

Министерство науки и высшего образования Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»
(Университет ИТМО)

Факультет безопасности информационных технологий

Управление мобильными устройствами

Лабораторная работа №2
Обработка и тарификация трафика NetFlow
Вариант 1

Работу выполнила:
студентка группы N3351
Петухова Дарья



Проверил:
Федоров И.Р

Санкт-Петербург
2020

Цель работы: обработать трафик NetFlow v5 из файла, написать программный модуль для тарификации пользователя, определить зависимость объема трафика от времени.

Задачи: преобразовать файл `nfcapd.202002251200` в пригодный для обработки, сформировать файл для тарификации, построить график зависимости объема трафика от времени, протарифицировать трафик в соответствии с вариантом задания.

Ход работы

Для реализации данного программного модуля был выбран Python 3.7.

1. С помощью утилиты `nfdump`, команда **`nfdump -r nfcapd.202002251200`**, было произведено чтение файла `nfcapd.202002251200`.
2. Т. к. в файле было большое количество данных и соответствующих им IP-адресов, была проделана выборка данных по IP-адресу, представленному в варианте задания.

`nfdump -r nfcapd.202002251200 -o csv 'ip 217.15.20.194' > net.csv`

```
ts,te,td,sa,da,sp,dp,pr,flg,fwd,stos,ipkt,ibyt,opkt,obyt,in,out,sas,das,smk,dmk,dtos,c
2020-02-25 11:30:02,2020-02-25 11:30:02,0.000,173.194.73.95,217.15.20.194,443,62596,I
2020-02-25 11:30:02,2020-02-25 11:30:02,0.000,108.177.14.94,217.15.20.194,443,56364,I
2020-02-25 11:30:03,2020-02-25 11:30:03,0.000,81.19.104.45,217.15.20.194,443,50297,T
2020-02-25 11:30:08,2020-02-25 11:30:08,0.000,23.226.231.226,217.15.20.194,56019,80,
2020-02-25 11:28:03,2020-02-25 11:30:09,125.800,13.94.251.244,217.15.20.194,443,5867
2020-02-25 11:21:32,2020-02-25 11:30:10,518.320,23.226.231.226,217.15.20.194,43635,8
2020-02-25 11:27:17,2020-02-25 11:30:17,180.020,68.232.34.200,217.15.20.194,443,5029
2020-02-25 11:30:07,2020-02-25 11:30:18,11.090,92.118.38.58,217.15.20.194,38220,25,T
2020-02-25 11:30:21,2020-02-25 11:30:21,0.000,23.226.231.226,217.15.20.194,43228,80,
2020-02-25 11:30:23,2020-02-25 11:30:23,0.000,173.194.222.95,217.15.20.194,443,59516
2020-02-25 11:28:08,2020-02-25 11:30:24,135.240,77.111.247.41,217.15.20.194,443,5135
2020-02-25 11:30:24,2020-02-25 11:30:24,0.000,172.253.5.2,217.15.20.194,35312,53,UDF
2020-02-25 11:30:24,2020-02-25 11:30:24,0.000,74.125.41.3,217.15.20.194,35120,53,UDF
-----
```

Рисунок 1. Фрагмент данных из преобразованного файла

3. С помощью программной реализации был построен график зависимости объёма трафика от времени.

