the Ripple equivalent of banks in the physical world. Their wallets maintain high connectivity. A newly created Ripple wallet that does not initially trust any existing wallet can create a credit link to a gateway and thereby interact with the rest of the network before creating direct links to other wallets.

Market Maker A market maker is a wallet that receives a certain currency on one of its credit links and exchanges it for another currency on another credit link, charging a small fee. Market Makers enable transactions where senders and receivers hold different currencies.

## 2.4 Main operations

To compose a path from a sender to a receiver, there are two possible operations Rippling and Exchanging.

Rippling Rippling is the process of moving debts. In the XRP Ledger, Rippling defines the redistribution of credit links amounts between multiple connected parties who have trust lines for the same currency. Rippling is an essential part of issued currencies because it allows users who don't have a credit link between them to send issued balances to each other using a passive third user who serves as an intermediary. Rippling only exist along the paths of a payment. Rippling can only occur between two credit links that belong to the same wallet and have credit in the same denomination.

Exchanging Exchanging is the process of converting one currency to another. A wallet that is willing to exchange one currency for another creates an exchange offer for this pair, this wallet will then be identified as a market maker. The typical exchange offers are from issued currencies to issued currencies (eg. USD/EUR). However, it is possible to offer exchanges from/to XRP. There are some path-based transactions that use exchanges offered by market markers involving XRP, then the transaction becomes sort of both path-based and 'direct' in the sense that is used both XRP and issued currencies.

Back to Figure 1, we now have Alice that wants to pay 100 USD to Charlie. Alice only has credit links only in EUR, however, suppose Bob publishes an offer EUR/USD with a rate 1 USD = 0.9 EUR. Bob plays the role of a Market Maker and enables Alice to pay Charlie by spending 90 EUR. The links are updated as follow: Alice  $\leftarrow$  Gatehub is decreased by 90 EUR, Gatehub  $\rightarrow$  Bob is increased by 90 EUR, Bob  $\leftarrow$  Bitstamp is increased by 100 USD and finally Bitstamp  $\rightarrow$  Charlie is increased by 100 USD. Notice

that the links of Gatehub and Bitstamp are shifted respectively by 90 EUR and 100 USD due to Rippling.

#### 2.5 Nodes and Consensus

The XRP network presented in the previous subsection is contained in a ledger. Every few seconds the XRP ledger has a new version which increments the index of the ledger by one. To operate properly and make the ledger progress with the arrival of new transactions, the XRP ledger relies on a set of servers. When a wallet wants to make a payment, it sends the transaction information to one server and it needs to operate according to rules so that the transaction may be applied and executed in the next version of the ledger.

The peer-to-peer XRP ledger network consists of many independent servers that receive and process transactions. The servers need to relay these candidates transactions to the other servers so that it can be applied to the ledger and validated by the validators. The server software that powers the XRP Ledger is called rippled[1], it enables servers to communicate in a peer-to-peer network. The XRP ledger servers are also called nodes. There are two main types for nodes in rippled:

**Stock server** Follows the network with a local copy of the ledger, propagates transactions and answers questions about the ledger.

Validating server validator for short - Participates in Consensus, a process designed to accept or deny a transaction based on some rules (and does everything a stock server does)

Servers propose transactions to apply to the current version of the ledger, this could lead to different versions with different content for the same ledger index. Of the many candidates, only one can become validated and thus there is only one valid ledger for each index. When a set of transactions can be applied to the ledger, and that validator has agreed on this set a new version of the ledger is created and it is marked as *validated*. In a perfect scenario, each new transaction would be known to each node allowing each to consider the same set of transactions to apply to the current version of the ledger. However, transactions take time to propagate and servers are not aware of all transactions. To answer this problem, the XRP Ledger uses a process called Consensus to ensure that all nodes have the same last version of the validated ledger.

