**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«Национальный исследовательский университет ИТМО»**

**Факультет безопасности информационных технологий**

**Дисциплина:**

**«Управление мобильными устройствами»**

**Лабораторная работа № 2 на тему**

**«Обработка и тарификация трафика NetFlow»**

**Выполнил:**

Студент гр. N3349

Нуралиев К. А.

**Проверил:**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Федоров И. Р.

Санкт-Петербург

2020г.

**Цель работы:** В данной работе необходимо обработать трафик NetFlow v5 из файла nfcapd.202002251200. Для чего изначально требуется привести данный файл в читабельный вид (проще всего это сделать с помощью утилиты nfdump), после чего нужно сформировать собственный файл для тарификации любого формата, с которым удобно работать (в соответствии с вариантом работы), после чего необходимо построить график зависимости объема трафика от времени (любым удобным образом) и наконец требуется протарифицировать трафик в соответствии с вариантом задания.

**Теория:**

NetFlow — это протокол, разработанный компанией Cisco и предназначенный для сбора информации об IP-трафике внутри сети. Маршрутизаторы Cisco анализируют проходящий через интерфейс трафик, суммируют данные и отправляют статистику в формате NetFlow на специальный узел, называемый NetFlow Сollector. NetFlow часто используется для ведения биллинга или для анализа трафика сети. Протокол существует в нескольких версиях, последняя версия 9 предназначена для учёта трафика между АС (Автономная Система) и в импортируемых данных имеет несколько дополнительных полей таких как АС источника, АС назначения и пр., но обычно, для биллинга в несложной сети внутри одной АС достаточно информации, содержащейся в данных NetFlow версии 5.

В данной работе предполагается обработка трафика NetFlow v5 из файла nfcapd.202002251200

Правила тарификации услуг “Интернет”:

X = Q \* k, где

• X – итоговая стоимость,

• Q – общий объем трафика NetFlow за отчетный период,

• k – множитель тарифного плана (у каждого варианта свой).

**Задание (1 вариант):** Протарифицировать абонента с IP-адресом 217.15.20.194

с коэффициентом k: 0,5руб/Мб

**Исходный файл**: «nfcapd.txt» & «new\_nfcapd.txt».

**Ход работы**:

Изначально с помощью команды «nfdump -r nfcapd.202002251200» был сформирован файл «nfcapd.txt». После обработки получился такой файл «new\_nfcapd.txt». Для написания ПО для тарификации был выбран язык программирования python (версия 3.8).

Используемые библиотеки: matplotlib, pandas

Основой выбора данного языка программирования является: простота понимания программного кода и мгновенная коррекция в случае изменения условий тарификации; наличие встроенных библиотек для обработки файлов csv формата и построения графиков.

**Выходные данные ПО**: total size, total cost и зависимость объема трафика от времени (сохраняется в виде картинки "plot.png»).

**Исходный код:**

#You should have these modules, otherwise program won't work.

#To install these modules, you should type these commands in terminal: pip install pandas; pip install matplotlib

import matplotlib.pyplot as plt

import pandas as pd

# My data

my\_ip = '217.15.20.194'

price = 0.5

def start():

# Getting started

text = []

with open('nfcapd.txt', 'r', encoding='utf\_8') as f:

text = f.readlines()

spaces = [' ' \* n for n in range(20, 2, -1)]

for i in range(len(text)):

text[i] = text[i].replace('->', '')

words = ['INVALID', 'XEvent', 'Ignore']

for word in words:

for i in range(len(text)):

text[i] = text[i].replace(word, ' ' + word + ' ')

for space in spaces:

for i in range(len(text)):

text[i] = text[i].replace(space, ' ')

for i in range(len(text)):

text[i] = text[i].replace(' ', ',')

with open('new\_nfcapd.txt', 'w', encoding='utf\_8') as f:

f.writelines(text)

# Getting needed data

df = pd.read\_csv('new\_nfcapd.txt', sep=',')

total\_traffic = df[(df['Src IP Addr:Port'].str.contains(my\_ip)) | (

df['Dst IP Addr:Port'].str.contains(my\_ip))].reset\_index(drop=True).copy()

total\_traffic['Date first seen'] = total\_traffic['Date first seen'].apply(pd.to\_datetime).copy()

total\_traffic['In Byte'] = total\_traffic['In Byte'].apply(

lambda x: float(x) if x[-1] != 'M' else float(x[:-1]) \* 1024 \* 1024).copy()

result = sum(total\_traffic['In Byte'].values)

result = round((result / 1024), 2)

print(f'Total size: {result}kb')

