

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

ТЕХНОЛОГІЇ ІНФОРМАЦІЙНОГО ПОШУКУ, Ч.2

МЕТОДИЧНІ ВКАЗІВКИ

до виконання лабораторної роботи №8
Робота з торрент-клієнтами. Google Trends

Мета: навчитися використовувати мережевий протокол для обміну файлами, отримати навички роботи з торрент-клієнтом, навчитися відстежувати тенденції в інформаційному пошуку за допомогою сервісу Google Trends.

Теоретичні відомості

Торрент це система обміну файлами, в основі якої лежить протокол BitTorrent і принцип P2P.

BitTorrent – це мережевий протокол, який розроблений для обміну файлами. Фактично, BitTorrent це набір правил, які використовують торрент-клієнти і торрент-трекери під час процесу обміну файлами. У такій мережі усі учасники маю однакове становище і права. Іншими словами, усі учасники мережі одночасно є як серверами, так і клієнтами. Завдяки використанню цього принципу дані можуть передаватися не лише від джерела до одержувача, але і від одного одержувача до іншого.

Торрент-клієнт – програма розроблена для обміну файлами за протоколом BitTorrent. Зараз існує безліч торрент-клієнтів. Найбільш популярні з них це: µTorrent, BitTorrent, BitComet, BitSpirit.

Торрент-трекер або BitTorrent-трекер – це сервер, який управляє файлообміном між окремими торрент-клієнтами. З точки зору користувача, торрент-трекер це сайт, на якому можна ознайомитися з доступними для скачування файлами. За допомогою торрент-трекера окремі учасники файлообміну обмінюються даними про частини файлів, які вони вже викачали і можуть передати іншим користувачам. При цьому торрент-трекер не зберігає файлів, якими обмінюються користувачі, і не бере участь в обміні файлами. Окрім цього торрент-трекер веде облік об'єму даних, який викачав кожен з користувачів. На багатьох торрент-трекерах існують обмеження, які не дозволяють викачувати файли, не беручи участь у файлообміні. Для цього обчислюється рейтинг.

Рейтинг – значення, яке обчислюється торрент-трекером для кожного користувача. Як правило, рейтинг обчислюється за формулою об'єм викачаних

файлів/об'єм відвантажених даних. Рейтинг може використовуватися для обмеження користувачів, які тільки викачують, але не відвантажують свої дані.

Торент-файл – це невеликий файл з розширенням .torrent. У цьому файлі зберігається інформація про дані, які поширюються через файлообмінну мережу BitTorrent, а саме:

- про розмір і кількість фрагментів (файл розбивається на кілька тисяч частин) і контрольну суму викачуваного файлу (файлів);
- про трекер, на якому можна отримати інформацію про сідерів і лічерів, що поширюють цей файл.

Фактично, торрент-файл – це ярлик, який однозначно ідентифікує вміст інформації, що передається, і адресу, де можна отримати цю інформацію. Зазвичай він має об'єм приблизно 100Кб. Для того, щоб завантажити певний файл, необхідно завантажити торрент-файл і відкрити його через торрент-клієнт.

Роздача – набір файлів і тек, які поширюються через торрент в одному комплекті. Роздача може складатися з одного файлу (наприклад, якщо це відео) або з великої кількості файлів (наприклад, якщо це музика).

Кожен користувач, який починає викачувати або роздавати файл на торренті автоматично стає піром. Пір – користувач, який бере участь в обміні файлами. Таким чином, кількість пірів – це загальна кількість користувачів, які викачують або роздають той чи інший торрент. Для кожного торрента (для кожної роздачі на торрент-трекері) є своя кількість пірів.

Піри поділяються на сідів і лічів.

Сід – учасник файлообміну, який має повний об'єм роздачі і відвантажує дані. Першим сідом є користувач, який створив роздачу.