Consensus comes in three steps:

**Deliberation** Nodes iteratively propose a set of transactions to apply to the prior ledger. Nodes receive proposals from other trusted nodes and adjust their set of transactions. When a node believes enough proposals agree, it applies the corresponding transactions to the prior

ledger according to the ledger protocol rules. It then issues a validation for the generated ledger.

Validation Nodes decide if they fully validate a ledger, given the validations received from trusted nodes. Once the needed amount of validations for the same ledger is reached, that ledger is fully validated and become irreversible.

**Preferred Branch** Nodes determine the most popular working branch of ledger history. There could be conflicts due to asynchrony, network difficulties and nodes could validate different ledgers for the same index. The nodes use the ledger ancestry of trusted validations to get back to the network's preferred branch.

# 3 Estimating wallet geographical position

The goal of this project is to investigate whether there is some locality in Ripple transactions. After having a brief overview of how the Ripple network operates we can now start to explore different techniques to estimate wallets geographical positions.

## 3.1 Using nodes

Nodes are the backbone of the Ripple Network. Our first focus was to dig into the possibilities that nodes could give us.

#### 3.1.1 Ripple crawler

We found a crawler[2] that crawl ripple nodes, and could give us geographical positions. At first, it seemed a good idea that could give us estimates of the positions of the wallets. However, as mentioned in the Ripple background section, the wallets and the nodes are not linked directly. This crawler finally turns out to be useless since it could only give us information about nodes and not on wallets.

#### 3.1.2 Connect to existing nodes

Another angle to tackle the problem was to find out if clients have preferred nodes they connect to (maybe the closest ones?). To explore this idea, we found the list of all nodes and their IP address of the ripple network using the Ripple Data API[3]. We used a Python implementation Data API calls to query the Ripple Data API to get all nodes IP and then get their URL. Then for all those nodes, we tried to connect to them on port 55233 (most used port) to see if we could access the Ripple Ledger. It turns out that we were not able to connect to any nodes. It seems that users that run a node

use it as their own copy of the ledger, and so don't have to rely on someone else. The only way we managed to connect to a node and query the ledger was to connect to the public Ripple node maintained by Ripple Inc.

#### 3.1.3 Running our own node

While trying to estimate wallet position, we thought about running nodes in various locations. Since we know nodes are responsible for propagating transactions, having several nodes in several parts of the world could enable us to hear transactions and tell that the sender is probably closer to the location of the first node who learns about the transactions. The problem is that it does not seems that in the peer protocol nodes connect to other nodes that are close to them. So the idea that transactions propagate along with geographic positions does not seem to fit here.

However, if we could get client connection to our nodes to send their transaction then we could know thanks to IP address position of the actual sender of the transactions. But how to get users sending their transactions to us? The only way would be to advertise our node since users can choose freely the node they want to use to send their transaction. We believe that we would get only a few users connection to our servers if any. In addition, we looked at the source code of some popular ripple clients and from what we could learn it seems those clients connect to the public nodes maintained by ripple. This promising idea finally turned out to be left out, as most of the users go through the same node to propagate their transactions then we would probably retrieve any information from running our own nodes.

#### 3.2 Running a gateway

Gateways are highly connected wallets in the Ripple network, they are the door between the Ripple network and the real world. A new wallet in the ripple network can create a credit link with a gateway then send (in the real world) some money to this gateway and then the user will have this credit available on the link with the gateway. The idea is that when a user goes on the gateway platform, creates a wallet and send money to a gateway, then the gateways are able to link wallet identifiers to a real-world person with their location thanks to either the IP address or either the postal address (usually gateways ask it due to regulations). Hence our first idea was then to look at whether it was easy to run a gateway and have clients connecting to us. It turns out the Ripple documentation[4] gives the steps to run a gateway but due to the complexity, it is outside of the scope of this project and left as further work.

# 3.3 Nodes positions

As explained above, the Ripple Data API[3] enables us to know all nodes and their IP. We can then process the IP addresses obtained and map each node to a country.

