

Vk-selections

Утилита для создания выборок пользователей на основе списков друзей и подписчиков групп.

Язык: C++

Кузнецов Владислав, г. 251

github.com/fizvlad

Содержание

- 1) Работа с API VK. Вызов методов и общение с ботом через диалог
- 2) Создание выборок. Сбор и хранение данных, обработка формулы запроса
- 3) Работа с ботом
- 4) TODO

Работа с API VK

Для работы с API используется отправка GET запросов с помощью функции.

```
apiRequest(method, parameters, token, version)
```

Отправка запроса осуществляется с помощью функций библиотеки CURL.

Результатом работы функции является json объект. Для работы с json используется header-only библиотека nlohmann/json.

Работа с API VK

Для общения с ботом используется объект класса ChatBot, способный ожидать сообщения и отвечать на них.

Ожидание сообщения реализовано в виде **longpoll-запросов**;

```
ChatBot bot(token, group_id);  
bot.setOnlineStatus(true);  
  
bot.start([&] (Message m, bool isOld) {  
  
    cout << (isOld ? "(Old message) " : "") <<  
    m.sender << " send message to " << m.receiver <<  
    " with id " << m.id << " at " << m.ts <<  
    ". Content: " << endl << m.text << endl;  
  
    bot.markAsRead(m);  
});
```

Создание выборок

Каждая выборка представляет из себя набор пар type-id.

Для хранения используется бинарный файл. Каждый Unit в нём занимает 5 байт. Таким образом, выборка из 15 миллионов пользователей (самая крупная группа на данный момент) будет занимать в памяти порядка 75МБ.

Источником выборки служит множество друзей пользователя или множество подписчиков группы.

```
Unit user = Unit(123456, true, stoken);  
Unit group = Unit("exclusive_muzic", gtoken);  
  
Selection s = music.members(gtoken) &&  
             !user.friends(gtoken);
```

Работа с ботом

Работа с ботом осуществляется через диалог сообщества.

Для запроса выборки надо отправить сообщения вида

!selection (id123.friends || mmspbu.members) && habr.auto

В ответ бот пошлёт информацию о выборке и текстовый файл, содержащий id пользователей.



Владислав 22:47
!selection fizvlad_temp.auto



Fizvlad | Bot 22:47
Запрос получен

Выборка не инвертирована. Размер: 114



Selection.txt
2 КБ

Дальнейшее развитие

- Разделение работы на два потока и создание „очереди обработки“
- Добавление новых критериев для фильтрации пользователей: страна, город, семейное положение, образование, место работы и пр.
- Работа с API: упрощение работы с token-ами
- Выборки из групп
- Оптимизация хранения данных
- Доработка кода