Ліч – користувач, який викачує роздачу. Після того, як ліч повністю викачає роздачу, він перетворюється на сіда. Також під терміном “ліч” може розумітися недобросовісний користувач, який не хоче ділитися викачаними файлами.

Швидкість скачування у файлообмінній мережі BitTorrent залежить від кількості пірів і сидів на роздачі, а також швидкості, з якою вони можуть роздавати файли. Оскільки кожен пір (ліч) може роздавати ту частину файлів, яку він вже завантажив, то швидкість скачування зростає разом зі збільшенням кількості пірів (лічів).

Схема роботи протоколу

Перед початком завантаження клієнт під'єднується до трекера за адресою, вказаною в торрент-файлі, повідомляє йому свою адресу і хеш-суму торрент-файлу, на що у відповідь клієнт отримує адреси інших клієнтів, що викачують або роздають той самий файл. Далі клієнт періодично інформує трекер про хід процесу і отримує оновлений список адрес. Цей процес називається **оголошенням**.

Клієнти з'єднуються один з одним і обмінюються сегментами файлів без безпосередньої участі трекера, який лише зберігає інформацію, отриману від підключених до обміну клієнтів, список самих клієнтів та іншу статистичну інформацію. Для ефективної роботи мережі BitTorrent необхідно, щоб якомога більше клієнтів були здатні приймати вхідні з'єднання.

При з'єднанні клієнти відразу обмінюються інформацією про наявні у них сегменти. Клієнт, що бажає викачати сегмент (лічер), посилає запит і, якщо другий клієнт готовий віддавати, отримує цей сегмент. Після цього клієнт перевіряє контрольну суму сегменту. Якщо вона збіглася з тією, що записана в торрент-файлі, то сегмент вважається успішно скачаним, і клієнт оповіщає всіх приєднаних бенкетів про наявність у нього цього сегменту. Якщо ж контрольні суми відрізняються, то сегмент починає викачуватися заново. Деякі клієнти банять тих бенкетів, які занадто часто віддають некоректні сегменти.

Таким чином, обсяг службової інформації (розмір торрент-файлу і розмір повідомлень зі списком сегментів) безпосередньо залежить від кількості, а значить, і розміру сегментів. Тому при виборі сегмента необхідно дотримувати баланс: з одного боку, при великому розмірі сегмента обсяг службової

інформації буде меншим, але у випадку помилки перевірки контрольної суми, доведеться завантажувати в кілька разів більше інформації. З іншого боку, при малому розмірі помилки не такі критичні, тому що необхідно заново завантажити менший обсяг, проте розмір торрент-файлу і повідомлень про наявні сегменти стає більше.

Переваги та недоліки BitTorrent

До переваг BitTorrent належить:

- Відсутність черг на скачування;
- Файл завантажується одночасно від всіх, хто завантажує і сіда;
- Користувачі обмінюються вже викачаними частинами, тому що більше користувачів, то більшою є швидкість;
- Завантажені фрагменти стають негайно доступні іншим клієнтам;
- Здійснюється контроль створюваних торрентів на унікальність.

BitTorrent має також низку недоліків.

- Недоступність роздачі. У випадку, якщо роздача непопулярна, то може виникнути ситуація, коли немає жодного сіда, а даних у присутніх бенкетів не вистачає, щоб завершити скачування. У такому випадку необхідно чекати появи або сіда, або бенкету, що має сегменти, які відсутні у інших.
- Відсутність анонімності. Принцип роботи BitTorrent-протоколу передбачає, що кожному клієнту відомі IP-адреси, як мінімум, інших клієнтів, отримані від сервера. Використання різноманітних розширень протоколу в деяких випадках дозволяє дізнатися також і адреси інших бенкетів з рою. Тому користувачі незахищених систем та клієнтів можуть зазнавати атак. Можна також дізнатися адреси користувачів, які обмінюються контрафактним контентом і подати на них до суду.
- Проблема лічерів. Деякі користувачі, особливо на трекарах, які не потребують реєстрації, не підтримують роздачу після завершення скачування, що веде до зменшення загальної продуктивності. Ця

проблема є однією з причин популярності приватних торрент-трекерів, що враховують кількість завантаженого/відданого .

