Improve the retention rate based on users' behavior analysis

Shenghua Du (MSc in Business Analytics and Data Scientice)

Essec Business School - Mozat Singapore

shenghua.du@essec.edu

June 25, 2018

I Abstract

As we know, Internet companies often launch some campaigns to acquire new users or reserve old users. So how to improve these users' s retention rate has become an important issue.In this project we will design a new user guide by analyzing the users behavior log data to help improve it.Below are one music app new users' natural retention rate curve and improved curve after operation campaign.

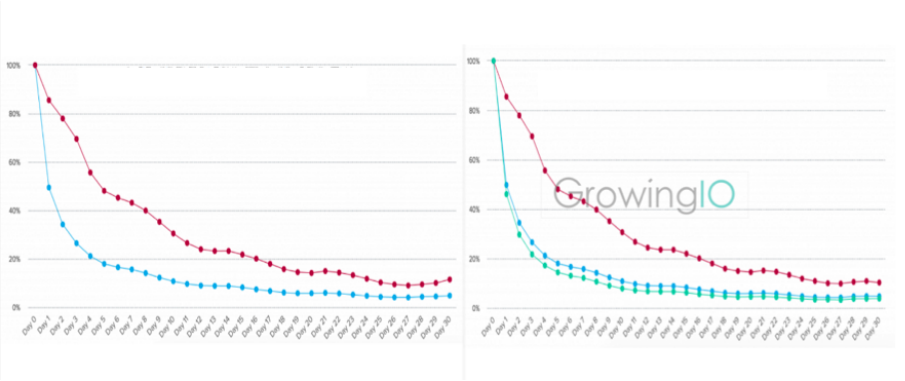


Figure 1: *An example of improvement of retetion rate*

II Data source and structure

Our data is from Mozat Singapore's mobile broadcast app Loops's database. We have used the users' basic information and behavior log,which includes 3000W records.

1-500 of 501+ Comma-...d (CSV) View Query

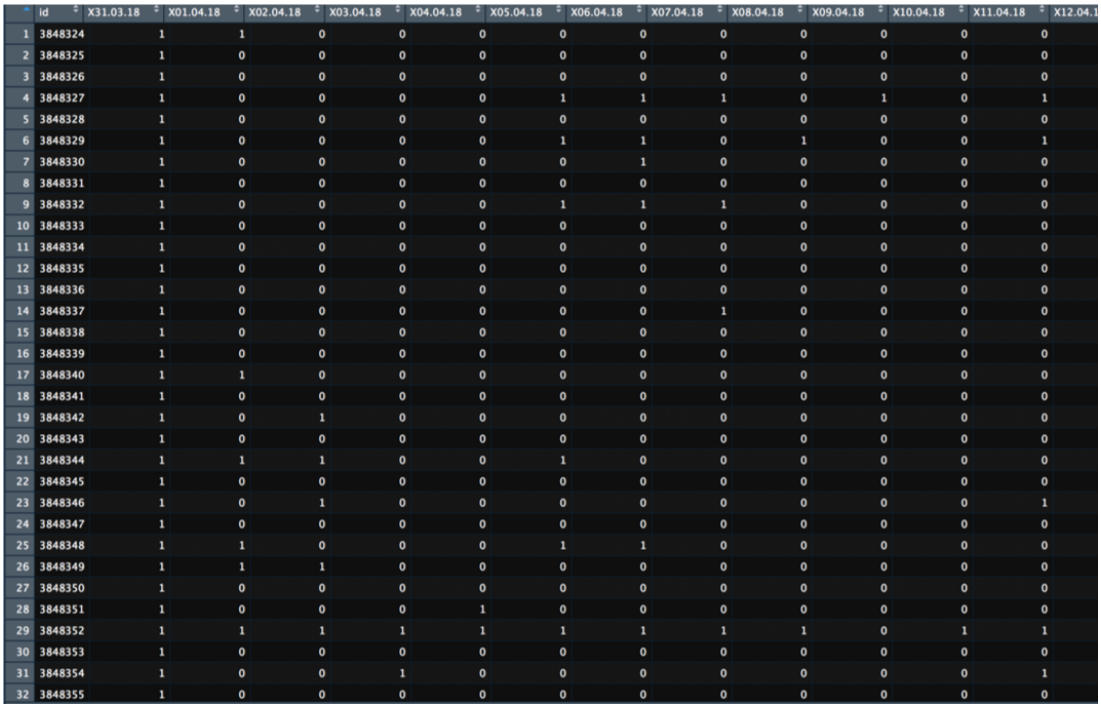
	id	register_time	last_online	ts	is_bdc_enter	is_chat_send	is_explore_click	is_gift_send	is_leaderboard_click	is_mylevel_enter	is_mytitle_enter
1	3848423	2018-03-31 00:28:39.000000	2018-03-31	2018-03-31 00:28:39.305000	0	0	1	0	0	0	0
2	3848463	2018-03-31 00:42:04.000000	2018-03-31	2018-03-31 00:42:04.393000	0	0	0	0	0	0	0
3	3848421	2018-03-31 00:28:15.000000	2018-03-31	2018-03-31 00:40:22.062000	0	0	1	0	0	0	0
4	3848506	2018-03-31 00:49:45.000000	2018-03-31	2018-03-31 00:54:02.972000	0	0	0	0	0	0	0
