Telegram chat statistics

Anastasiia Chumak
ČVUT–FIT
chumaana@fit.cvut.cz

7. ledna 2024

1 Introduction

This report describes the coursework created for the BI-PYT course in winter term B231. The coursework deals with the topic of web scraping, which is a technique for obtaining web data by extracting it from the pages of web resources. Web scraping can be done manually by a computer user, however the term usually refers to automated processes implemented using code that performs GET requests to a target site. In our program we will use telegram development tools(telethon).

Telegram is a cloud-based instant messaging application and communication platform that prioritizes speed, security, and user privacy. It allows users to send messages, photos, videos, and files, as well as create groups and channels for communication and dissemination of information.

Key Features of Telegram:

Security and Privacy: Telegram employs end-to-end encryption for Secret Chats, ensuring only the intended recipient can access the messages. It also offers features like self-destructing messages and passcodes for accessing the app.

Cloud-Based: Messages and media are stored in the cloud, enabling seamless synchronization across devices without compromising data security.

Versatile Communication: Apart from individual and group chats, Telegram supports channels for broadcasting messages to an unlimited number of subscribers, making it suitable for news dissemination or community interaction.

File Sharing: Users can share files of various formats (documents, videos, images, etc.) up to a specified size limit.

Bots and APIs: Telegram provides a platform for bot development, allowing developers to create automated services, games, and more within the app. It also offers APIs for third-party integrations and app development.

Cross-Platform Support: Telegram is available on multiple platforms, including mobile devices (iOS, Android), desktop (Windows, macOS, Linux), and web browsers.

The purpose of this work is to create a CLI program that will receive a telegram chat on analyses, analyses it and display statistics data using tkinter.

2 Used models

We are using telethon here to get access to telegram chats. Telethon is a powerful Python library de- signed to streamline interactions with the Telegram API. It serves as a bridge between your Python code and the Telegram messaging platform, enabling developers to create bots, automate messaging tasks, manage chats, and perform various actions programmatically.

For GUI (Graphical User Interface) we are using tkinter. Tkinter is a standard GUI toolkit that co- mes bundled with Python. It allows developers to create graphical applications with a consistent look and feel across platforms. Tkinter is based on the Tk GUI toolkit and provides a set of tools for building desktop applications with widgets like buttons, labels, entry fields, and more.

SQlite is used to store data. This is done in order not to wait every time for the end of the query of all chat messages, however, it does not solve the problem: when a new message arrives in the chat, it can be stored in the table and analysed only in a repeated query. SQLite is a popular lightweight, self- contained, server less, and zero-configuration SQL database engine. It's embedded directly into the application that uses it, eliminating the need for a separate server process or setup. SQLite databases are stored as a single file on disk and are widely used in various applications and embedded systems due to their simplicity and efficiency.

Also for more beautiful displaying we have used tkkbootstrap

3 Main functionality

Analyses and tests were performed on the basis of a single chat.

After requesting all messages of the selected chat

and recording them into the database, the work goes faster - statistics on this chat is displayed: total number of messages, number of messages broken down by type, number and types of messages for each user, etc. Here we do not use queries to telegram, we work with the previous data.
there is a field for entering a period of time, in order to get statistics for a certain period, the period can be changed: enter new data and press the book and the information will be updated.

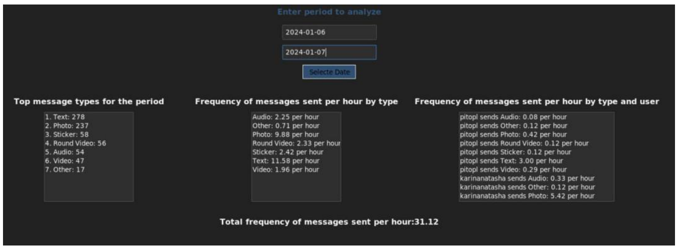
At the end of the program all database tables are cleared, but the client’s session is not cleared, if desired you can log in again by phone number and there will be no need to auth.

4 Results

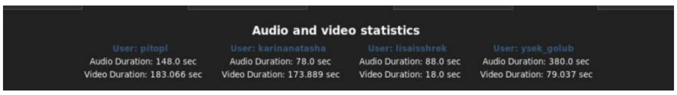
Analysis and tests was one chat based. Statistics is showed Picture1-Picture4. There are also displayed statistics per different period

5 Conclusion

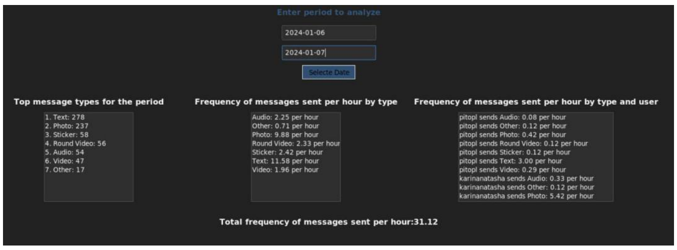
As a result we get statistics for one selected chat, this statistics can be extended with more positions. Also there are implemented feature to get statistics per period, but to update data and display newer data, you have to restart program, which is not convenient. Program works fast with small chats, for the bigger ones it can take some time. It can be improved in the future



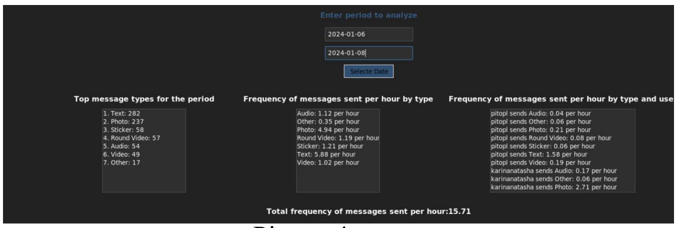
Picture 1



Picture 2



Picture 3



Picture 4

Reference

[1] How Many People Use Telegram in 2023? 55 Telegram Stats Backlinko. [Electronic resource] // – Access mode: https://backlinko.com/telegram-users

[2] Python GUI Programming With Tkinter [Electronic resource] // – Access mode: https://realpython.com/pytho gui-tkinter/

[3] Telethon API [Electronic resource] // – Access mode: https://tl.telethon.dev/index.html

[4] Largest catalog of Telegram channels. Statistics, analytics, TOP chart. Telegram Analytics. Largest catalog of Telegram channels. Statistics, analytics, TOP chart. Telegram Analytics. [Electronic resource] // – Access mode: https://tgstat.com/