

Blocking malicious users

...

Loris D'Acunto

Final Project - W205-4

Context

My company provides
Audio Content Recognition
as a service – think about a
Shazam for TV with B2B
model



A Customer

Reward app that gives points to the users for how much TV and commercial they watch



Customer Data

- 60,000 MAU
- 3.7 Billion events
- 1.4 TB of data



Log File

- 45 fields tab separated
- One field is a JSON with 5 different formats

```
1457395200090 LogRequestTVLive
9c6c2092a3a04828 137975332-
50d8z7NSrdjFLOQP 1 0 49 1457395177504
29.8501077 -90.1087257 2 null null - 4
104.48.148.156 1457394918285 -6 en US android
4.2.2 GT-P3113 - 70 0.5 17 espresso
espredssosasue com.perk.livetv.aphone 7d1dc183-
25ac-4e95-8103-700cab37d812 false 100.0 true 86
WnpjeEprQkNVUTVZSUJ0TA== true
Y1VKZkYwaE9XMTVVQms5VEVHNWJRVWM9
{"image":"http://labs.fyitelevision.com/FYIMediaServ
ices/v_3_3_6_1/Station.svc/107/99/Logo/120x120","
iab_program_topics":"","stop":"2016-03-08
00:00:00","topics":[],"channel":"E!
Entertainment","start":"2016-03-07
23:00:00","genre":"","id":"49","program":"Keeping
Up with the
Kardashians","iab":"","type":"livetv","program_topics
":[ ]} 3.3 2.0.1 TVLIVE H E! Entertainment Keeping
Up with the Kardashians
```

