

A Look at Mobile Video Streaming Data: Korea

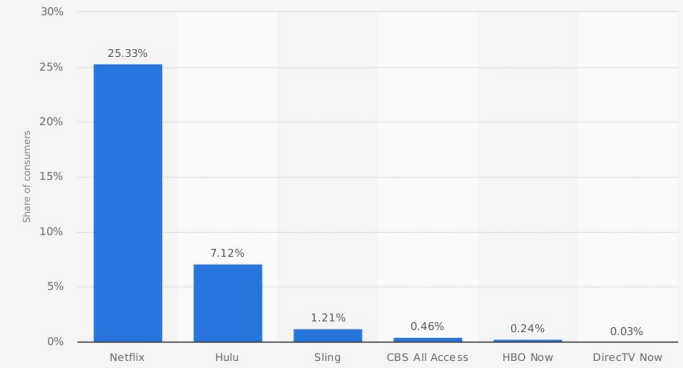
Youngseok Lee

cnulee@ucdavis.edu

Mobile Video Streaming Service

- US
 - Netflix, Amazon Video, Hulu, Sling, Google Play, Apple TV
- China
 - iqiyi, youku, le, fun, mgmtv, ppt
- Korea
 - Pooq, tving, oksusu

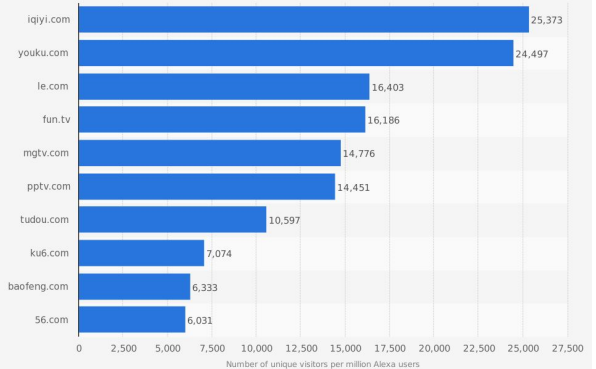
Share of consumers who subscribe to selected video streaming services in the United States as of September 2017



Sources
Recorder: Second Measure
© Statista 2017

Additional information:
United States; Second Measure September 2017

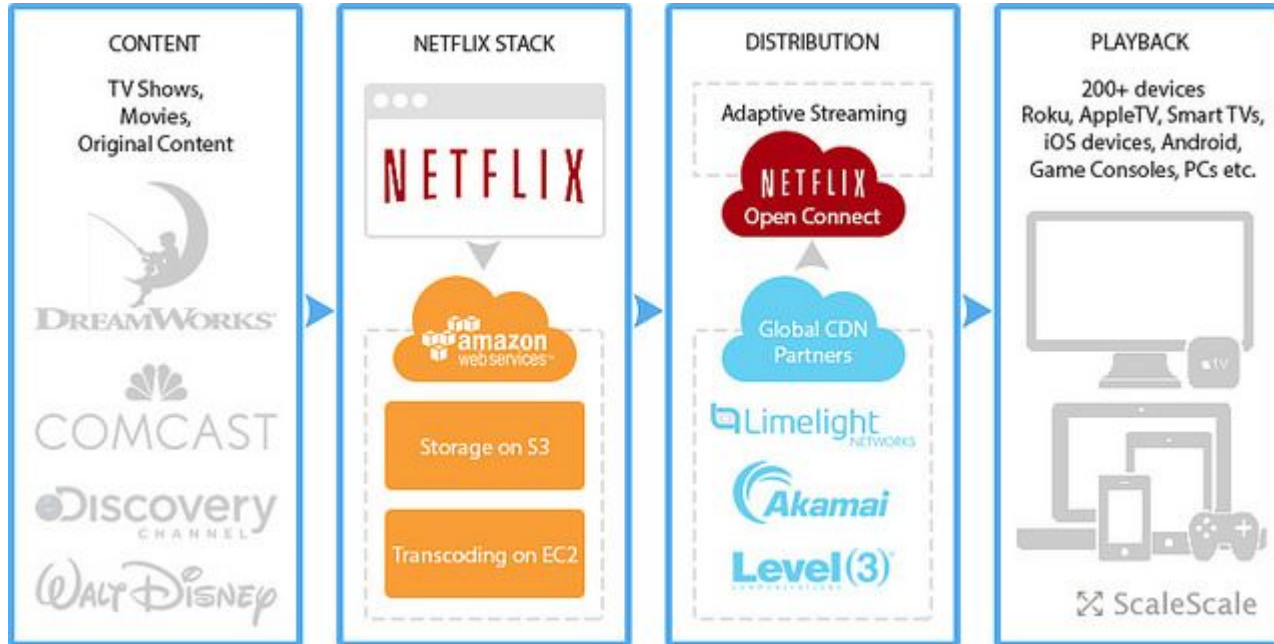
Leading online video platforms in China in April 2017 (in monthly unique visitors per million Alexa users)



Source
Alexa
© Statista 2017

Additional information:
China; Alexa; April 2017

Netflix Architecture



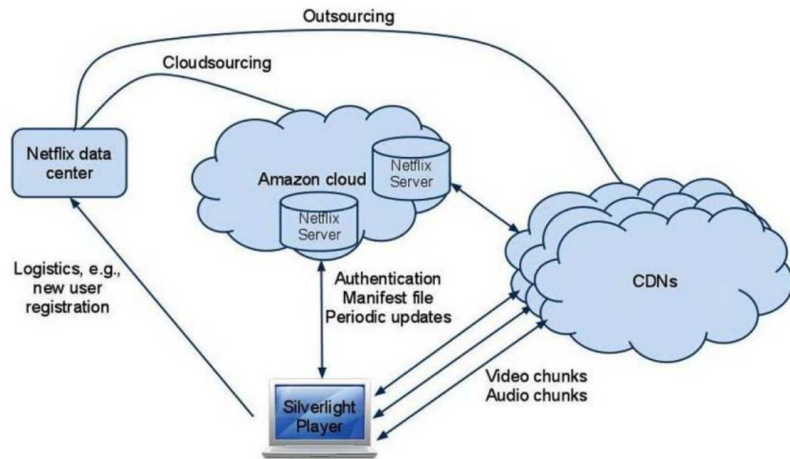


Fig. 1. Netflix architecture.