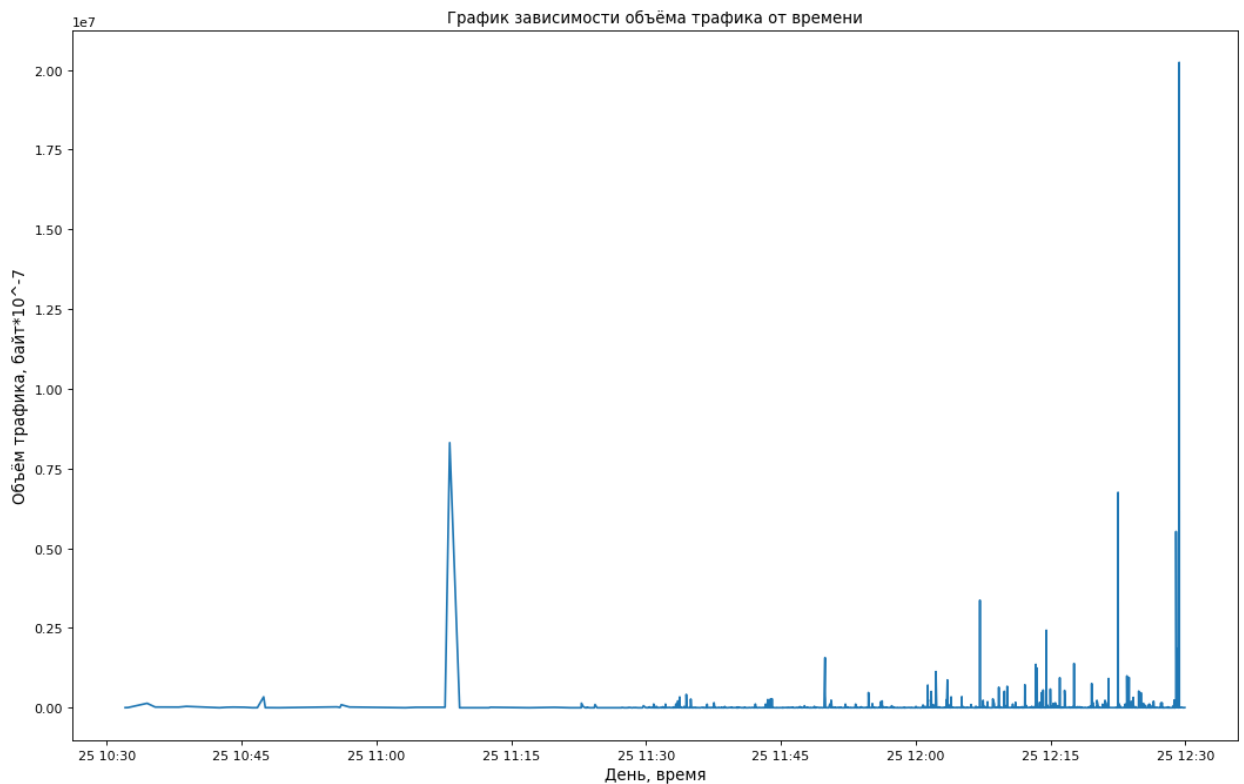


Рисунок 2. График зависимости объёма трафика от времени

4. Был протарифицирован трафик пользователя с IP-адресом 217.15.20.194.

Результат тарифицирования представлен на рисунке 3.

Стоимость услуг: 88.33 рублей

Рисунок 3. Результат тарифицирования

Листинг программы

```
import csv
from datetime import datetime
import numpy as np
from matplotlib import figure
from matplotlib import pyplot as plt
%matplotlib inline

filepath = 'net.csv'

sum_bytes = 0
k_traffic = 0.5
```

```

bytes_arr = []
date_time = []
list_d_t = []

#открытие файла и обработка данных в нем
with open(filepath, newline='') as file:
    reader = csv.DictReader(file, delimiter=',')
    for row in reader:
        bytes_arr.append(row['ibyt'])
        bytes_arr = bytes_arr[:4823]
        date_time.append(row['ts'])
        date_time = date_time[:4823]
    for byte in bytes_arr:
        bytes_arr = [int(byte) for byte in bytes_arr]
    for b in range(len(bytes_arr)):
        sum_bytes += bytes_arr[b]
        b += 1
    for t in range(len(date_time)):
        date_time[t] = datetime.strptime(date_time[t], '%Y-%m-%d
%H:%M:%S')
        list_d_t.append(date_time[t])

#тарификация
Q = sum_bytes / 10**6
X = round(Q * k_traffic, 2)
print('Стоимость услуг: ' + str(X) + ' рублей')

#построение графика
dict_t_b = dict(zip(list_d_t, bytes_arr))
lx = list(dict_t_b.keys())
lx.sort()
for i in lx:

```

```
d[i] = dict_t_b[i]
ly = list(d.values())
for l in ly:
    l = l / 10**6
x = np.array(lx)
y = np.array(ly)
plt.figure(figsize=(16,10), dpi= 80)
plt.title('График зависимости объёма трафика от времени',
fontsize=12)
plt.xlabel("День, время", fontsize=12)
plt.ylabel("Объём трафика, байт*10-7", fontsize=12)
plt.plot(x, y)
plt.show()
```