5	3848324	2018-03-31 00:00:42.000000	2018-03-31	2018-03-31 00:53:20.537000	0	0	0	0	0	0	0
6	3848414	2018-03-31 00:26:24.000000	2018-03-31	2018-03-31 01:08:18.295000	0	14	0	0	0	0	0
7	3848581	2018-03-31 01:16:30.000000	2018-03-31	2018-03-31 01:16:30.280000	0	0	0	0	0	0	0
8	3848602	2018-03-31 01:21:29.000000	2018-03-31	2018-03-31 01:24:13.666000	8	0	0	0	0	0	0
9	3848607	2018-03-31 01:23:12.000000	2018-03-31	2018-03-31 01:28:49.750000	10	2	0	0	0	0	0
10	3848593	2018-03-31 01:19:02.000000	2018-03-31	2018-03-31 01:33:29.567000	0	7	0	0	0	0	0
11	3848657	2018-03-31 01:41:26.000000	2018-03-31	2018-03-31 01:41:26.834000	6	0	0	0	0	0	0
12	3848718	2018-03-31 02:09:34.000000	2018-03-31	2018-03-31 02:09:34.986000	4	0	0	0	0	0	0
13	3848714	2018-03-31 02:07:16.000000	2018-03-31	2018-03-31 02:24:38.604000	0	0	0	0	0	0	0
14	3848763	2018-03-31 02:27:25.000000	2018-03-31	2018-03-31 02:27:25.433000	2	0	0	0	0	0	0
15	3848622	2018-03-31 01:28:18.000000	2018-03-31	2018-03-31 02:25:08.099000	0	0	0	0	0	0	0
16	3848766	2018-03-31 02:29:08.000000	2018-03-31	2018-03-31 02:35:15.095000	1	7	0	0	0	0	0
17	3848434	2018-03-31 00:32:49.000000	2018-03-31	2018-03-31 02:34:31.790000	0	0	0	0	0	0	0
18	3848678	2018-03-31 01:51:01.000000	2018-03-31	2018-03-31 02:49:39.704000	5	6	0	0	0	0	0
19	3848778	2018-03-31 02:34:39.000000	2018-03-31	2018-03-31 02:54:39.154000	0	6	0	0	0	0	0
20	3848874	2018-03-31 03:22:38.000000	2018-03-31	2018-03-31 03:22:38.044000	0	0	0	0	0	0	0