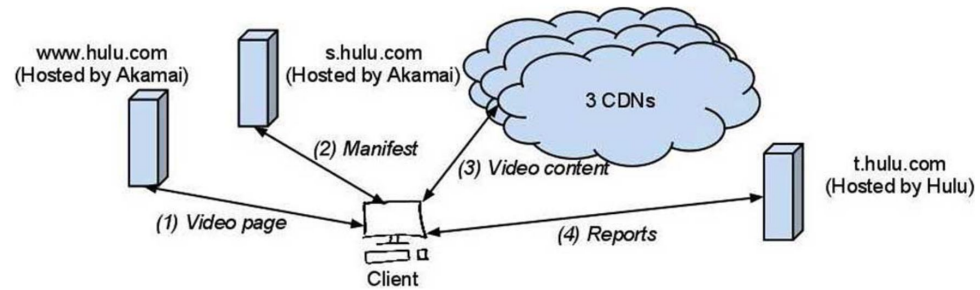


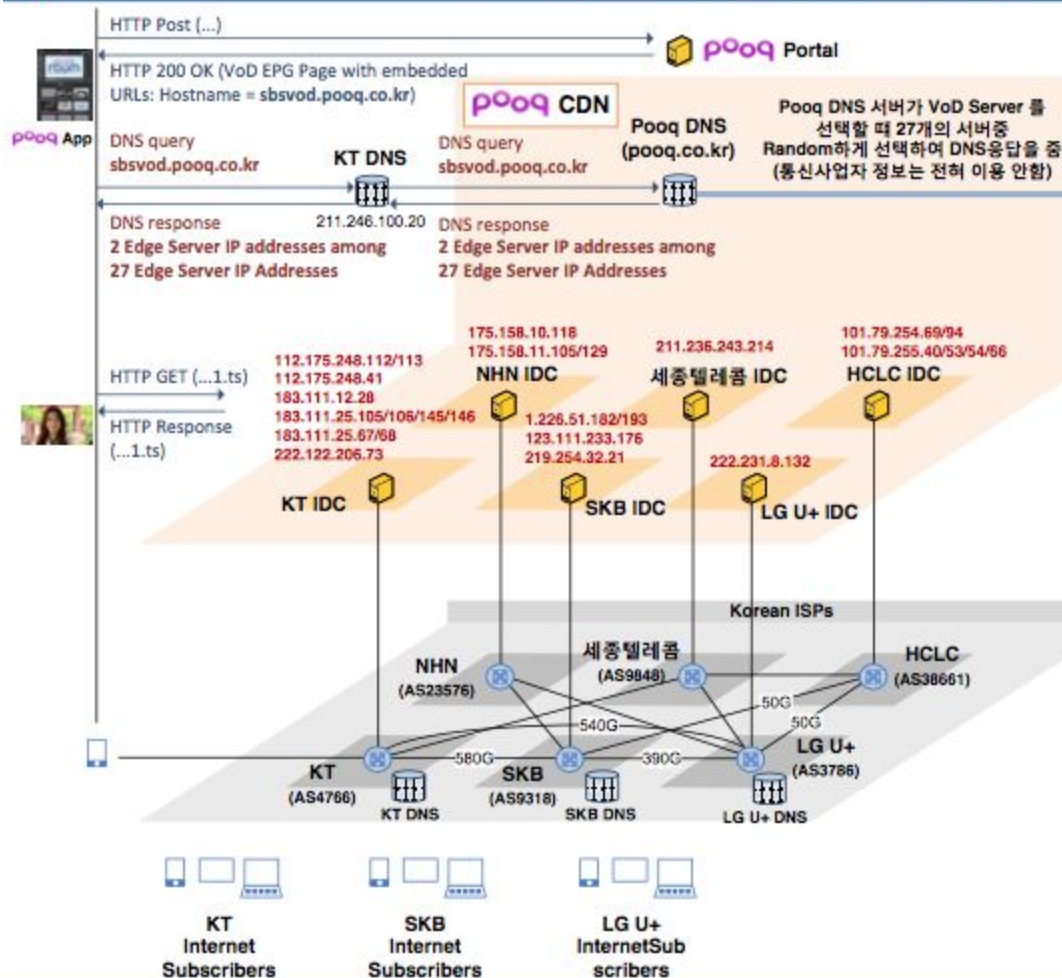
Fig. 6. High-level Hulu architecture.

Preliminary Study

- User behavior of video streaming applications in Korea
 - Real data from one content provider, Korea
- Applications of video streaming-related problems
 - recommendations of contents
 - network engineering issues

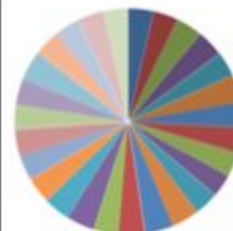
Pooq Data and Analysis Environment

- Date: 2017-01-01 23:00 ~ 24:00
- Volume: 4.4GB(csv)
 - Contents : live.csv(2.9K), movie.csv(504K), vod.csv(12M)
 - *live channels: 70, movies: 7K, vod: 280K*
 - Uers : member.csv(163M)
 - *subscribers: 3M*
- Analysis environment
 - Lenovo X1 Carbon(i7-6600U, 16GB RAM)
 - Spark 2.2.0, Zeppelin 0.7.3
 - Docker container



CP	SBS		
VoD Server Hostname	sbsvod.pooq.co.kr		
	1	112.175.248.112	1642
	2	112.175.248.113	1583
	3	112.175.248.41	1751
	4	183.111.12.28	1771
	5	183.111.25.105	1604
	6	183.111.25.106	1793
	7	183.111.25.145	1646
	8	183.111.25.146	1671
	9	183.111.25.67	1620
	10	183.111.25.68	1472
	11	222.122.206.73	1570
	12	1.226.51.182	1774
	13	1.226.51.193	1840
	14	123.111.233.176	1723
	15	219.254.32.21	1626
	16	222.231.8.132	1842
	17	211.236.243.214	1712
	18	101.79.254.69	1686
	19	101.79.254.94	1710
	20	101.79.255.40	1723
	21	101.79.255.53	1599
	22	101.79.255.54	1732
	23	101.79.255.55	1891
	24	101.79.255.66	1678
	25	175.158.10.118	1688
	26	175.158.11.105	1715
	27	175.158.11.129	1606