- Персоналізація. Протокол не передбачає використання ніків. Відсутній чат між бенкетами. Неможливо переглянути список файлів бенкету (в пошуках інших файлів , які могли б зацікавити). Більшість цих функцій реалізовано в інших протоколах.

- Чітерські програми (англ. to cheat – шахраювати). Існують програми, що дозволяють проводити махінації з рейтингом на торрент-трекерах (RatioMaster, GiveMeTorrent). Такі програми відсилають трекеру сфальсифіковані дані про кількість завантаженого і відданого.

Користування торрент-клієнтом

Існує велика кількість торрент-клієнтів, але найбільш популярним і простим є BitTorrent. Свою популярність BitTorrent отримав завдяки простому і зрозумілому інтерфейсу, а також високій швидкості роботи і не вимогливості до системних ресурсів.

BitTorrent – програма розроблена творцем протоколу BitTorrent. Існують версії для Microsoft Windows і Mac OS. Після того, як компанія BitTorrent Inc придбала програму µTorrent, програма BitTorrent ґрунтується на початкових кодах µTorrent і практично повністю повторює її інтерфейс.

Для початку роботи з торрент-файлами необхідно перейти на офіційний сайт BitTorrent (<http://www.bittorrent.com/>) і завантажити безкоштовну програму (Free Download).

Встановлення всіх додатків, які пропонує інсталяційний пакет, не є обов'язковим. Після швидкого встановлення, відкриється головне вікно програми (Рис. 1).

Головне вікно програми має такі вкладки:

- Файли – інформація про файли, що входять в торрент;
- Інфо – інформація про завантажуваний файл;
- Піри – інформація про пірів, які беруть участь в роздачі;

- Оцінки – оцінки цього файлу іншими пірами;
- Трекери – трекери, які керують активною роздачею;
- Швидкість – графік швидкостей завантаження/відвантаження.

Стрічка команд містить такі кнопки (Рис. 2):

- Add Torrent – додати торрент шляхом додавання попередньо збереженого файлу з жорсткого диску, наприклад;
- Add Torrent from URL – додати торрент шляхом додавання URL-адреси, за якою розміщений файл;
- Create New Torrent – створити роздачу;
- Remove – видалити обраний файл;
- Start Torrent – запуск завантаження;
- Stop Torrent – зупинка завантаження;
- Move Up Queue – пересування обраного файлу на 1 позицію вище у списку, що означає його більшу пріоритетність при завантаженні;
- Move Down Queue – пересування обраного файлу на 1 позицію нижче у списку, що означає його меншу пріоритетність при завантаженні;

