Recent trends in NU delinquency

Author: Vivek Pandey
May 28, 2023

I have been analyzing NU delinquency trends for the past week or so. After a good amount of work which spanned looking at dozens of time series from across the systems, talking to a lot of you and drawing many many charts, I have a bunch of reasons which I think together explain the differences. in delinquencies. Raw data for most of the analysis is in these three sheets.

Summary

1. NU delinquency has been rising over past several cycles: From a min of **DPD30 20.01% in Jan 2nd cycle**, it has recently crossed **24% for Apr 2nd cycle** (projected DPD 30 numbers). So, there is an increase of ~4+pp.

2. The situation is complex with many moving parts, but to the best of my estimates and educated guesses, following are the dominant reasons for delinquency increase along with how much pp they explain out of a total of 4pp.
(a) **CTUF deprecation,** and staleness of the substitute network model**:** ~2 pp
(b) Retraining of **cumulative amount filter** [Not fully sure] ~1 pp
(c) Increased fraction of more delinquent Tachyon: 0.3pp
(d) Control set delinquency: 0.4pp

3. RTA delinquencies, which have always been higher than delinquencies of other lines, have not deteriorated any further during the times of increased NU delinquencies. Infact, RTA delinquents have gotten closer to other lines during Jan 2nd cycle vs Apr 2nd cycle.

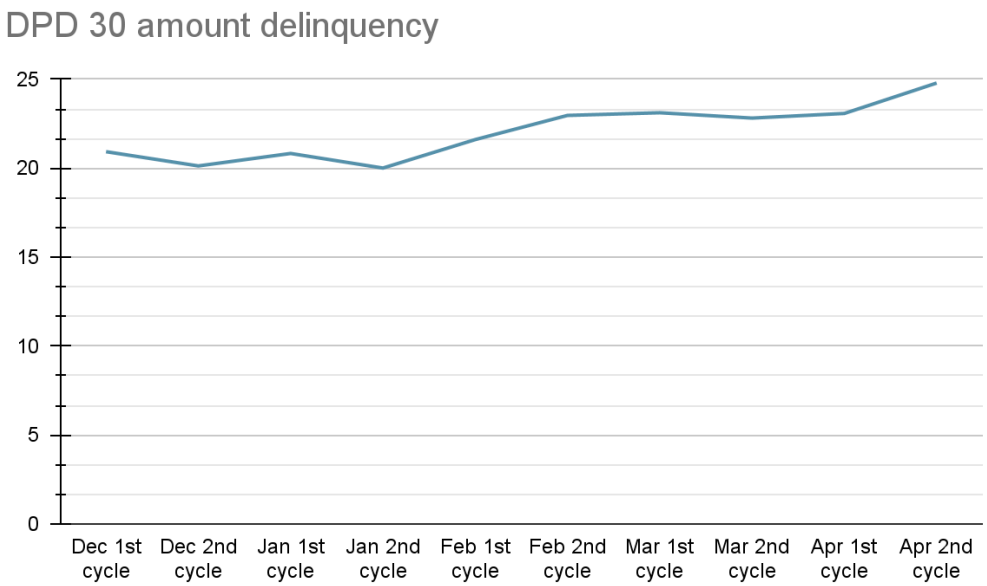
Note
This is the best fitting story I could construct based on the available data. If you have alternative explanations for the phenomenon backed by data, or else you can find big flaws in the analysis, please discuss with me.

Detailed analysis

While I have mentioned the finding in the summary, it was difficult to converge to that finding. So, we will go through everything step by step, and also look at some of the data, which should be interesting though it does not provide the reason for the delinquency increase.

Simpl wide delinquency

First let's look at headline, Simpl wide delinquencies:



Dec 1st cycle	20.93
Dec 2nd cycle	20.13
Jan 1st cycle	20.83
Jan 2nd cycle	20.01
Feb 1st cycle	21.61
Feb 2nd cycle	22.97
Mar 1st cycle	23.12
Mar 2nd cycle	22.82
Apr 1st cycle	23.08
Apr 2nd cycle (DPD27)	24.79

Note that Apr 2nd cycle delinquency is D27.
For the purpose of this document, we will focus on comparing the rise in NU delinquency between the Jan 2nd cycle (20.01) to Apr 2nd cycle (24.79), though Apr 2nd cycle number is D27.