VoD Server IP	# of Selected by Pooq DNS	Server Location
112.175.248.112/113		KT
112.175.248.41		
183.111.12.28		
183.111.25.105/106/145/146		NHN
183.111.25.67/68		
222.122.206.73		SKB
1.226.51.182/193		
123.111.233.176		LG U+
219.254.32.21		
211.236.243.214		HCLC
101.79.254.69/94		
101.79.255.40/53/54/66		NHN
101.79.255.53		
101.79.255.54		HCLC
101.79.255.55		
101.79.255.66		NHN
175.158.10.118		
175.158.11.105		NHN
175.158.11.129		



- * 112.175.248.112
- * 112.175.248.113
- * 112.175.248.41
- * 183.111.12.28
- * 183.111.25.105
- * 183.111.25.106
- * 183.111.25.145
- * 183.111.25.146
- * 183.111.25.67
- * 183.111.25.68
- * 222.122.206.73
- * 1.226.51.182
- * 1.226.51.193
- * 123.111.233.176
- * 219.254.32.21
- * 211.236.243.214
- * 101.79.254.69
- * 101.79.254.94
- * 101.79.255.40
- * 101.79.255.53
- * 101.79.255.54
- * 101.79.255.55
- * 101.79.255.66
- * 101.79.255.66
- * 175.158.10.118
- * 175.158.11.105
- * 175.158.11.129

Questions

- Who
 - Who are watching video?
- When
 - When do users watch video?
- *Where*
 - *Are they watch video at home or outside?*
- What
 - What contents are popular to users?
- Why
 - Why some users watch this/that video or quickly stop watching video?
- How
 - How (which device) do users watch video?

Sample Data

date, hour, minute, second, uno,
channeltype, programId, contentId, cornerId, mediatime, devicetype, bitrate, ipaddress

20170313,09,02,08,9e51254c591e4f59d291a1f32c72d627a346f6f0ca08bc73bbc7862534637e86,
L,S01,,,01:13:20,Android Phone,1000,125.190.*.*

20170313,09,18,53,e829baa186f48bdd5d98162c999578d405ff92fdaa43e2f7730d207a0835246f,
L,M01,,,00:02:20,Android Phone,1000,175.223.*.*

20170313,09,38,58,1823af847183b2c182ef90c6425bf095010125c60a9a1896342104574ca741f8,
V,K02_T2017-0043,K02_PS-2017033426-01-000,1,01:06:00,Android Phone,1000,122.42.*.*

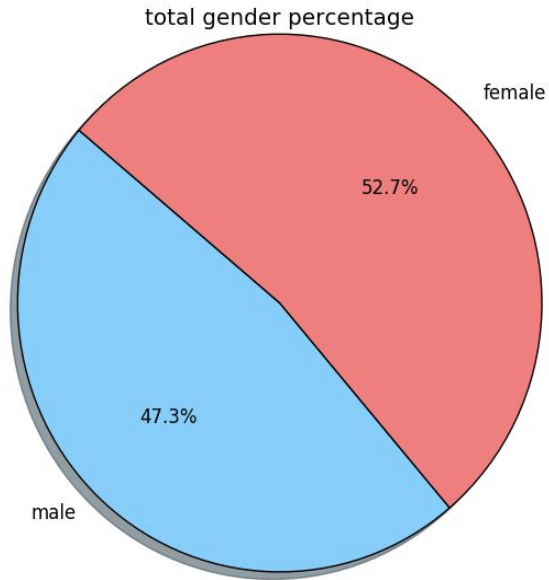
uno,gender,ageGroup,isCheckedName

8d4dc4006fa2990b93243277fbc0ff0799feaa20688bad17be6829f16b0c5561,M,35,Y

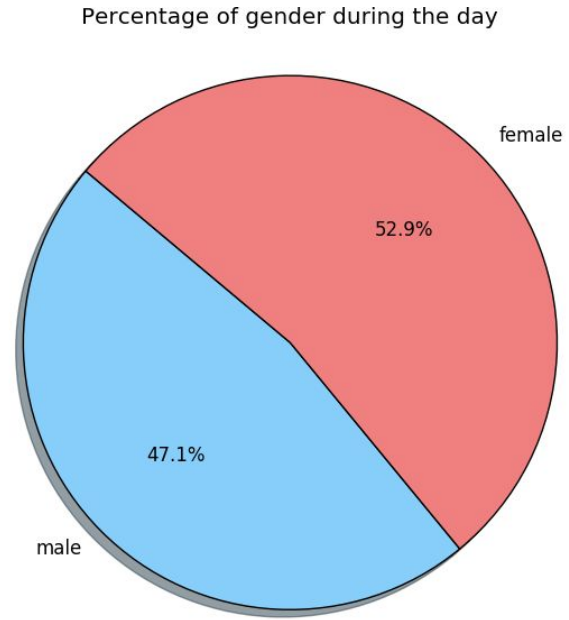
9e51254c591e4f59d291a1f32c72d627a346f6f0ca08bc73bbc7862534637e86,F,20,Y

