МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

Национальный исследовательский университет ИТМО

ФАКУЛЬТЕТ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Управление мобильными устройствами

Лабораторная работа № 2

«Обработка и тарификация трафика NetFlow»

Работу выполнил:

Студент группы №3352
Александрович Д. А. _____

Работу проверил:

Федоров И. Р. _____

Санкт-Петербург, 2020

Цель работы: изучение технологии работы протокола NetFlow, а также разработка и реализация программного модуля обработки трафика NetFlow v5 и тарификации абонента.

Исходный код реализован на языке Python версии 3.7.1 в формате Jupiter тетради. Для работы исходного кода требуется наличие модулей Pandas, NumPy и MatPlotLib.

Файл nfcapd.202002251200 был преобразован в файл формата csv для работы с Pandas при помощи команды nfdump -r nfcapd.202002251200 -o "fmt:%ts,%sa,%da,%ibyt,%obyt" | sed s/'\s'//g | head -n -4 > lab2.csv

За обработку и просмотр статистики отвечает функция plot_traffic. Ниже представлен исходный код программы:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from matplotlib.dates import DateFormatter
%matplotlib inline
```

```
IP = '217.15.20.194'
```

```
def plot_traffic(times, values, format=None, save_to=None):
    fig, ax = plt.subplots(figsize=[15,8])
    ax.plot(times, values)
    myFmt = DateFormatter(format if format is not None else "%H:%M:%S")
    ax.xaxis.set_major_formatter(myFmt)
    plt.gcf().autofmt_xdate()
    plt.ylabel('Количество байт в пакете')
    plt.xlabel('Временная шкала')
    plt.title('Зависимость объема трафика от времени')
    if save_to:
        plt.save(save_to)
    else:
        plt.show()
```

```
df = pd.read_csv('lab2.csv', skiprows=1, header=None)
df.columns = ['t', 'sa', 'da', 'ibys', 'obys']
df = df[np.logical_or(df.sa == IP, df.da == IP)]
df.ibys = df.ibys.apply(lambda row: int(row) if 'M' not in row else (int(float(row[:-1])*10**6)))
df.t = df.t.apply(lambda row: row[10:18])
print(f'Всего транзакций с IP = {IP}: {df.shape[0]} шт.')
df.head(3)
```

```
outgoing_traffic = df[df.sa == IP].ibys.sum() / 10**6
ingoing_traffic =  df[df.da == IP].ibys.sum() / 10**6

print(f'Исходящий трафик составил {outgoing_traffic:0.2f} Мб.')
print(f'Входящий трафик составил {ingoing_traffic:.02f} Мб.')
```

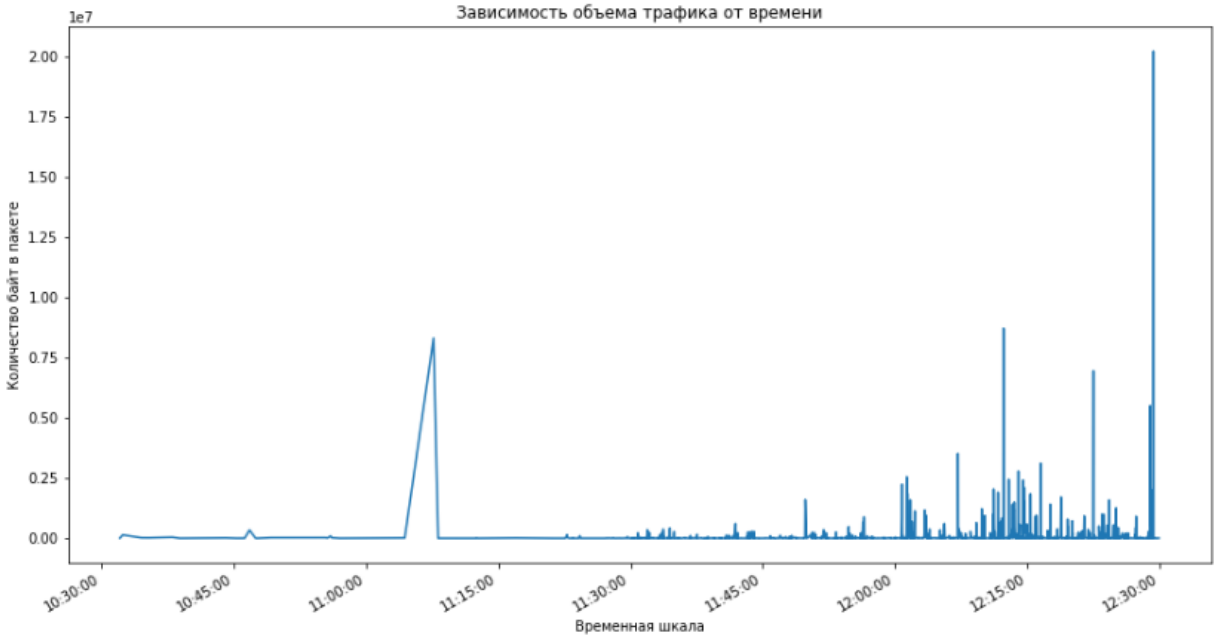
```
times = np.sort(df.t.unique())
values = []
for t in times:
    values.append(df.loc[df.t == t, ['ibys', 'obys']].sum().sum())