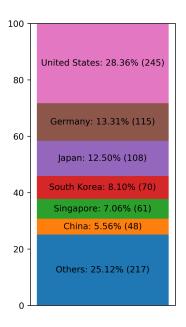


Figure 2: Repartition of Ripple nodes

Figure 2 on the left shows the repartition of the 868 nodes we were able to locate over 1018 that are known to the API. For each candidate country, we show the percentage and in parenthesis the actual number of nodes situated in the country. We regrouped all countries that count for less than 5% each in the 'Others' category. We used a stacked bar chart because it helps to visualize proportion and compare values.

We see that the country that has the most node is the United States then comes Germany and Japan. Together they represent more than 50% of all nodes.

We want to see if the number of nodes in a country is correlated to how this country's currency is used in the Ripple ledger.

Figure 3 shows the volume, in XRP, for

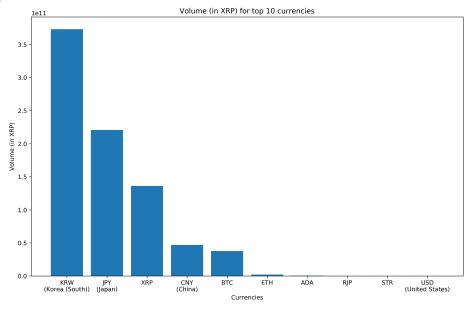


Figure 3: Top 10 currencies in XRP volume

the 10 currencies with the highest volumes. We chose to represent the result as a bar chart to ease the comparison between currencies. We see that KRW, JPY, and CNY have a much higher volume that USD but on Figure 2 we saw that the United States had more nodes. We can also say that Japan has more nodes than South Korea but that KRW has higher volume. So our assumption that we could extract some locality information from the position of the nodes does not seem to fit here since the number of nodes in a country does not reflect how this country's currency is used in the Ripple Network.

#### 3.4 Using transactions data

Let's try a different approach to gain further information about the localization of Ripple transactions. What could be better than looking directly at the actual transactions data? The data we used was provided by Pedro Moreno-Sanchez who is one of the publishers of the paper Mind Your Credit: Assessing the Health of the Ripple Credit Network[5] which motivated us for this semester project. We have 3 files with transactions data, one that contains transactions from January 2017 to August 2017, one that has transactions from 2013 to 2016 and another file with some missing transactions from 2013 to 2016.

In the next subsections we want to look at path in the Ripple network and sender/receiver currency pairs, the file with data from 2013 to 2016 was not really useful for this step because we didn't have any information about the path of transactions and it shows contains only the receiver's currency not the sender's.

We also did a little bit of cleaning on the files that were usable because for some transaction the sender's currency was missing so we had to infer it from the path.

#### 3.4.1 Looking at credit links

For the reader that will look at the Jupyter notebooks, notice that credit links are called trust lines in the Ripple Data API[3]. Thanks to this API and our additional work, we could recover all credit links for the wallets that appear in our transactions data. We then computed a few statistics on those credit links.

We have that 43.34% of wallets have at least one credit link. Over the wallets that have at least one credit link 69.59% of them have one with a gateway (that is 30.17% of all wallets). Overall credit links, 69.24% of them are with gateways.

It seems that gateways play an important role in credit links. Can we get any information from those credit links to help in our goal of localizing transactions?

As we explained before, each credit link is associated with a currency. We looked for each wallet how many different currencies link them to gateways. For the wallets that have credit links with gateways, we have that 75.95% of them that are linked to gateways via only 1 currency and we have 19.94% that are connected via only 2 currencies. From those results, it seems that wallets tend to keep a small number of currencies on links with gateways. Why wallets would behave like this? One hypothesis we made to answer this question was that wallets tend to use their local currency in the Ripple network. However, this assumption can't be easily verifiable and only a few of all wallets are concerned by this hypothesis.