35a5cb2f30bfcedb01a2c274952fd01a50f464fca7aeccaa9496f4ddeed6f88d,F,25,Y

Who) Gender

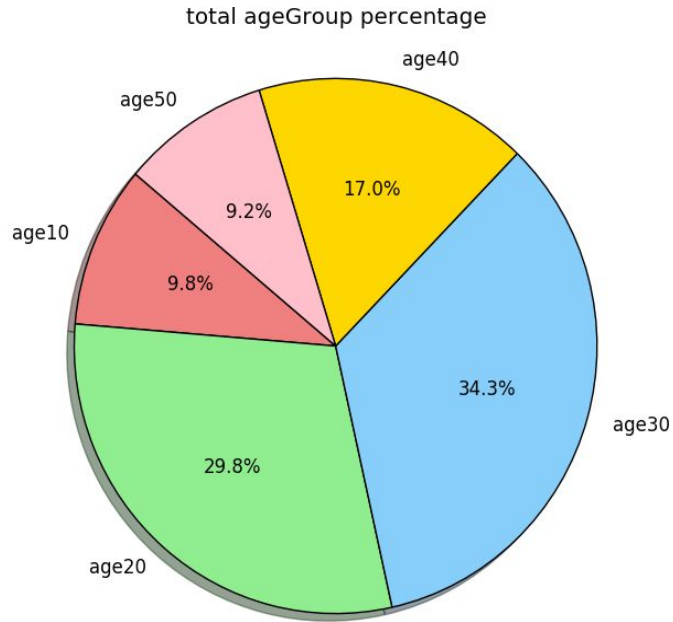


Subscribers

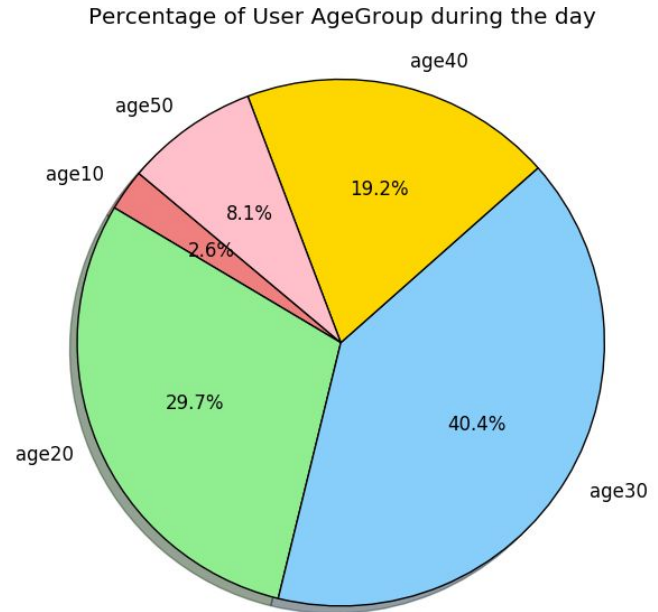


Viewers

Who) Age

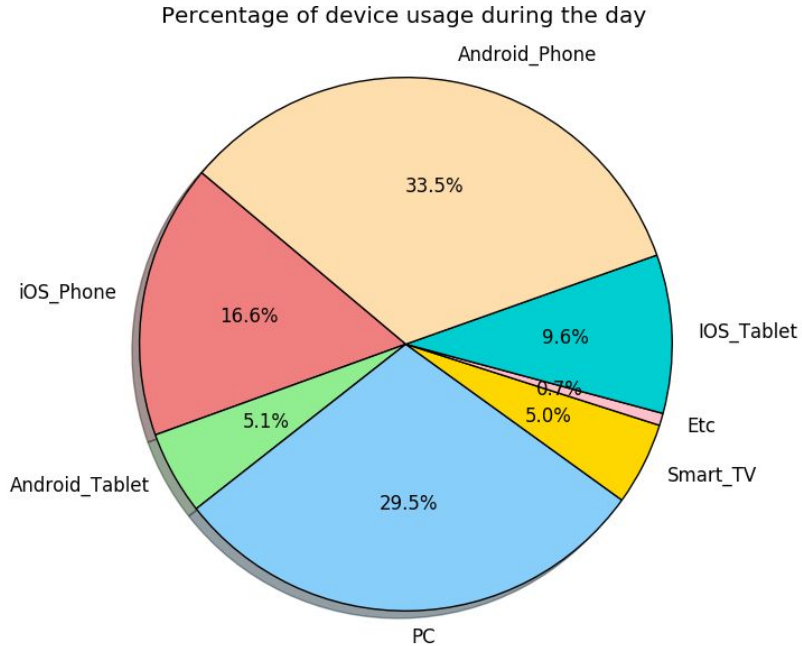


Subscriber

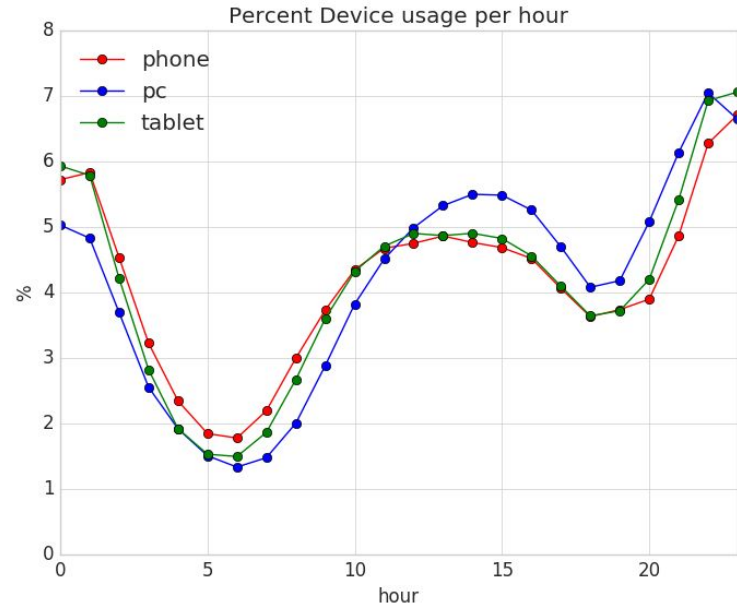


Viewers

How, When) Which Device?