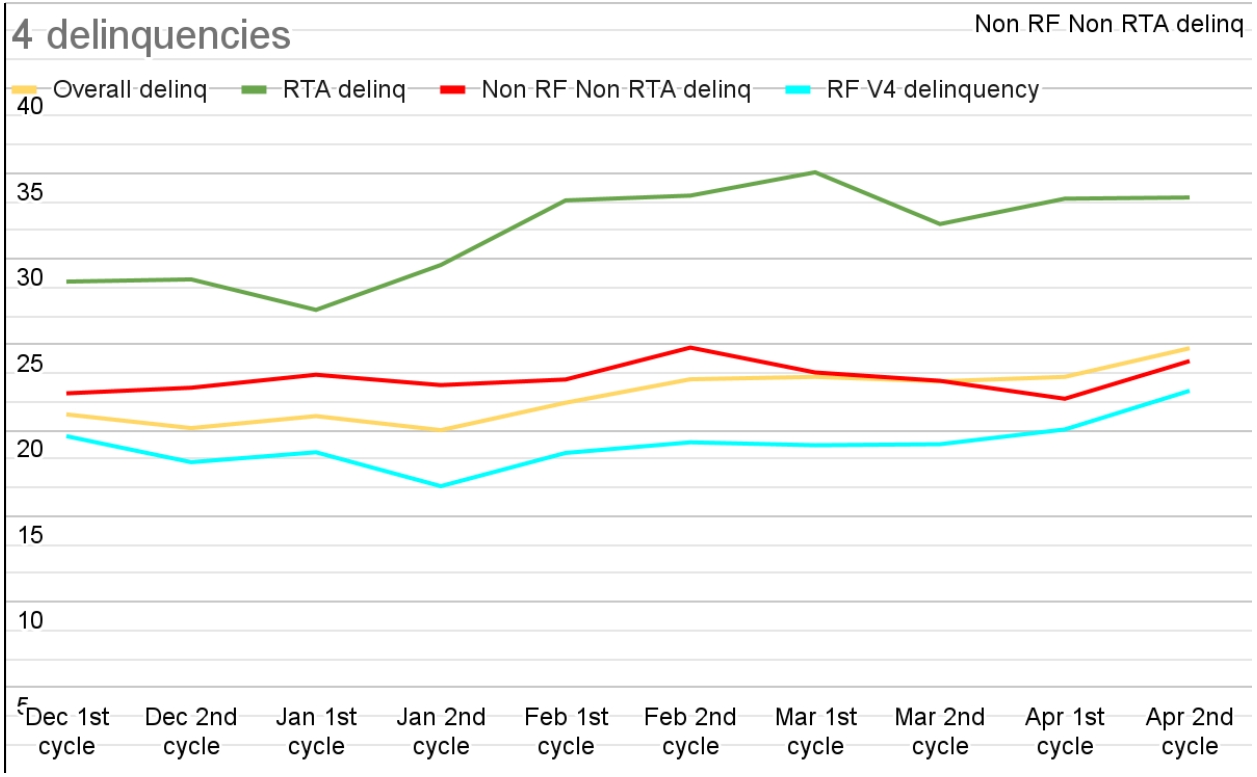
What is noticeable is:
1. There was a jump in NU delinquency Jan 2nd cycle -> Feb 1st cycle -> Feb 2nd cycle
2. There has been a noticeable jump Apr 1st cycle -> Apr 2nd cycle (though some of it will be bridged as D27 mature to D30)

In the rest of the document, we focus on Jan 2nd cycle vs Apr 2nd cycle numbers to have a focused analysis.

1. Delinquencies of various lines

Next we look at following four delinquency series:
1. Overall Simpl wide delinquency
2. Delinquency of RF V4 line
3. Delinquency of RTA lines (all RTA lines combined)
4. Delinquency of non RF non RTA lines (i.e. onboarding lines etc)

Note that (1) is a weighted average of (2), (3) and (4), weighted by the TPV of (2), (3), (4) respectively.