We now present another look at credit links where we looked at the currencies that linked the sender and gateway on the first hop of a transaction (and respectively the gateway and the receiver on the last hop). A gateway is a first hop (respectively a last hop) if it is the first wallet after the sender on the path from sender to receiver (respectively the last wallet before the receiver).

The stacked-bars in Figure 4 show as a percentage the currencies of the link where the gateway appears as first or last hop. Each subgraph represents a gateway, we distinguish the first hop on the left and last hop on the right. On top of each subgraph, you can see the name of the gateway along with a parenthesis with two countries. The first country is the supposed fiscal location of the gateway found on this website[7] and the second country is the where the gateway's servers are situated. We got the server's location thanks to the Ripple Data API[3] that could give us the domain of the gateways then from those domains we got the corresponding IP addresses and mapped them to a country. What is important to look at here is the colors, each color corresponds to a currency.

On this graph, we can see that most of the subgraph contains one or two colors. This means that gateway seems to be active on at most 2 preferred currencies. We can also note that there seems to be a correlation on the location of the gateway and its currencies. For example, the first one 'Ripple Fox', supposedly situated in China, seems to process mainly CNY which is the Chinese currency. The same applies to Ripple Trade Japan where the servers of this gateway are situated in Japan and it processes only JPY the currency of Japan.

It seems that gateways tend to specialize. We know they are financial institutions that are looking for profit, so if there must be an explanation as to why they seem to stay local. We think that people try to connect to the gateway that will offer their local currency and that hence that is situated the closest to them. Again we don't know if this is true, but otherwise, gateways will tend to offers more and more currencies to attract more clients in order to have higher profits.

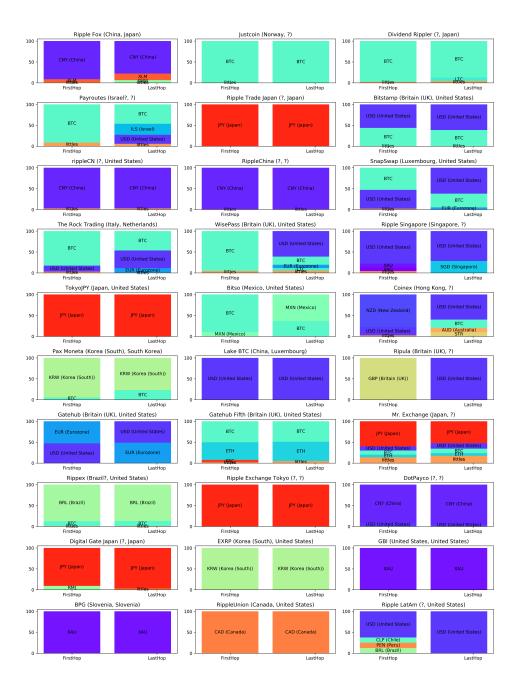


Figure 4: Currency of transaction for first and last hop per gateway

#### 3.4.2 Evolution of credit links

Following up with the results we obtained on credit links, we wanted to investigate the behavior of those lasts over time. The previous analysis was just over the credit links we could get from the Ripple Data API[3]

and does not reveal evolution in time. To palliate this problem we looked at transaction paths which uses credit links. We only treated transactions from 2017 because it was the year with the most information about transaction paths. We computed for each currency and for each month the wallets appearing as first or last hops. For those wallets, we then computer the total volume and number of transactions routed trough them per month. We only showed a wallet if it had a month were the amount was significant enough (1% of the biggest value over the year). We plotted the results for CNY, JPY, USD and EUR the 4 top currencies in terms of number of transactions. All graphs are visible on the Github repo[?]. We show the ones for USD in this report, as they are the most interesting.