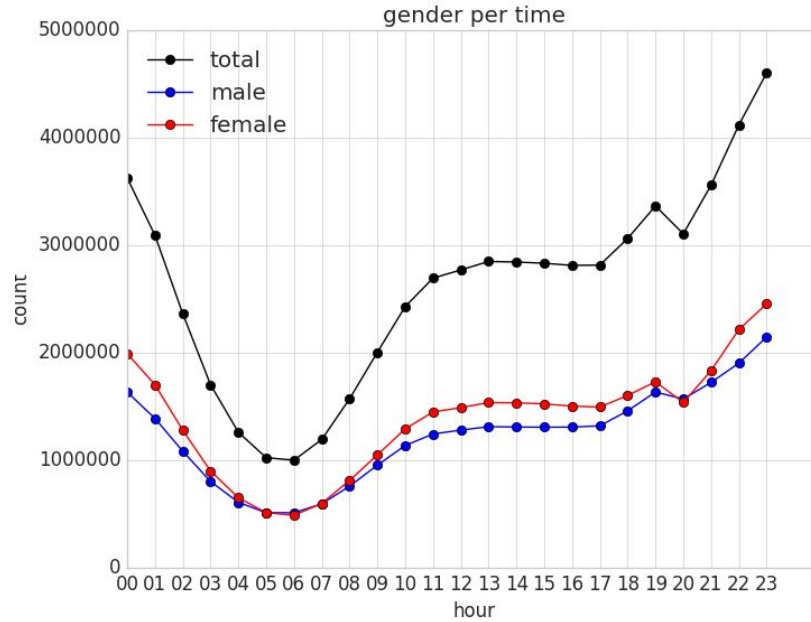


Device



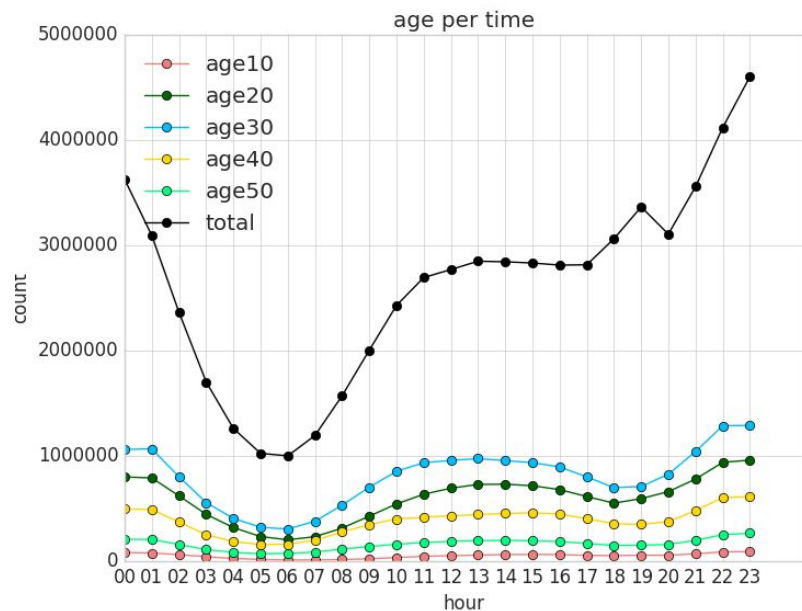
Hourly

Who, When) Hourly

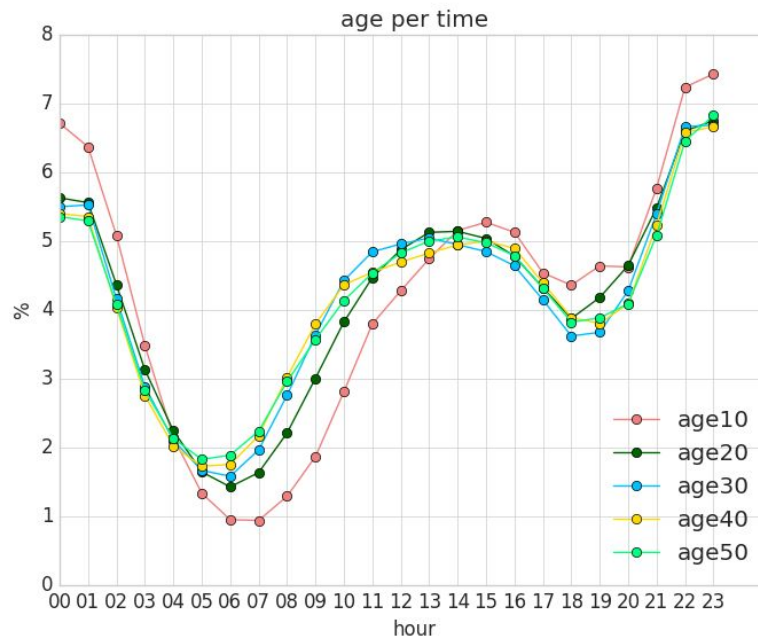


User Count

Who, When) Hourly Plot by Age



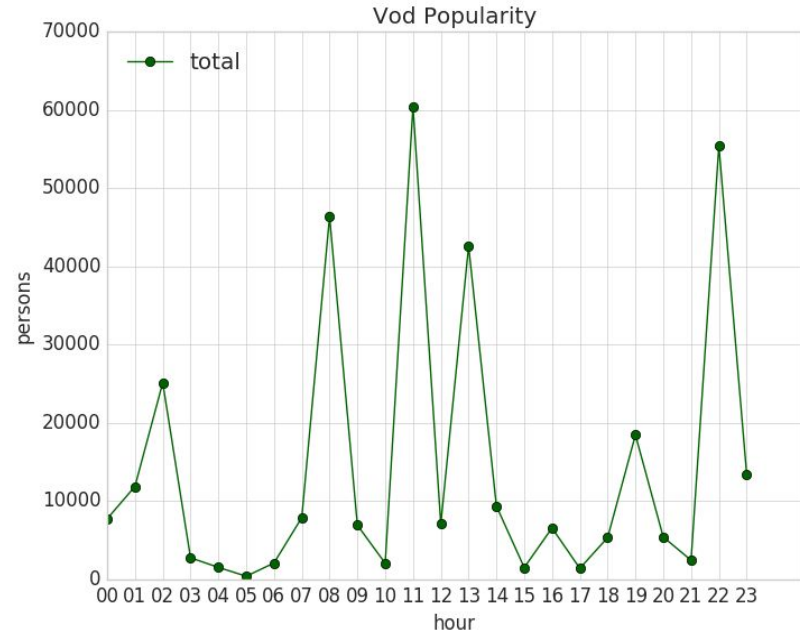
User views per hour



Percentage

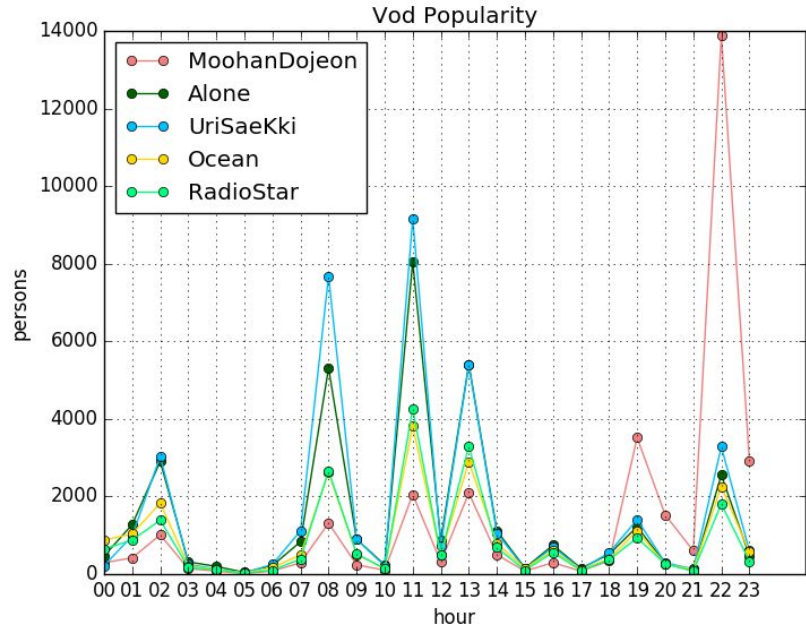
Who, When) Viewer Count per Hour

- Peak at 8, 11, 13, 22 o'clock



Who, What, When) Viewer Count for Top 5 Videos

- Top 5 video contents hit the peak time
 - Entertainment contents