Comparing Jan 2nd cycle vs Apr 2nd cycle, here are the relevant data points

	Jan 2nd cycle	**Apr 2nd cycle (DPD 27)**
Overall	20.01	24.79
RF line	16.74	22.31
RTA lines	29.65	33.59
Remaining lines	22.64	24.04

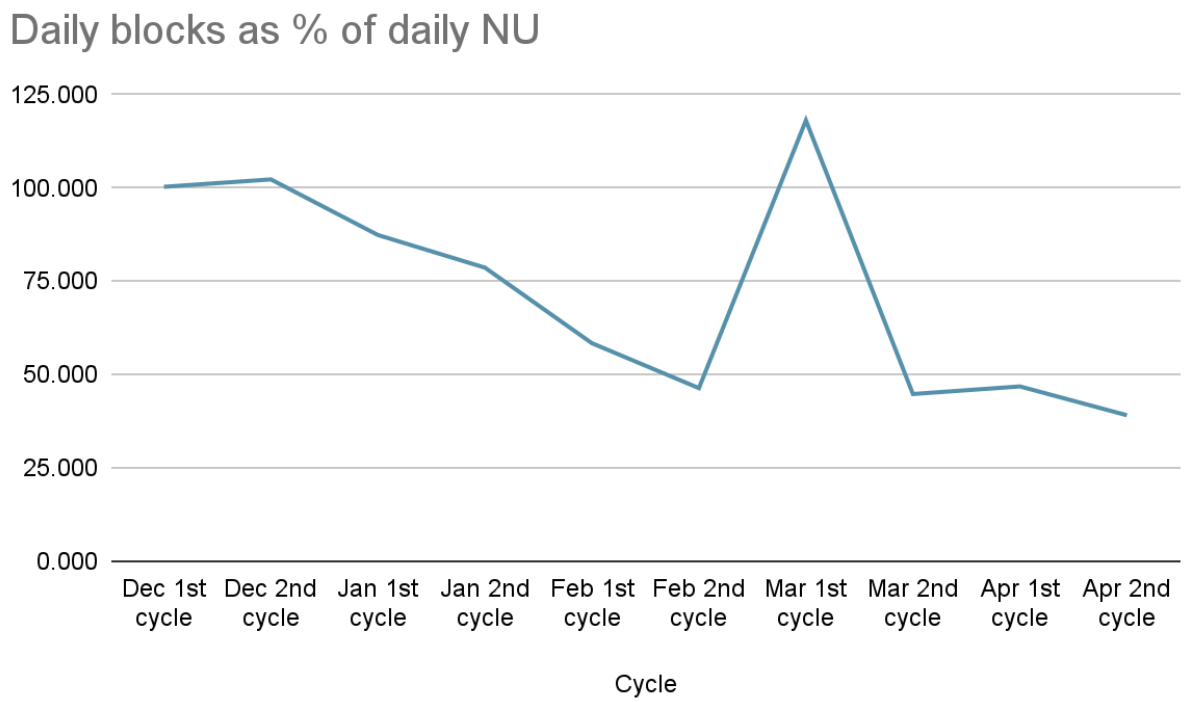
So, summary here is: Increase in delinquencies is not explained by increase in delinquency of some one line. However, the RF line has been hit the worst, and non RTA non RF lines have been hit the least.

2. Fraud blocking behavior

Next check the metrics of our anti fraud systems. This is where we "solve" the puzzle. There are two metrics we consider:

1. How many users are we blocking per day?
2. How many transactions are we denying per day?

Users being blocked per day



Each data point in above chart is x / y where

x = number of users we blocked in a given cycle divided by the number of days in the cycle
y = number of new users who transacted in the given cycle divided by the number of days in the cycle

Note that a vast majority of the blocked users are non transacting, so it is not alarming to have numbers 50%, 60% or even 100%+

The peak in Mar 1st cycle is because some filters were unusually active during the ATO phase.

We see a steady decline in daily blocks / daily NU from Dec 2nd cycle to Feb 2nd cycle.