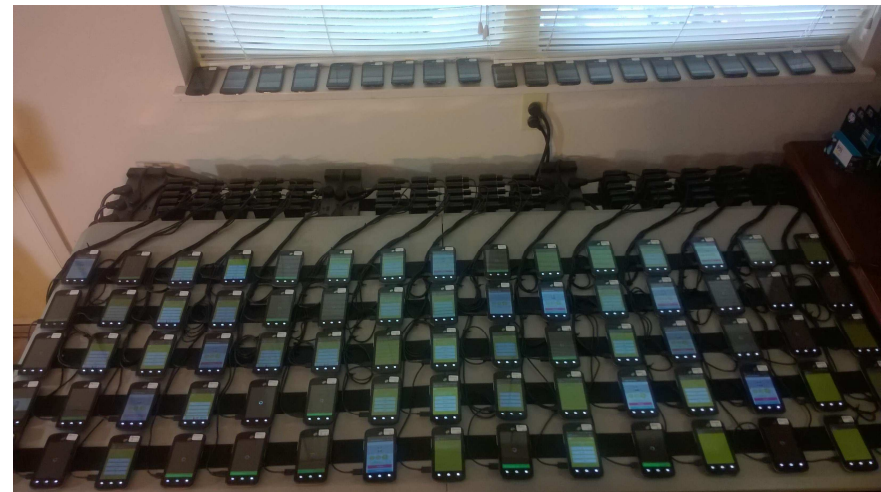
The Problem

Find and block malicious
users

Entry level



Pro level

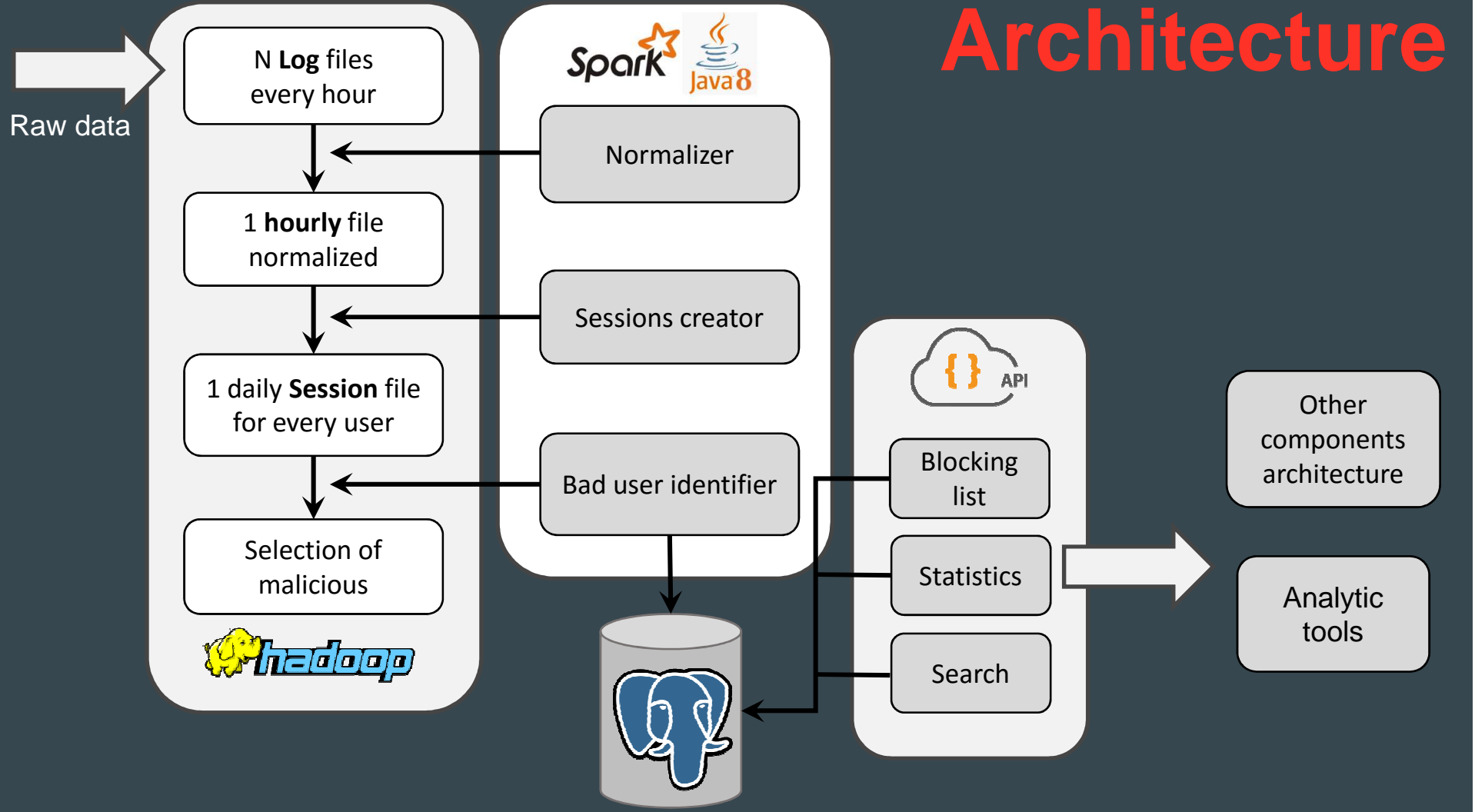


How to find them?

Search for unusual behaviours like watching the same commercial all the day



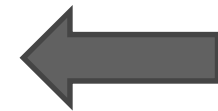
Architecture



Findings

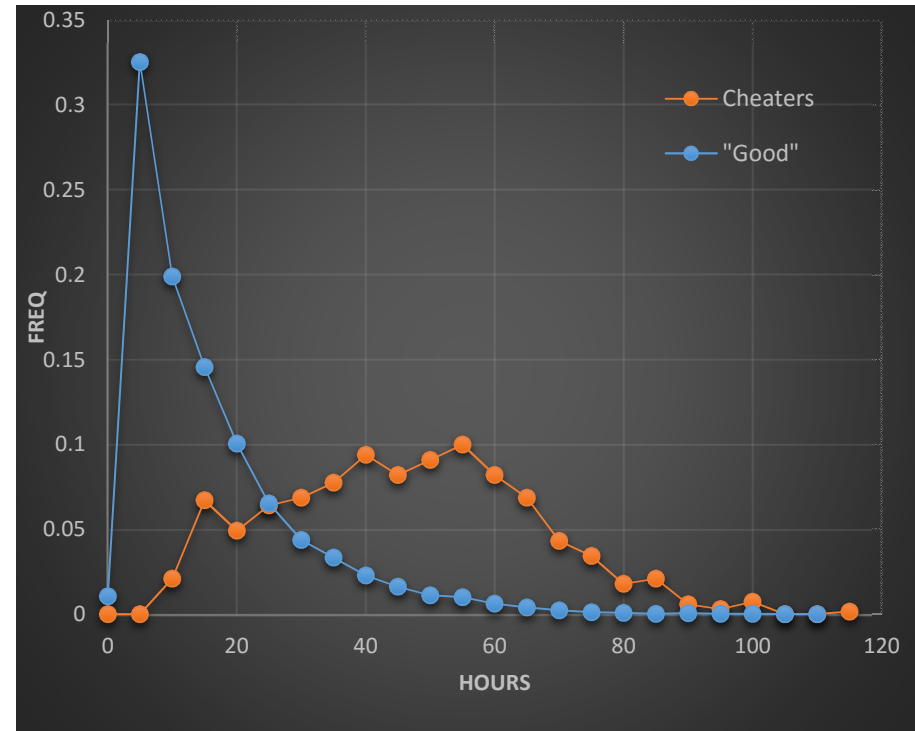
8% of devices are
malicious and generate
30% of queries

Devices	Users
1	644
2	56
3	17
4	14
5	3
6	5
7	3
9	2
10	2
11	2
15	1



Findings

A normal user watches less TV than a malicious one



Next steps

Prototype -> Production