Who) User Classification

- User behavior
 - Early Leaving viewer: stop watching within 5 minutes
 - Steady Viewer: watching a single video over 5 minutes
 - Highlighter: browsing interesting parts of a video
 - Surfing Watcher: watching several videos within an hour

Who) Four User Classification

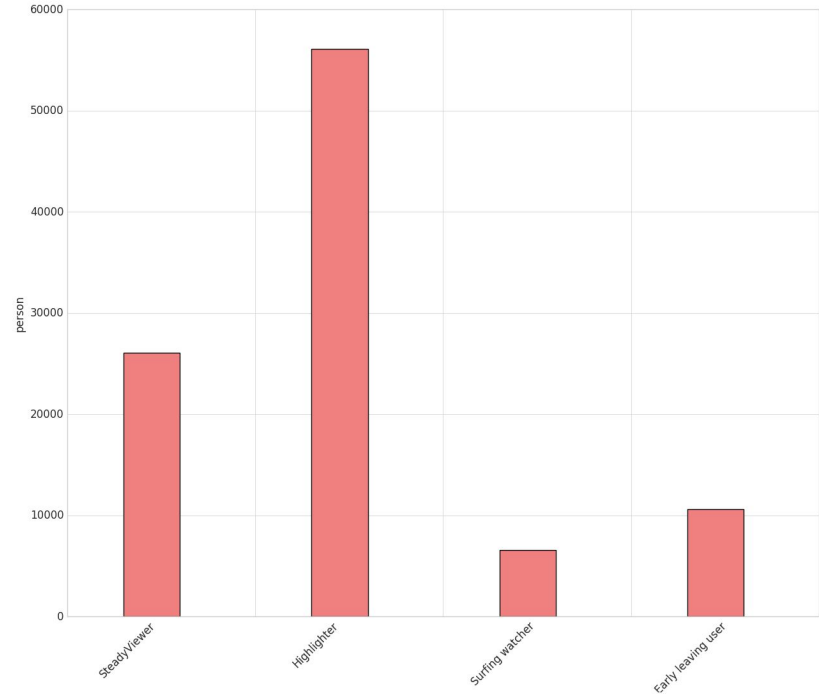
1. Highlighter!!!

- a. Users tend to watch only interesting parts

2. Steady viewers

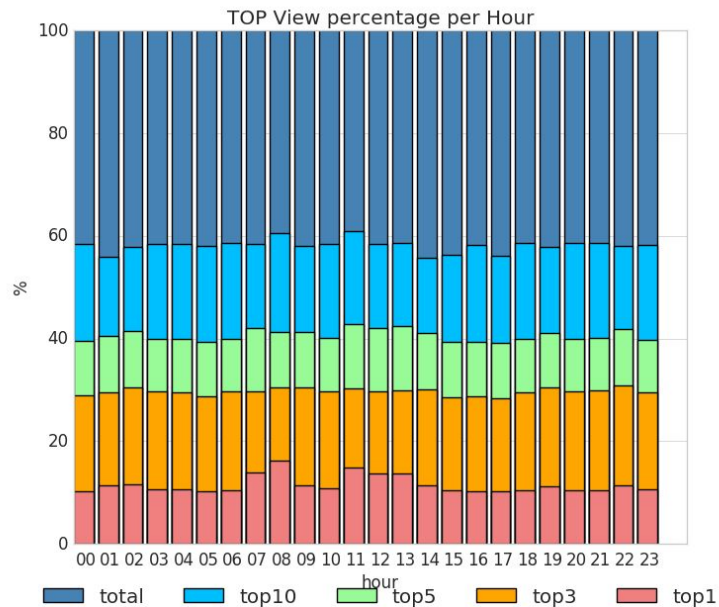
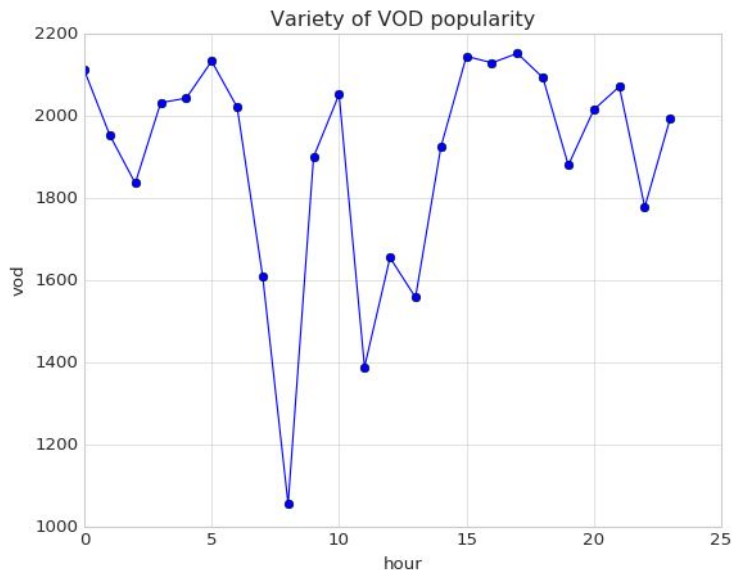
3. Early leaving viewers