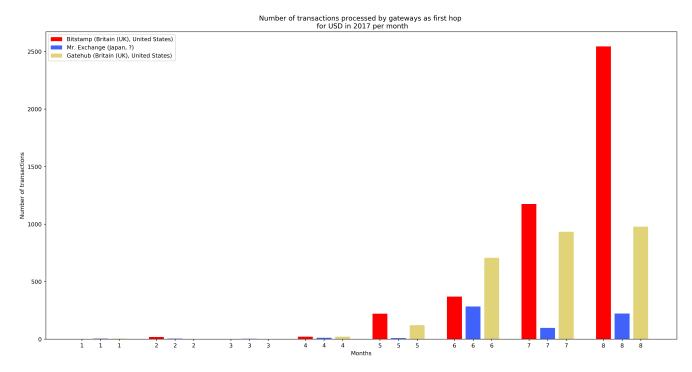


Figure 5: Number of transactions routed trough wallets as first hop for USD in 2017

Figure 5 shows the number of transactions, with sender using USD, routed trough each wallet and Figure 6 (on the next page) shows the same transactions but in terms of volume. On the x-axis you can see numbers that represent the month of the year. There is one tick by wallet and by month. We see that the number of transaction increase over time, we can see this increase in the volume plot to but it seems let smooth. It is also interesting to notice the difference between the wallets that appears in Figure 5 and Figure 6. Some wallets are insignificant in terms of number of transactions but appears in the volume graph. You can also see that for wallets that

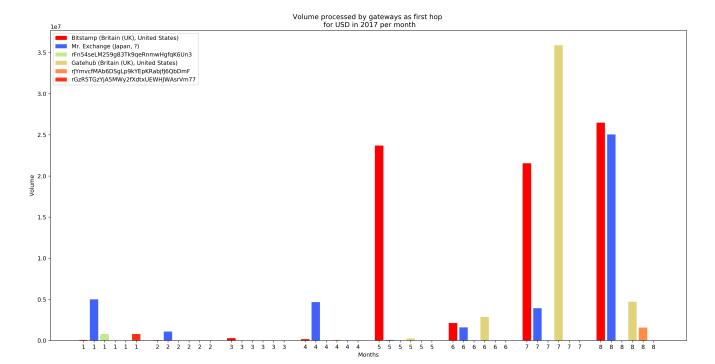


Figure 6: Volume routed trough wallets as first hop for USD in 2017

are gateways we showed their names and their respective fiscal and servers geographical positions (as in Figure 4). We saw in this last figure that gateways seemed to specialized but remark here that none of the gateways for USD are fiscally situated in the USD.

Let's now look as the same data but for last hops in USD over 2017. The x-axis of Figures 7 and 8 (on the next page) represent also the time progression but instead of grouping by month as in Figures 5 and 6 we group by gateway and for each gateway have the eight months.

The results are similar to what we could get from the first hops. However this time there is much more wallets that are not know as gateways. We see that the gateway Bitstamp is the most popular wallet for first and last hop. For the last hop the gateway Mr. Exchange that does appear in the number of transactions plot is absent form the graph with the volume. In this graph we have only gateways with their fiscal location outside of the United States. After looking at the plot for the other currencies, it seems that USD is the only currency to behave like this.

In this subsection, we wanted to dig into the evolution of credit links. From what we see is seems that the most popular wallets, that turns out to be gateways, appears usually in each month over 2017. We can conclude

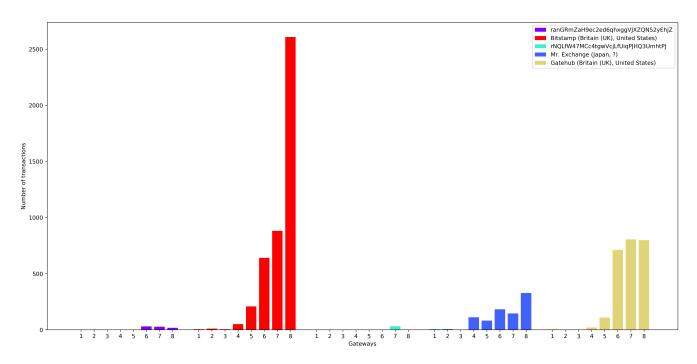


Figure 7: Number of transactions routed trough wallets as last hop for USD in  $2017\,$ 