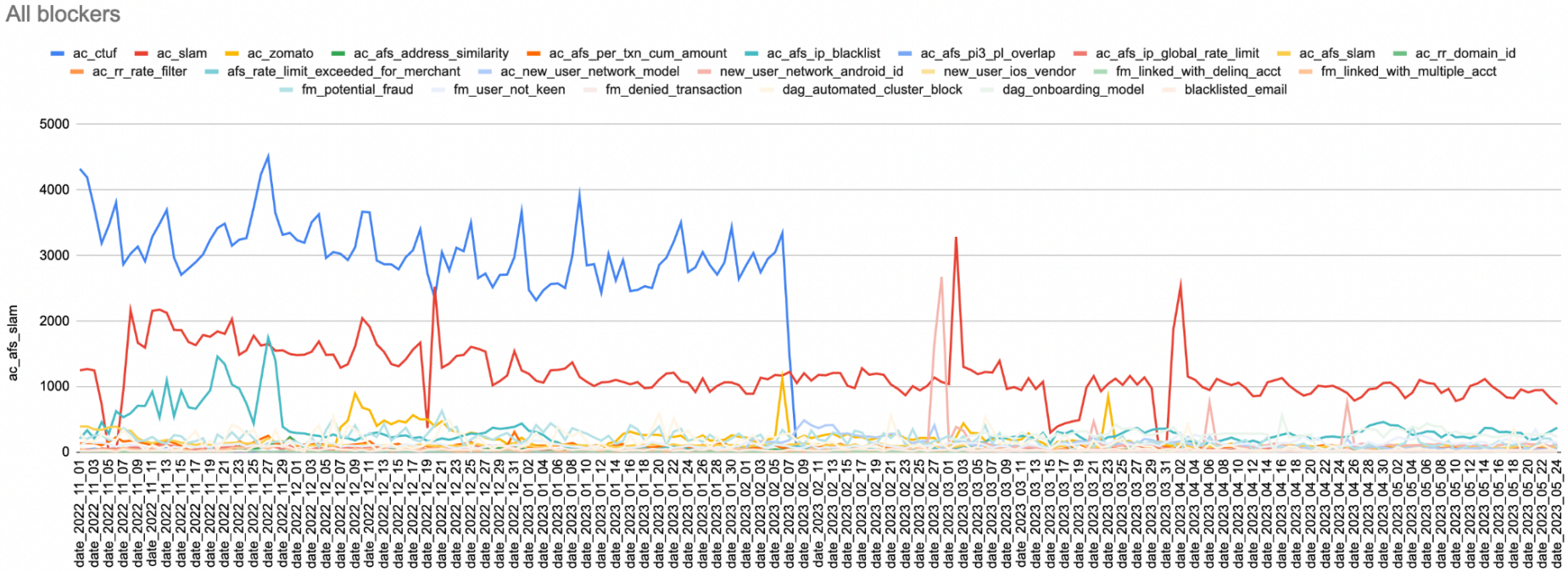
Let's zoom into constituents of this decline.

Cycle	Daily NU	Daily blocks	daily blocks / daily NU
Dec 2nd cycle	5814	5948	102.298
Jan 1st cycle	5900	5156	87.401
Jan 2nd cycle	6578	5175	78.663
Feb 1st cycle	6691	3909	58.419
Feb 2nd cycle	5864	2720	46.387

There is a steady drop in daily blocks / daily NU.

For Dec 2nd cycle -> Jan 1st cycle -> Jan 2nd cycle, daily blocks don't decline that much.

But Jan 2nd cycle -> Feb 1st cycle -> Feb 2nd cycle, they decline 5.1K -> 3.9K -> 2.7K. They have become stable after that (except for ATO)



It is a noisy diagram, but we need to look at it closely. [This is the link] where original resides.

What is noticeable is this:

1. CTUF shutdown happened on Feb 8. And Jan 2nd cycle -> Feb 1st cycle -> Feb 2nd cycle we observe the first bit of increase in NU delinquency:

Dec 1st cycle	20.93
Dec 2nd cycle	20.13
Jan 1st cycle	20.83
Jan 2nd cycle	20.01
Feb 1st cycle	21.61
Feb 2nd cycle	22.97
Mar 1st cycle	23.12
Mar 2nd cycle	22.82
Apr 1st cycle	23.08
Apr 2nd cycle	24.79

This is a full 20.01 => 22. 97, roughly 3pp increase. (Though there can be other factors, including odd vs even cycle).

This number matches a back of envelope calculation:
As per this document, CTUF used to lead to a loss of ~600 users per day.
Assuming a precision of 0.33, the number of fraudsters blocked everyday = 200.