plot_traffic(pd.to_datetime(times), values)
```

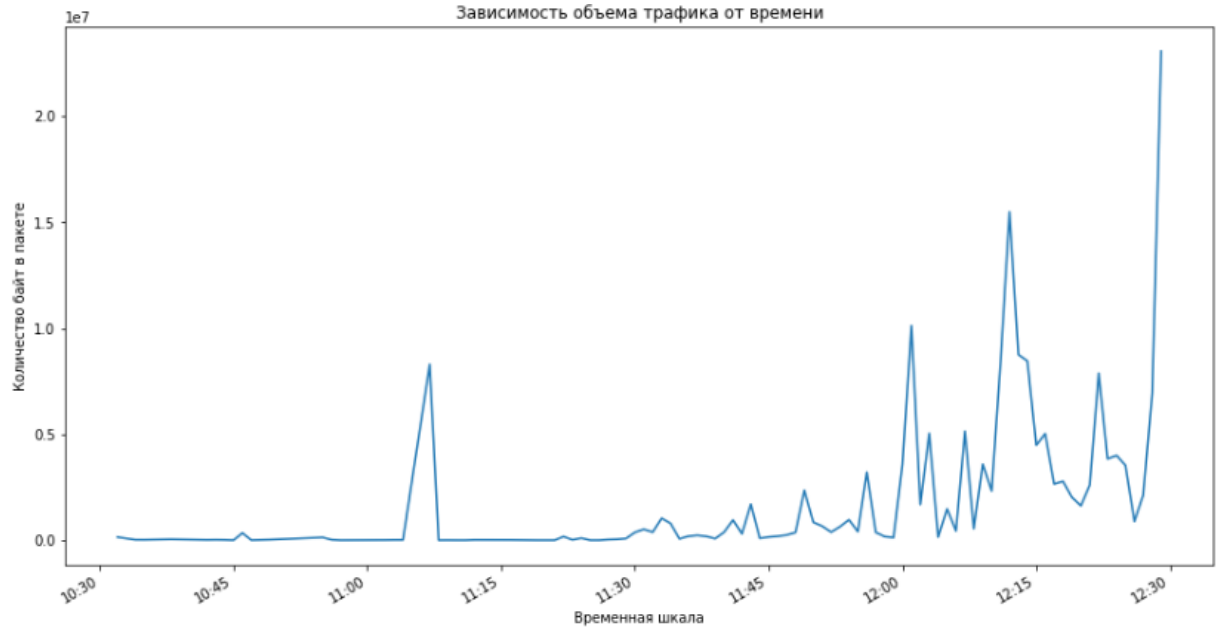
```
hm = df.t.apply(lambda row: row[:5])
times_hm = np.sort(hm.unique())
values_hm = []
for t in times_hm:
    values_hm.append(df.loc[hm == t, ['ibys', 'obys']].sum().sum())

plot_traffic(pd.to_datetime(times_hm), values_hm, format="%H:%M")
```

```
k = 0.5
print(f'Счет за входящий трафик: {k * ingoing_traffic:.2f} р.')
print(f'Счет за исходящий трафик: {k * outgoing_traffic:.2f} р.')
```

В ходе работы программы было выявлено, что исходящий трафик составил 0 байт, входящий – 176.81 Мб.

Также в ходе работы было построено два графика зависимости объема трафика – посекундно, и, так как для построения посекундного графика не хватило данных, поминутно. Ниже представлены оба графика:





Далее была произведена тарификация с коэффициентом k=0.5. Счет за входящий трафик составил 88.40 рублей, за исходящий – 0 рублей.

Вывод

В ходе данной работы были изучены технологии работы протокола NetFlow, а также разработан и реализован программный модуль обработки трафика NetFlow v5 и тарификации абонента.