Содержание

Идентификация корпоративных абонентов через CLI

Рекомендуем использовать инструмент QoE и GUI для составления отчетов по поиску перепродажи услуг. Ниже приведен пример оценки путем анализа сырых данных, полученных с DPI.

Существуют разные подходы к решению задачи идентификации корпоративных абонентов, которые сидят на физических тарифах. Например, это можно сделать на основе количества сессий, генерируемых абонентом, для чего можно снять средствами DPI полный NetFlow и построить по нему соответствующий отчет в инструментах анализа NetFlow.

В данном сценарии предлагается воспользоваться оценкой количества и типа устройств, с которых осуществляется доступ в интернет.

Сначала запишем метаданные Clickstream за сутки или хотя бы несколько часов.

Добавляем в файл /etc/dpi/fastdpi.conf настройки

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

При идентификации абонентов по login (динамических ip) в настройке формата вместо ipsrc укажите login

Можно расширить формат дополнительными интересующими нас полями, например:

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

Сделайте рестарт сервиса.

Запись данных осуществляется по умолчанию в каталог /var/dump/dpi

Убедимся что на этом диске достаточно места и не забудем выключить запись про окончанию работы, установив настройку ajb_save_url=0 или закомментировав ее и сделав reload.

В результате получим набор файлов с именами url_*.txt в котором заданные нами поля разделены табуляцией.

Выполним над ним запрос, который покажет сколько у каждого абонента абонентских устройств, которые идентифицируем по User-Agent из http запросов):

для короткого формата ajb_save_url_format=ipsrc:uagent он будет выглядеть следующим образом

```
```

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

а для длинного формата и идентификации по login предварительно вырежем нужные нам поля с помощью cut

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

Получим отчет вида:

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

где первое поле - количество различных устройств у абонента

Дальше, при желании, можно посмотреть типы устройств по интересующим абонентам, и иногда из этого можно сделать вывод о типе бизнеса. Например, если там в основном мобильные телефоны (каждый день разные), то это кафе.

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

где SUBS_9802207 - login интересующего нас абонента

```
...
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)
...
```

наличие 1С, IP-телефонов, множества разношерстных компьютеров и телефонов явно говорит о бизнес характере данного абонента