

**Содержание**

**14 Identification of corporate subscribers ..... 3**



# 14 Identification of corporate subscribers



We recommend using the [QoE tool](#) and the [GUI to report on the resale service search](#). Below is an example of an evaluation by analyzing raw DPI data.

There are several approaches to solve the problem of identification corporate subscribers that use the tariff plan. For example it can be done based on the number of sessions generated by the subscriber. The ISP can collect the full NetFlow and generate a relevant report with the help of DPI and netflow analysis tools.

This use case suggests estimating the number and the type of devices that were used to provide the Internet access.

First, let's record [Clickstream metadata](#) during a day or at least few hours.

Then add to the `/etc/dpi/fastdpi.conf` configuration file the following settings:

```
ajb_save_url=-1
ajb_save_url_format=ipsrc:uagent
```

If subscribers are identified by their login (dynamic IP), please specify `login` instead of `ipsrc` in the format option.

You can expand the format with additional fields according to your needs, for example:

```
ajb_save_url=-1
ajb_save_url_format=ts:login:ipsrc:ipdst:host:path:uagent:ref:cookie:tphost:
blockd:method
```

Please restart the service.

The data is written to the `/var/dump/dpi` directory by default

Make sure that there is enough disk space and keep in mind that you have to turn off the recording whenever you'd finish a job by setting the `ajb_save_url = 0` option or just by commenting corresponding `ajb_save_url=-1`; reload it upon completion. So we get a set of files with the `url_*.txt` names as a result. The user-defined fields are separated by tab therein.

Let's execute a query applying to the `url_*.txt` file. The output shows how many subscriber devices each subscriber has. They are identified by the User-Agent from http requests:

for a short format `ajb_save_url_format = ipsrc: uagent` it will look like this

```
sort -u /var/dump/dpi/url_*.txt|cut -f1|uniq -c|sort -n
```

and for a long format and identification by login, let's cut out relevant fields using the `cut` utility:

```
cut -f2,7 /var/dump/dpi/url_*.txt|sort -u|cut -f1|uniq -c|sort -n
```

We get a report like:

```
...
  87 SUBS_9987153
  97 SUBS_9802207
 105 SUBS_4924486
 107 SUBS_4979880
...
```

here the first field corresponds to the number of different devices subscriber uses

Further you can check out the types of devices used by subscribers, and sometimes it can lead you to a conclusion about the type of business. For example, if the mobile phones are mostly presented (every day is different), then this is likely a cafe.

```
cut -f2,7 /var/dump/dpi/url_*.txt|grep "^SUBS_9802207[[:blank:]]"|sort -u
```

here SUBS\_9802207 is a login of the subscriber of our concerns

```
...
1C+Enterprise/8.3
C530A IP/42.199.00.000.000;C530H/107.037.00.000.000
C530A IP/42.207.00.000.000;C530H/107.037.00.000.000
Dalvik/1.6.0 (Linux; U; Android 4.4.2; 6037Y Build/KOT49H)
Dalvik/2.1.0 (Linux; U; Android 5.1.1; D5503 Build/14.6.A.1.216)
Dalvik/2.1.0 (Linux; U; Android 6.0.1; Redmi Note 3 MIUI/7.3.9)
iPhone8,1/10.3.1 (14E304)
...
```

The presence of ERP-software, VoIP phones, a lot of different computers and phones provides proof that such subscriber is a corporation.