# Showing graphics

total\_traffic['Date first seen'] = total\_traffic['Date first seen'].apply(lambda x: str(x)[:str(x).find('.')]).copy()

total\_traffic['Date first seen'] = total\_traffic['Date first seen'].apply(pd.to\_datetime).copy()

total\_traffic['Date first seen'] = total\_traffic['Date first seen'].apply(lambda x: str(x)[:-3]).copy()

total\_traffic['Date first seen'] = total\_traffic['Date first seen'].apply(pd.to\_datetime).copy()

traffic\_after\_group = (total\_traffic.groupby('Date first seen').aggregate(sum)).copy()

plt.figure(figsize=(15, 10))

plt.title('Minutes')

plt.plot(traffic\_after\_group.index, traffic\_after\_group['In Byte'].values)

plt.show()

# Calculating

total\_price = result \* price

total\_price = round(total\_price, 2)

print(f'Total cost: {total\_price}')

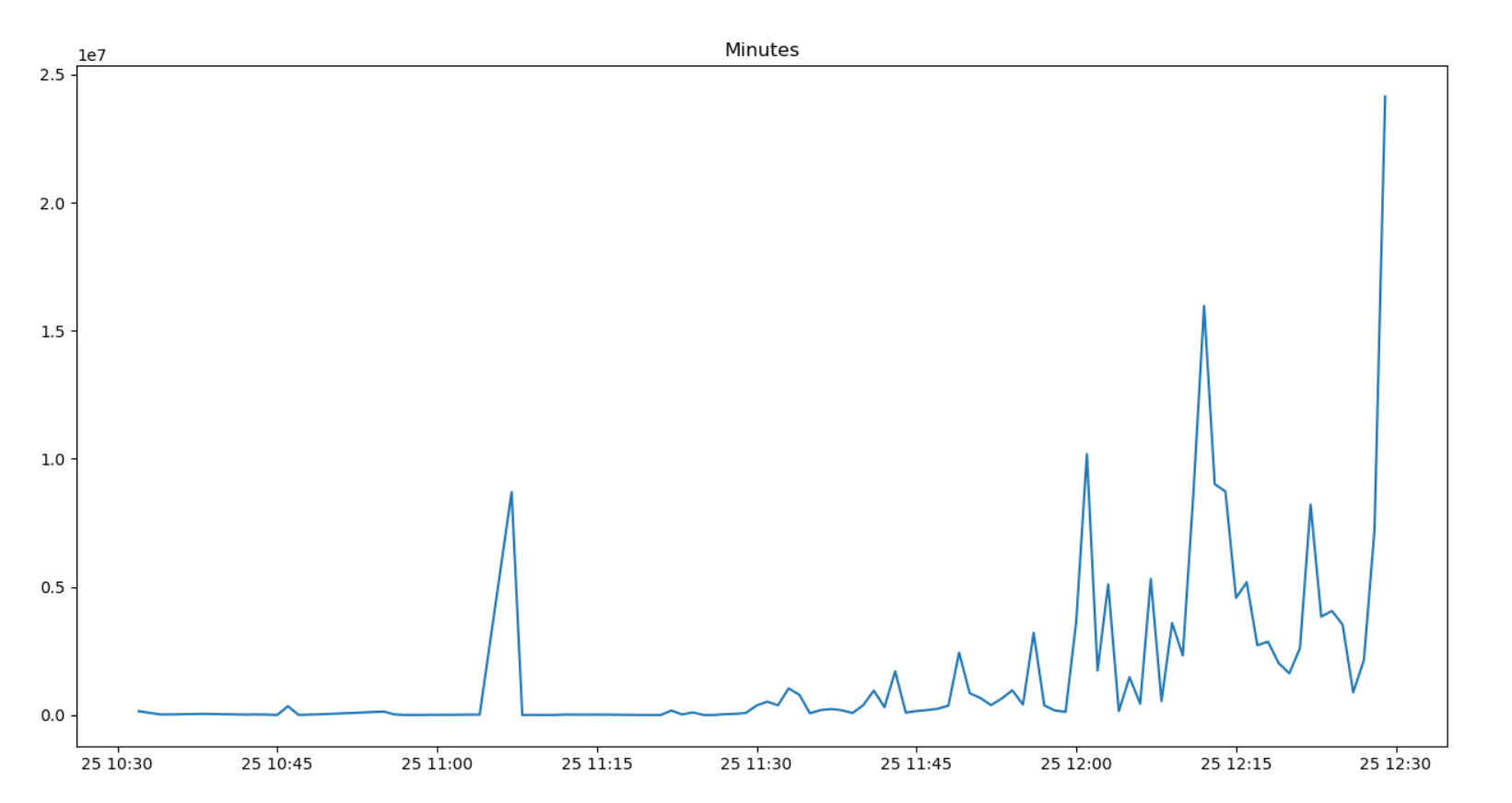
start()

**Вывод программы**:

Первые и последние строки вывода программы:



График (файл «plot.png»):



**Вывод**: в результате работы была написана программа, позволяющая протарифицировать абонента и подсчитать итоговую стоимость использования услуги «Интернет» с учётом определённого тарифа, также позволяющую пострить график зависимости объема трафика от времени.