Figure 2: *The data structure*

III Definition and Benchmark

Nine users' behavior includes:watch live, send chat,send gift,go to upcoming,register upcoming, go to explore center,enter my level. These users need to do the behavior in his first online time cause we need to analyze the influence to the retention rate of the first time's click.

Retention rate:We need to account all new users' retention ratio and users' who had done the behavior active ratio in one month separately. If a user is online one day we account 1 active point in one day Then we use the accumulated point divided by the number of users We can get the one month's retention ratio.



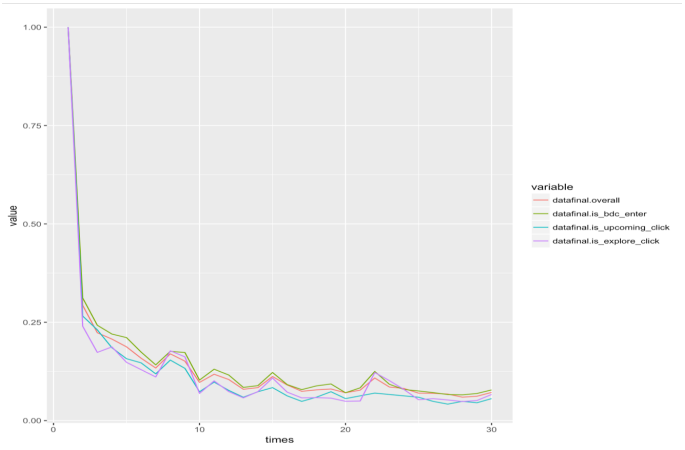
	id	X31.03.18	X01.04.18	X02.04.18	X03.04.18	X04.04.18	X05.04.18	X06.04.18	X07.04.18	X08.04.18	X09.04.18	X10.04.18	X11.04.18	X12.04.1
1	3848324	1	1	0	0	0	0	0	0	0	0	0	0	
2	3848325	1	0	0	0	0	0	0	0	0	0	0	0	
3	3848326	1	0	0	0	0	0	0	0	0	0	0	0	
4	3848327	1	0	0	0	0	1	1	1	0	1	0	1	
5	3848328	1	0	0	0	0	0	0	0	0	0	0	0	
6	3848329	1	0	0	0	0	1	1	0	1	0	0	1	
7	3848330	1	0	0	0	0	0	1	0	0	0	0	0	
8	3848331	1	0	0	0	0	0	0	0	0	0	0	0	
9	3848332	1	0	0	0	0	1	1	1	0	0	0	0	
10	3848333	1	0	0	0	0	0	0	0	0	0	0	0	
11	3848334	1	0	0	0	0	0	0	0	0	0	0	0	
12	3848335	1	0	0	0	0	0	0	0	0	0	0	0	
13	3848336	1	0	0	0	0	0	0	0	0	0	0	0	
14	3848337	1	0	0	0	0	0	0	1	0	0	0	0	
15	3848338	1	0	0	0	0	0	0	0	0	0	0	0	
16	3848339	1	0	0	0	0	0	0	0	0	0	0	0	
17	3848340	1	1	0	0	0	0	0	0	0	0	0	0	
18	3848341	1	0	0	0	0	0	0	0	0	0	0	0	
19	3848342	1	0	1	0	0	0	0	0	0	0	0	0	
20	3848343	1	0	0	0	0	0	0	0	0	0	0	0	
21	3848344	1	1	1	0	0	1	0	0	0	0	0	0	
22	3848345	1	0	0	0	0	0	0	0	0	0	0	0	
23	3848346	1	0	1	0	0	0	0	0	0	0	0	1	
24	3848347	1	0	0	0	0	0	0	0	0	0	0	0	
25	3848348	1	1	0	0	0	1	1	0	0	0	0	0	
26	3848349	1	1	1	0	0	0	0	0	0	0	0	0	
27	3848350	1	0	0	0	0	0	0	0	0	0	0	0	
28	3848351	1	0	0	0	1	0	0	0	0	0	0	0	
29	3848352	1	1	1	1	1	1	1	1	1	0	1	1	
30	3848353	1	0	0	0	0	0	0	0	0	0	0	0	
31	3848354	1	0	0	1	0	0	0	0	0	0	0	1	
32	3848355	1	0	0	0	0	0	0	0	0	0	0	0	

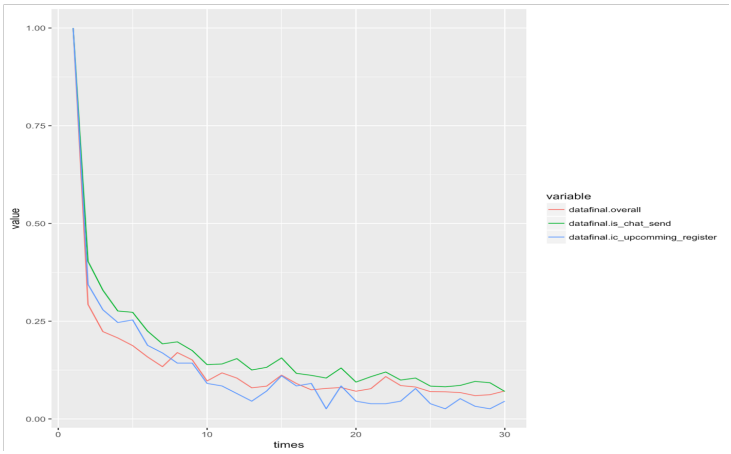
Figure 3: *How to difine he retention rate*

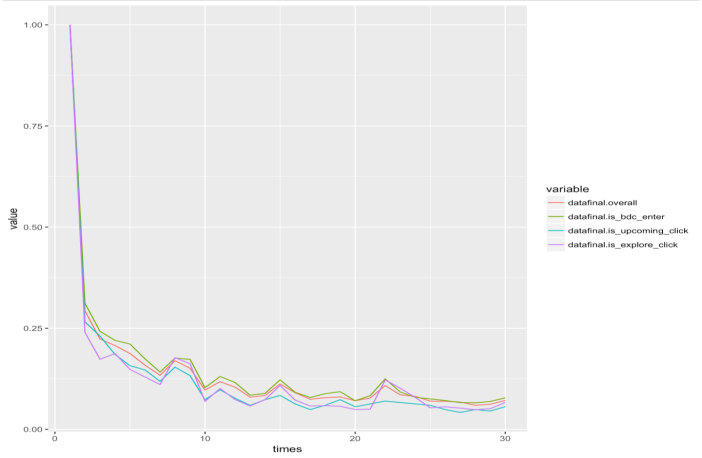
And another question:how to divide people who had been familiar with broadcast platform and who originally have high retention rate and who are motivated by the behavior?In order to reduce the influence of these people we remove these people who had given gift in one month. We think them as people who originally have high retention rate.

IV Result and conclusion

Below three figures are the final result what we get. Behavior of same dimension has been put into same figure.







V Strategy

1. Encourage broadcasters to welcome the new users in policy. Then watcher is more likely to interactive with broadcasters like send chat. (broadcaster guide)
2. First ten message rewards like experience to help user achieve level 2 more quickly.
3. New user guide to force new users to send chat.

VI Attachment

R code website(ssh): git@github.com:PPPPParadise/MSc-Capstone-Project.git