So, number of fraudsters blocked in a cycle = 200 * 15 = 3000

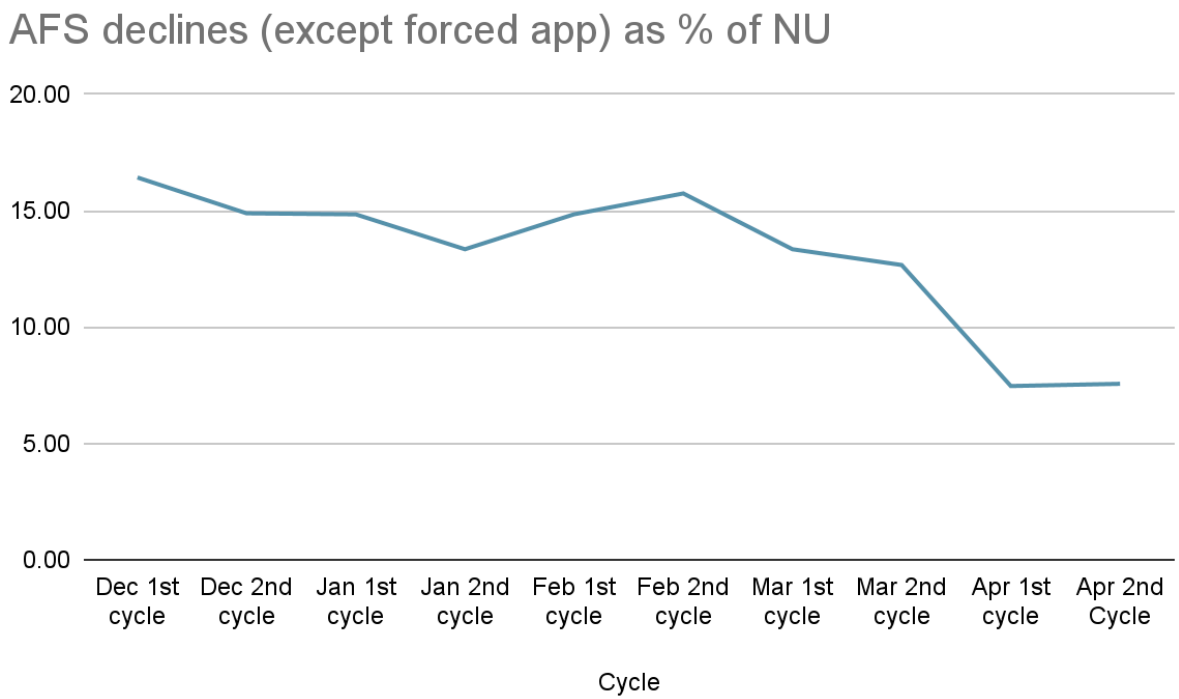
Assume 100K NU per cycle.
So, delinquency increase expected when CTUF goes away = 3000 / 100000 = 3%.

So, this matches with observed increase in delinquency.