4. Surfing viewers



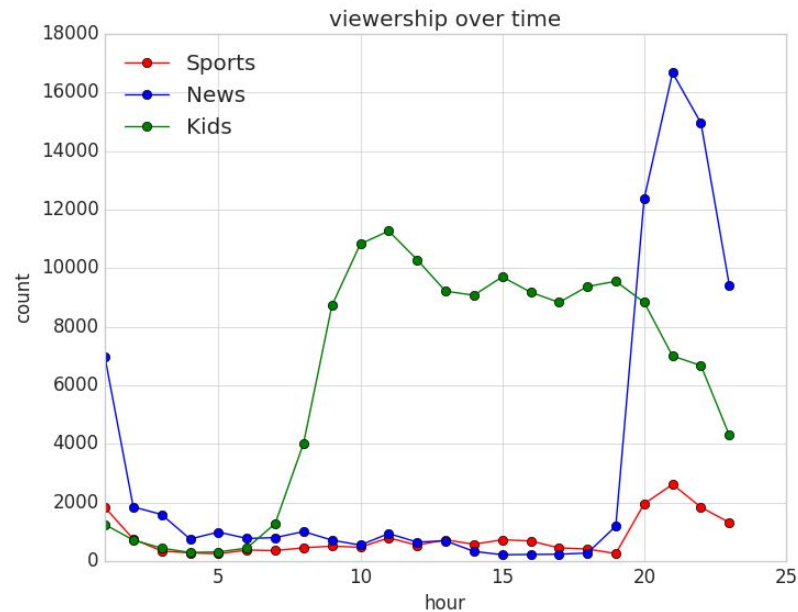
What, When) What are Popular Videos?

- Video count at 8, 11, 1 o'clock
 - Popular video (live)



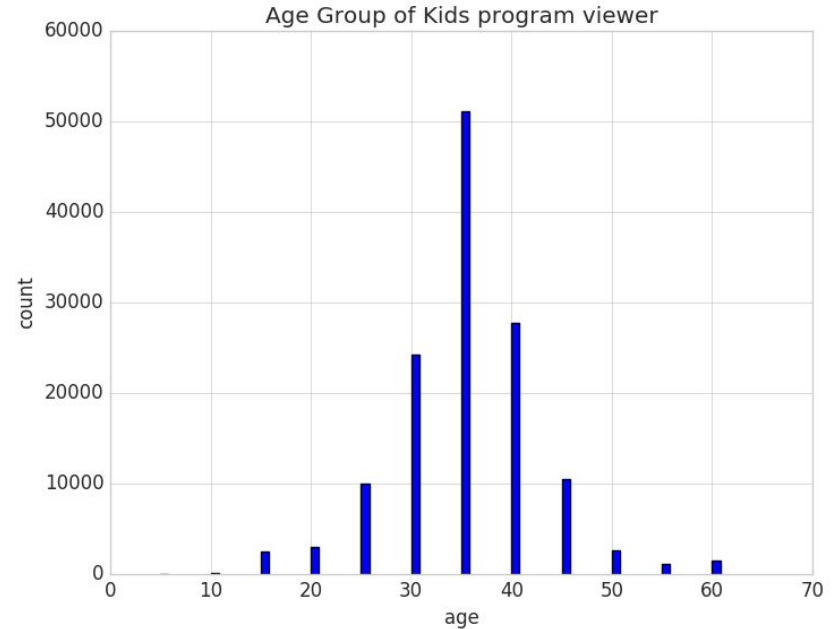
What, When) Content Category

- **Kids** video steady popularity
- **News** in the evening
- Sports in the evening