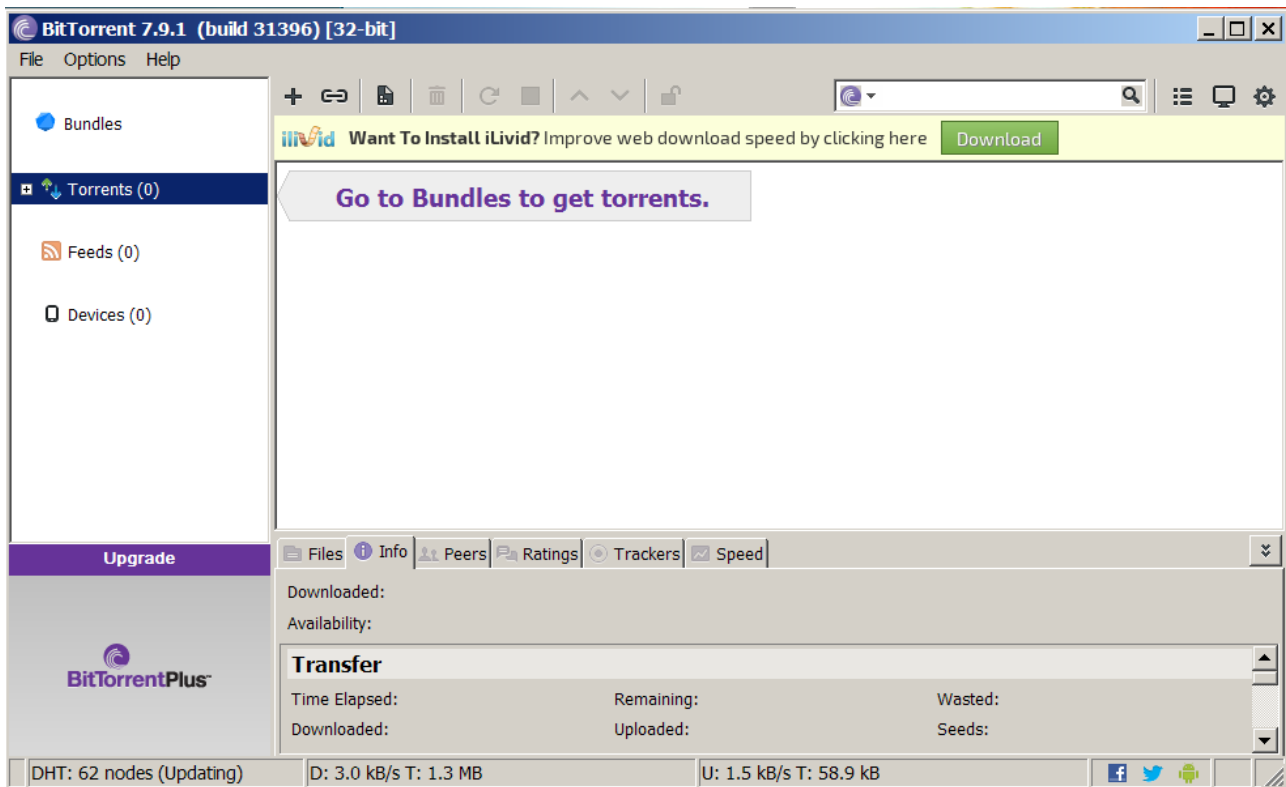


Рис.1. Головне вікно програми BitTorrent



Рис.2. Стрічка команд

З торрент-трекера, наприклад, Гуртом (<http://toloka.hurtom.com>) завантажуюмо необхідний файл (Рис. 3), використовуючи лінк-магніт та встановлений торрент-клієнт (Рис.4).




Skryabin, Taras CHubay i druzi - Nashe Rizdvo [1999] FLAC.torrent		
Трекер:	Зареєстрований	 Завантажити  Дивитись онлайн
Зареєстрований:	2010-06-30 10:45	
Розмір:	204 MB	
Подякували:	97 Дякую! (список)	
Оцінка музики:	9.7/10 (Голосів: 29) 	
Список файлів завантаження		