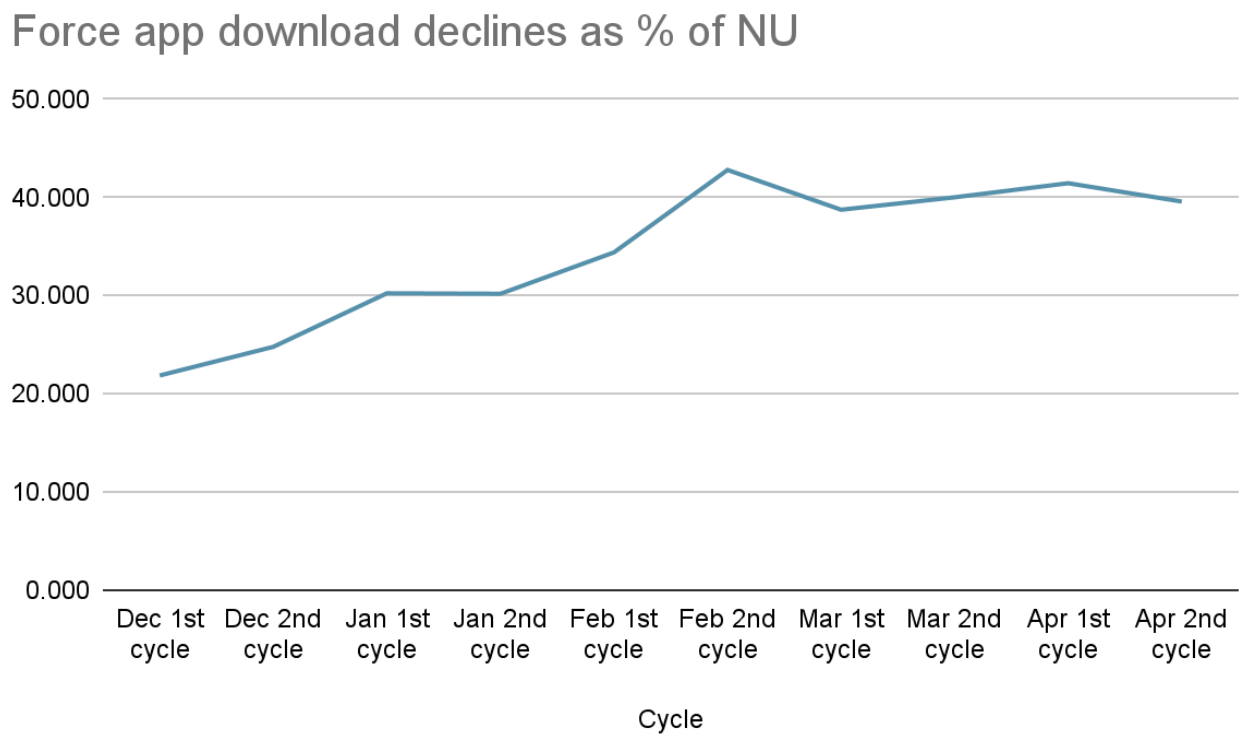
Note that when CTUF was active, many users being blocked by CTUF were later getting unblocked. Since this unblocking was happening at hourly level, it is possible that many of these transactions were denied in the interim and thus delinquency was controlled. Now, there is no block to begin with.

Users being denied transactions everyday

Following is the number of transaction denials / daily NU per day, not counting forced app download filter.