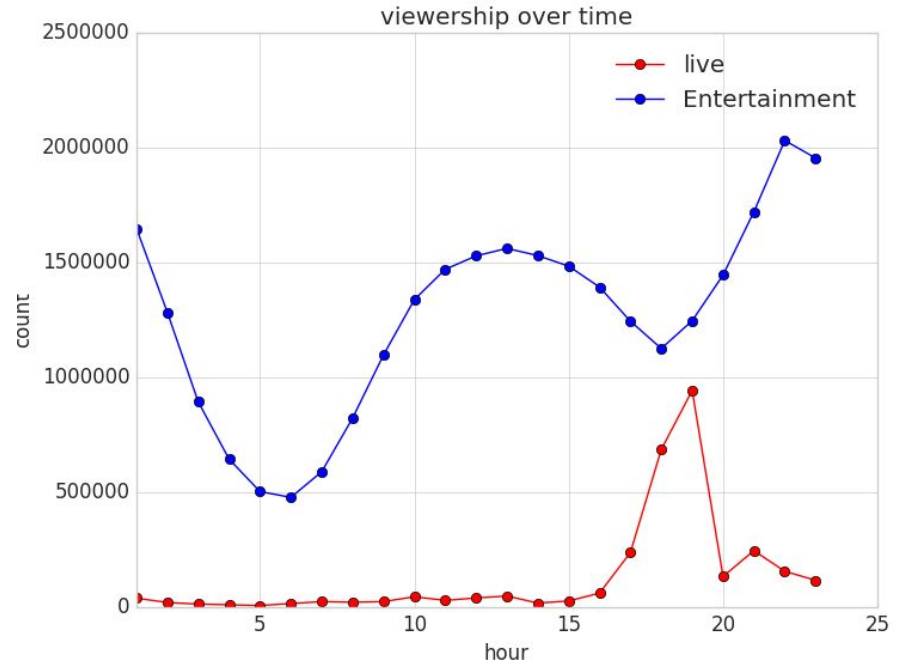
What, When) Kids by Age

- Age 30~40 viewers
 - for their children



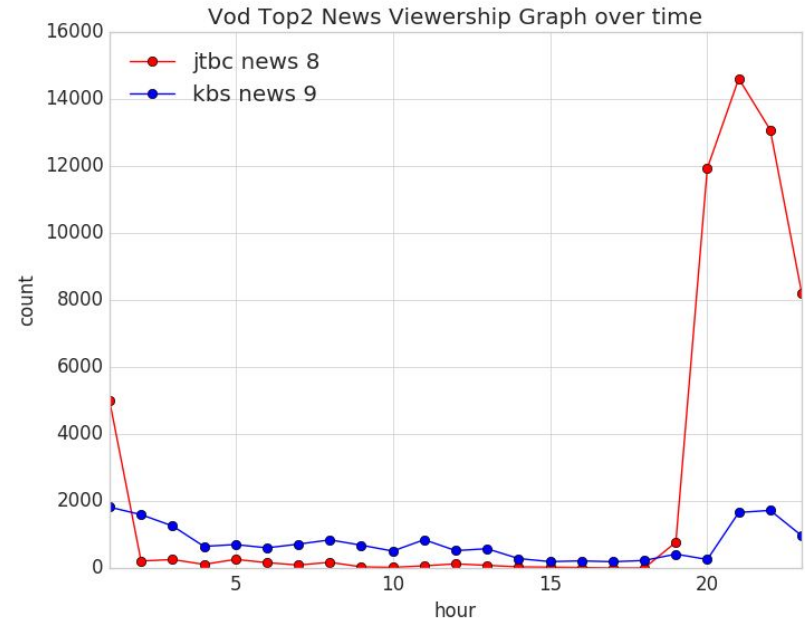
What, When) Live vs VoD ?

- Live top 1 content at 7 - 8 PM
- Entertainment vod at 10 - 11 PM



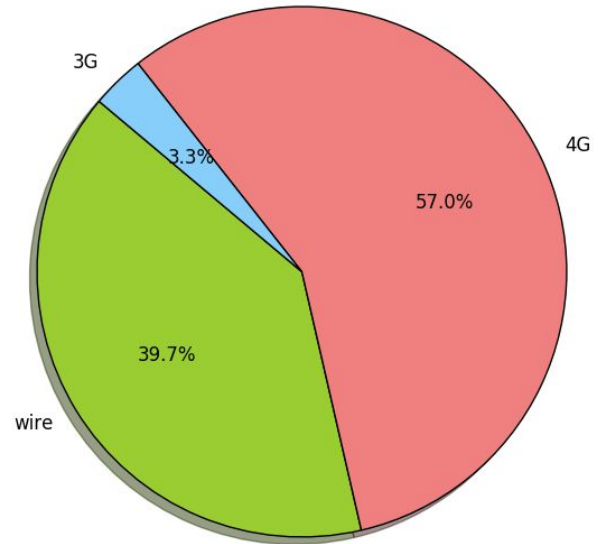
What, When) Content Category: News

- News peak at 20 o'clock
 - Live



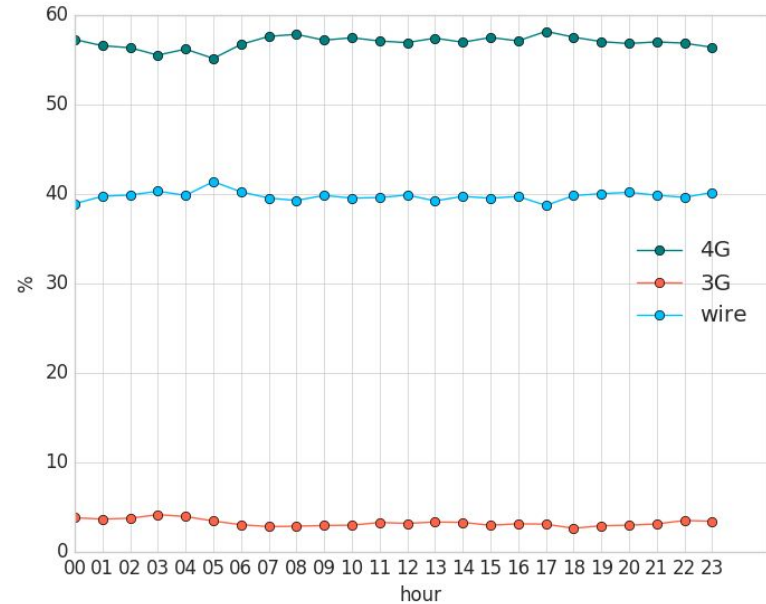
How) 4G, 3G, Wired(WLAN)

- 4G > Wired(WLAN) > 3G
 - 4G: LTE
 - Wired(WLAN): broadband network



How) Hourly 4G, 3G, Wired(WLAN) Breakdown

- Steady state
 - LTE: 57%
 - Wired(WLAN): 40%
- WLAN traffic offloading?



Things under Work and To Do

1. Data analysis and visualization on the cloud and big data platform
 - Detailed analysis
 - Spark and R on Amazon EC2
2. Content recommendation problem
 - Automatic video highlight generation
3. Where to place contents for users by OTT providers
 - Prefetching/caching contents by recommendation
 - Device vs. edge vs. CDN vs cloud
 - LTE/WLAN offloading