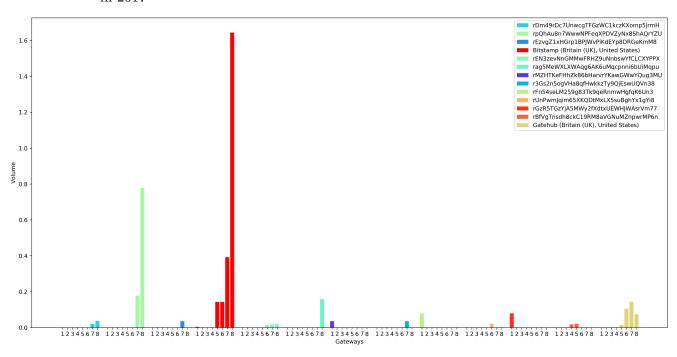


Figure 8: Volume routed trough wallets as last hop for USD in 2017

that credit links seems quite regular for the wallets with high activity in the Ripple Network. However we also found wallets, that are not gateways, which appears to have high activity. We didn't have time to investigate further on those wallets and this is left out a further work.

#### 3.4.3 Looking at currencies

Next step in our analysis of transaction locality. We looked at the sender and receiver currency pair and if there were any correlation. We expect to find here a lot of cross-currency payment in the Ripple Network. We looked at all transactions in our data and counted the number of transactions that correspond to each possible sender and receiver currency pairs. The heat-

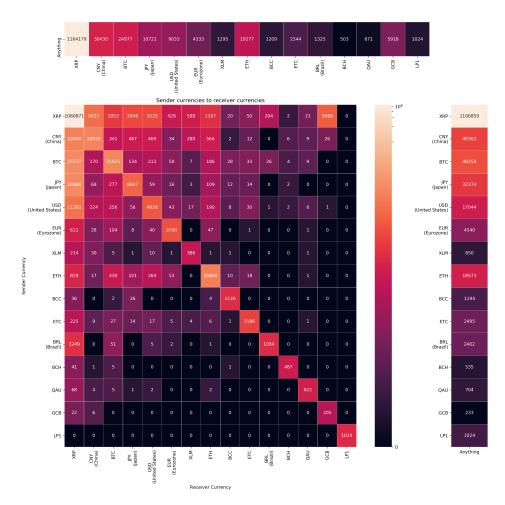


Figure 9: Sender Receiver Currency Paris in number of transactions

map in Figure 9 shows the result for the fifteen most popular currencies in terms of the number of transactions. The x and y-axis are currencies where

the most popular currency is on the left (respectively on the top). Each cell of the heat-map shows the number of transactions for the corresponding sender/receiver pair. We used a logarithmic scale because when we first plotted this heat-map everything was black but the top left cell. This is because there are much more transactions that goes from XRP to XRP than any other pair of currencies. Again what is important is not the actual numbers but the colors. The more bright a cell is then the more there are transactions for the corresponding currency pair. The diagonal pops out from this visualization, what does that mean? It means that there is a good amount of transactions that start and end in the same currency. Said differently, we don't have that many cross-currency payments. There are more same currency payments than cross-currency payments. We can also note that transaction from one currency to XRP seems to be quite common too, at least for the top 5 currencies without considering XRP. You might have noticed the top and left heat-maps. They represent respectively the number of transaction from any currency to a given currency and the number of transaction from a given currency to any currency.

We can also look at the heat-map for the same currencies but instead of the number of transactions, we have the volume in XRP. The heat-map in Figure 10 shows the result for the fifteen most popular currencies in terms of volume. The scale is also logarithmic. The results are quite similar to the previous ones. As for the number of transactions, the diagonal pops out from this visualization. Those results reinforce that there are a high number of transactions that start and end in the same currency.