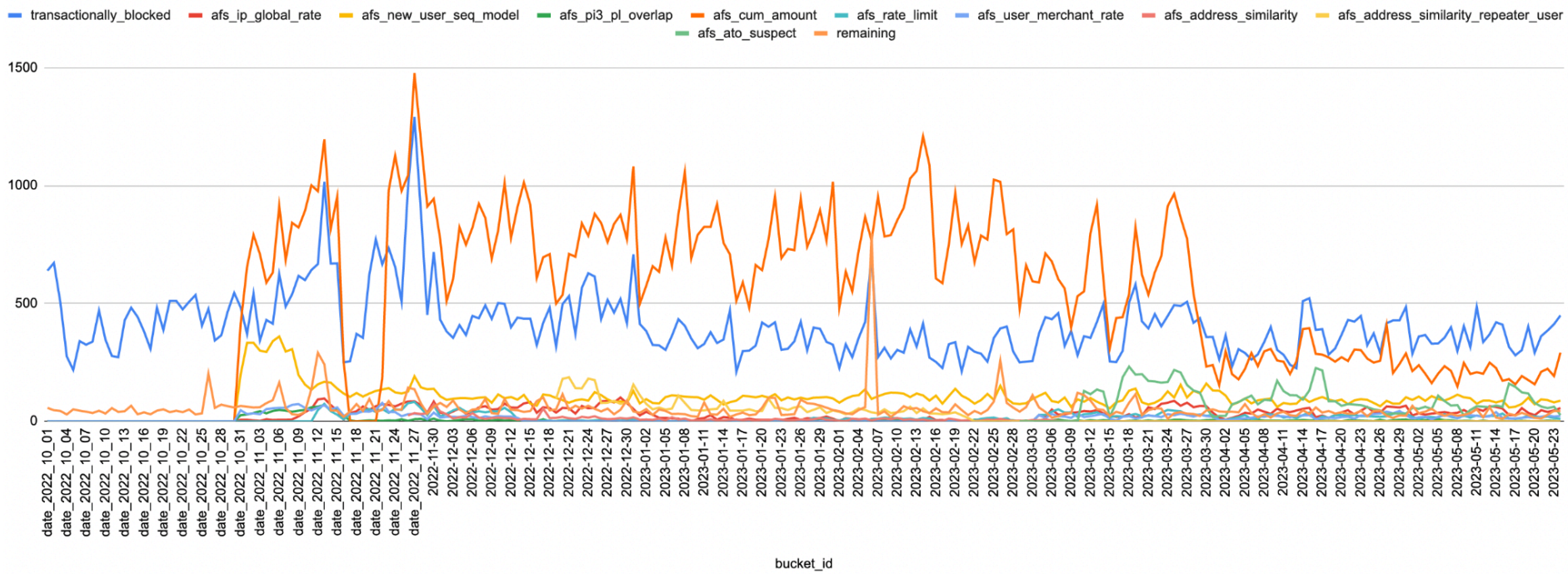


For completeness, force app download denials / NU has been rising consistently as we would expect:



Back to non force app AFS declines, May 2nd cycle -> Apr 1st cycle, there has been a steep decline in number of transaction denials.

Let's dig deeper into filter specific trends to isolate the filter which contributes to this decline. Here is the original excel sheet for the same, which is also copied below.



The orange line which dips to low values, from ~700 per day to ~300 per day, on March end. is cumulative amount filter. The filter was retrained because its precision had become low.

Do lower blocks by this precision explain the second episode of delinquency increase? There are two opinions:

1. No, it does not: Retrained model with fewer blocks was live in all of Apr 1st cycle. But delinquency dips happened only in Apr 2nd cycle. So, this model does not explain Apr 1st cycle -> Apr 2nd cycle delinquency.

2. Yes, it does: The math is hard to refute. As per adhoc precision numbers we had, before training precision of filter was 32%, and after training, precision became 41%. When the filter was blocking 700 filters at 32% users, it was blocking 700 * 0.32 = 224 fraudsters every day. Retrained model similarly blocks around 350 * 0.41 = 144 users. So, we are letting 224 - 144 = 80 fraudsters go

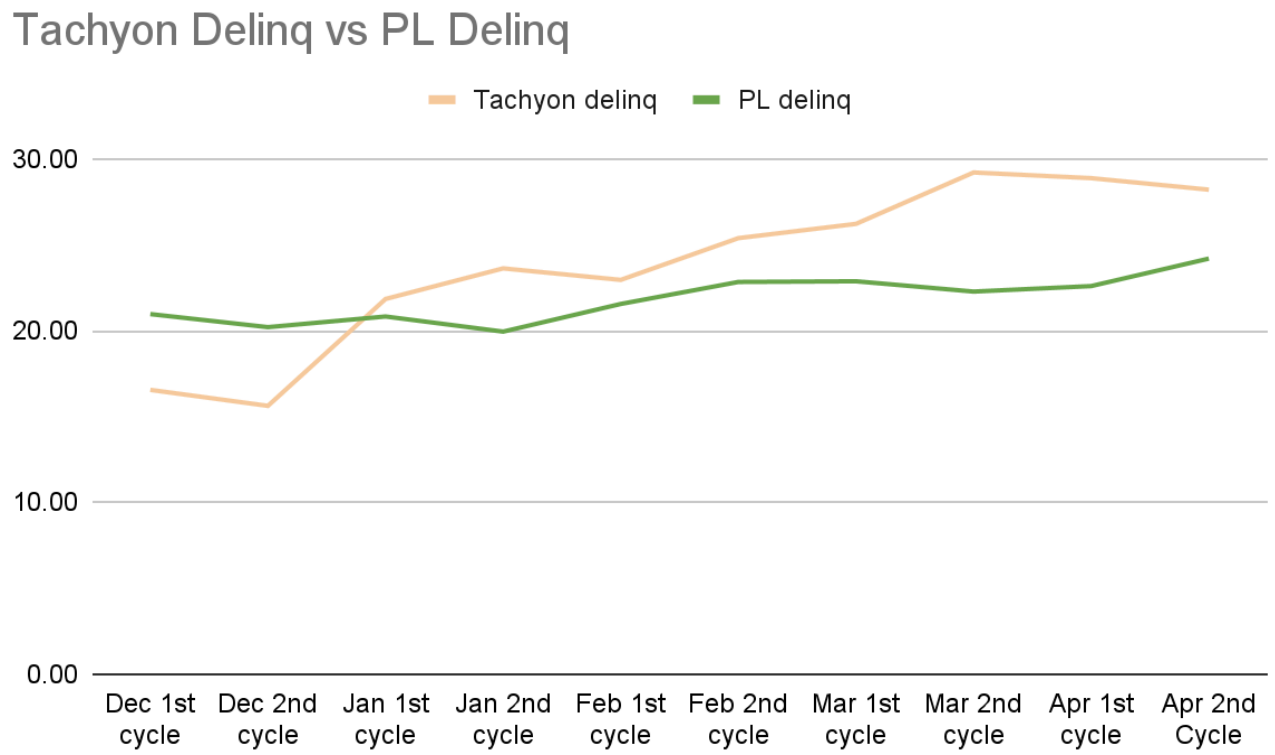
through every day. Over a cycle, this is like 80 * 15 = 1200 fraudsters per cycle. In a cycle we have like 100K users, which means a delinquency increase of 1200 / 100K = 1.2%. This 1.2% will be reflected in our delinquency numbers.

