Darren Hong

Exploring the Most Active Bots and Their Community

During the r/place event, many people argue about whether or not using bots to paint pixels automatically is “fair” or not. Some argue that bots undermine the collaborative and competitive nature of the event, while others see them as a tool for coordination or protection. I wanted to find out which community used them the most and the reason why.

In order to distinguish between active users and bots with some order of certainty, I selected the top 1% active users being 103,811 unique users, and then predicted that out of that pool, 1,803 of them were bots due to their average time interval in between pixel placements being less than 7 minutes. I set the limit to 7 minutes because the normal cooldown for pixel placement is 5 minutes, and normal humans would have to take long intervals due to sleep, therefore probably increasing their average time interval greater than 7 and being out of the top 1% of active users.

Out of the suspected bots, the most placed pixels from them were at two artworks which are shown at the table to the right. 12 of the top 20 most placed pixel locations from the suspected bots were at these artworks. Both of these artworks were from the *My Little Pony* (MLP) fandom. The two artworks were located at (892, 1830) to (961, 1886) and (1611, 212) to (1691, 277), the left and right picture below respectively.

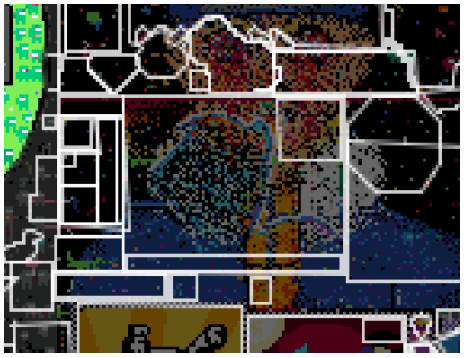
Most Placed Pixels From Suspected Bots
Only Showing Pixels Part Of MLP Artworks

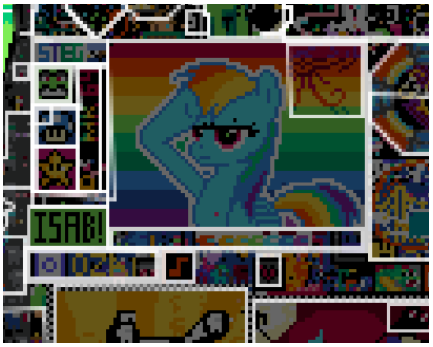
Rank	Coordinates	# of Placements
1	918,1862	1296
3	931,1852	818
4	924,1856	814
5	919,1853	615
9	1651,252	364
10	1652,253	353
13	1651,230	328
14	1651,253	326
16	1611,254	307
17	893,1831	301
19	917,1858	271
20	1652,252	270

The MLP reddit community openly stated and confirmed that they used bots and their reason on why they used bots was because they kept getting attacked by multiple online influencers. They were attacked several times by big streamers such as Mizkif, xQc, and Asmongold, each with millions of followers overwriting their pixel art. The below pictures are the artworks being attacked by the streamers and their viewers.

The most active time where the bots were used was during attacks. The table to the right shows the top 4 most hourly changes made from the suspected bots. Using the r/place Atlas to verify my analysis, it showed that an attack happened when the suspected MLP bots were most active.Below are pictures from the time frame in the table, showing the start of the attack to the quick recovery of the artwork. This example showed how useful the bots were in rapidly recreating the artwork after being vandalized.

Time	# of Changes
2022-04-04 04:00:00	3373
2022-04-04 06:00:00	3067
2022-04-04 07:00:00	3032
2022-04-04 05:00:00	3010





DuckDB vs. PySpark

My framework of choice was DuckDB. When trying to change from DuckDB to PySpark, I had trouble with out of memory errors, which I solved by allocating more memory when creating the spark session. I also had to configure the session due to me not removing the "UTC" in the timestamp. For my code, DuckDB's execution time was 68900.6665ms and PySpark's execution time was 172869.4386ms, making my PySpark implementation slower by a factor of 2 compared to my DuckDB implementation. It was easier for me to use DuckDB since I'm familiar with it compared to PySpark which is why my PySpark configuration is unoptimized. From my time learning about PySpark to use for this analysis, I know that it allows for more functionality like Polars and is made for distributed processing, but since I'm processing small amounts of data, DuckDB would be better since I'm more used to it.