# 4 Limitations

We could come up with the result that there is some locality in ripple transactions, unfortunately, our analysis is not really precise. It is hard to map a given wallet to a position in the world. The limitations that stood in front of us where actually the protocol that is used by Ripple, since in the transaction data we only have a wallet identifier, no IP address or any information that could help us to link a wallet to a position. It is really hard to de-anonymize or get any personally identifiable information about the ripple wallets given how the transactions are propagated.

#### 5 Conclusion

We tried different approaches to estimate if there is locality in Ripple transactions. We first tried to look at rippled nodes, unfortunately, we were not able to gain much information about the locality of transactions.

After those setbacks, we looked at the data we had about transactions. The results were more promising here. We have found that cross-currency

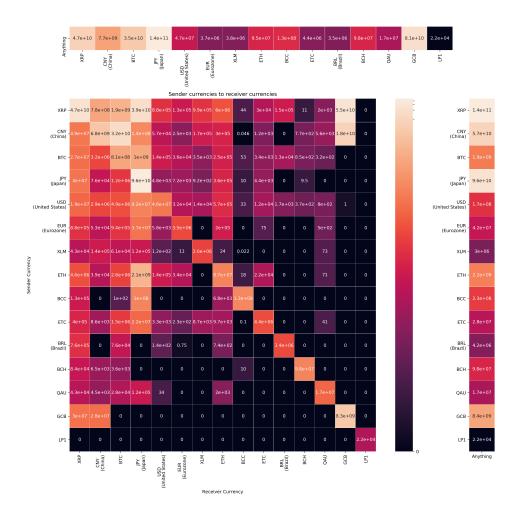


Figure 10: Sender Receiver Currency Paris in volume (XRP)

payments do not seem really developed and that gateway tends to specialize. With our assumption that users use their local currency, it seems that the closer people are the more there are susceptible to interact in the ripple network. That the finest analyses we can get so far with the information we have. This work highlights the efforts from Ripple to make transactions and wallet owners unidentifiable.

# 6 Further work

To gain further information about localization of ripple transactions, it could be a good starting point to try to run rippled servers. As we explained in part 3.1.3, with good advertisement it might be possible to have clients connecting to those nodes.

In the analysis of the evolution of credit lines, we have found wallets that are not referenced as gateways in the API that appears as last hops of transactions. It could be a good idea to investigate further on those wallets they might turn out to be useful.

You can also read Listening to Whispers of Ripple: Linking Wallets and De-anonymizing Transactions in the Ripple Network by Pedro Moreno-Sanchez et al.[6] were they use different heuristics to gain information about wallet and transaction locality in Ripple.

# 7 Reproduce results

How to run the project

#### References

- [1] rippled https://github.com/ripple/rippled 2012, 2013, 2014 Ripple Labs Inc.
- [2] XRPL Network Crawler https://github.com/ripple/rippled-network-crawler Ripple Inc. 2018
- [3] Ripple Data API v2 https://developers.ripple.com/data-api.html
- [4] Become an XRP Ledger Gateway https://developers.ripple.com/become-an-xrp-ledger-gateway. html
- [5] Pedro Moreno-Sanchez, Navin Modi, Raghuvir Songhela, Aniket Kate, Sonia Fahmy: Mind Your Credit https://www.cs.purdue.edu/homes/pmorenos/public/ ripple-study-www.pdf
- [6] Pedro Moreno-Sanchez, Muhammad Bilal Zafar, and Aniket Kate: Listening to Whispers of Ripple https://petsymposium.org/2016/files/papers/Listening\_to\_ Whispers\_of\_Ripple\_\_Linking\_Wallets\_and\_Deanonymizing\_ Transactions\_in\_the\_Ripple\_Network.pdf
- [7] Know your gateway
  https://wipple.devnull.network/research/gateways.html Dev
  Null Productions LLC 2018-2019