As to why delinquency increase does not show up in March 2nd cycle -> Apr 1st cycle, may be there were other factors that helped decrease the delinquency (though I can't think of what these factors could have been)

3. Tachyon Delinquencies

Tachyon delinquencies are higher than PL delinquencies and the fraction of Tachyon TPV as a fraction of all TPV is rising. How much could Tachyon delinquency help explain Jan 2nd cycle -> Apr 2nd cycle difference? Let's check.

First, let's see Tachyon vs non tachyon delinquencies



We are comparing Jan 2nd cycle vs Apr 2nd cycle, and during this time period, delinquency gap between PL delinq and Tachyon delinq has widened from 3.68% to 4.01%. More importantly, fraction of TPV of more delinquent Tachyon has increased from 1.87% to 9.76%.

Here are some relevant numbers:

	Jan 2nd cycle	Apr 2nd cycle
Tachyon delinq	23.62	28.20
PL delinq	19.94	24.19
Tachyon traffic %	1.87	9.76
PL traffic %	98.13	90.24

So,
Jan 2nd cycle overall delinq = (23.62 * 1.87 + 19.92 * 98.83) / 100 = 20.12
Apr 2nd cycle overall delinq = (28.20 * 9.76 + 24.19 * 90.24) / 100 = 24.58

(24.58 is lower than DPD27 number quoted above for Apr 2nd cycle, because DPD30 is available by the time I write this, but when I wrote earlier numbers, only DPD27 was available)

What would Apr 2nd cycle overall delinq be if Tachyon fraction stayed where it was:

(28.20 * 1.87 + 24.19 * 98.13 / 100 = 24.26

So, Tachyon increased traffice explains 24.58 - 24.26 = 0.32%, i.e. only a small part of the delinquency increase.

4. Increased control set delinquency

We computed the delinquency of control set, and while I am not giving all computations here, its impact on overall delinquency is like 0.3-0.4%.

Action Items (to be discussed with team)

Based on the above analysis, action items are clear.

1. Retrain network model
   If there is a dire need to reduce NU delinquency, CTUF can be brought back, but we will lose like 500 NU every day.
[Plan: Analyze network data for repeat users, build repeat user filter, then retrain network model for new users]

2. Make cumulative spend filter a bit more liberal
   i.e. make it block more users

3. Retrained SLAM is already live. That should anyway help. (0.5pp)
3(i) We can increase the rollout % of SLAM.

4. RTA team is making some changes to reduce RTA delinquencies. That will help overall delinquencies by 0.5%. Note that while RTA has always had higher delinquency than other lines, gap between RTA lines and other lines has reduced during the time overall delinquencies have increased.

5. Tachyon fraction is rising. At current levels of 10% it does not hurt appreciably. But it is growing fast and in a few months it will start hurting us.

Notes
=====
24.7 (Apr 2nd cycle) ==> May 1st /2nd cycle (reduction from force app) ==> June 1st cycle (24.7 - 0.5 from RTA improvement - 0.75 SLAM)