Рис.3. Опис торрент-файла

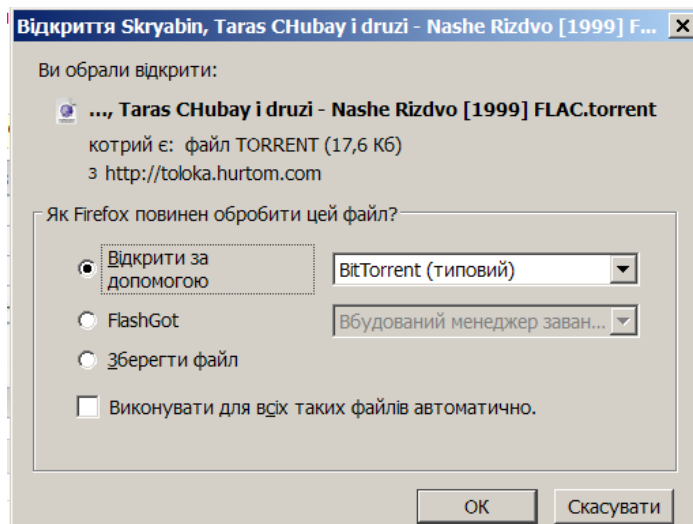


Рис.4. Відкривання через лінк-магніт

Після завантаження, відкривається вікно програми (Рис.5.), яке дозволяє налаштувати місце зберігання завантаженого файлу, його ім'я, невідкладний початок завантаження, а також перелік файлів, які вкладені і які необхідно завантажувати.

Коли відповідні налаштування встановлені, розпочнеться завантаження. Інформацію щодо активного завантаження можна отримати на вкладках, призначення яких описано вище (Рис.6).

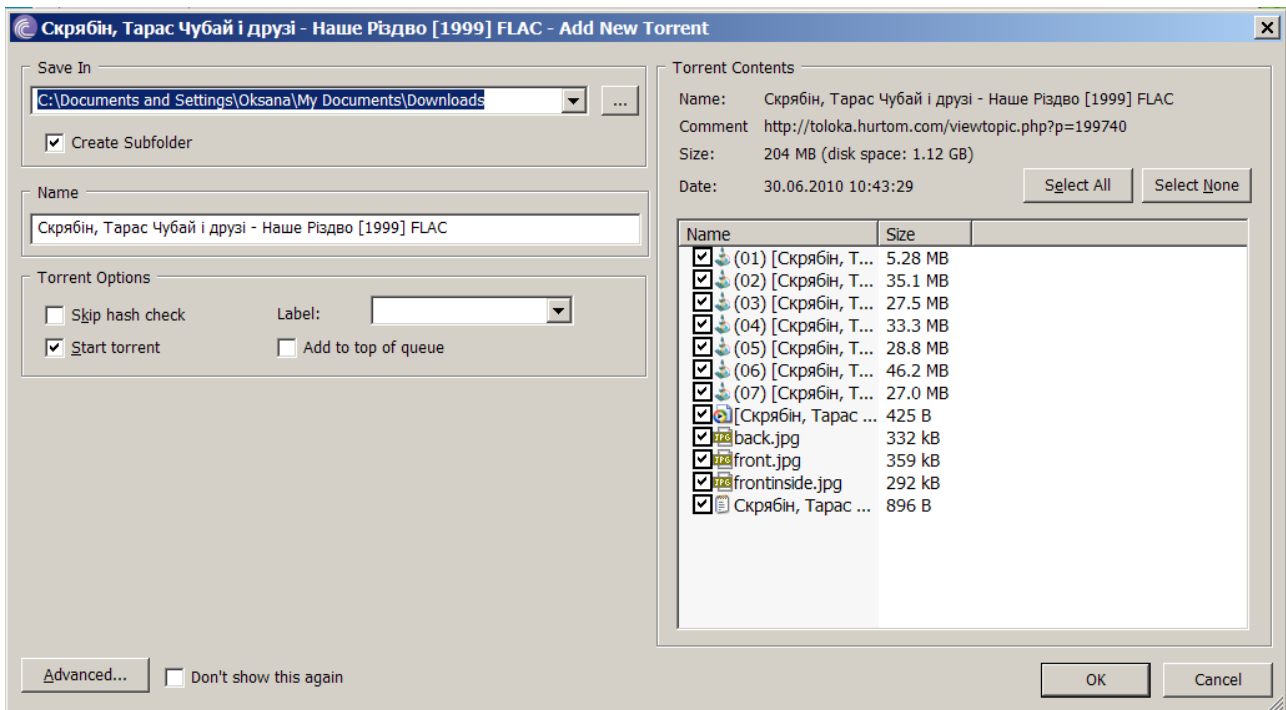


Рис.5. Додавання нового торрент-файла для завантаження через лінк-магніт

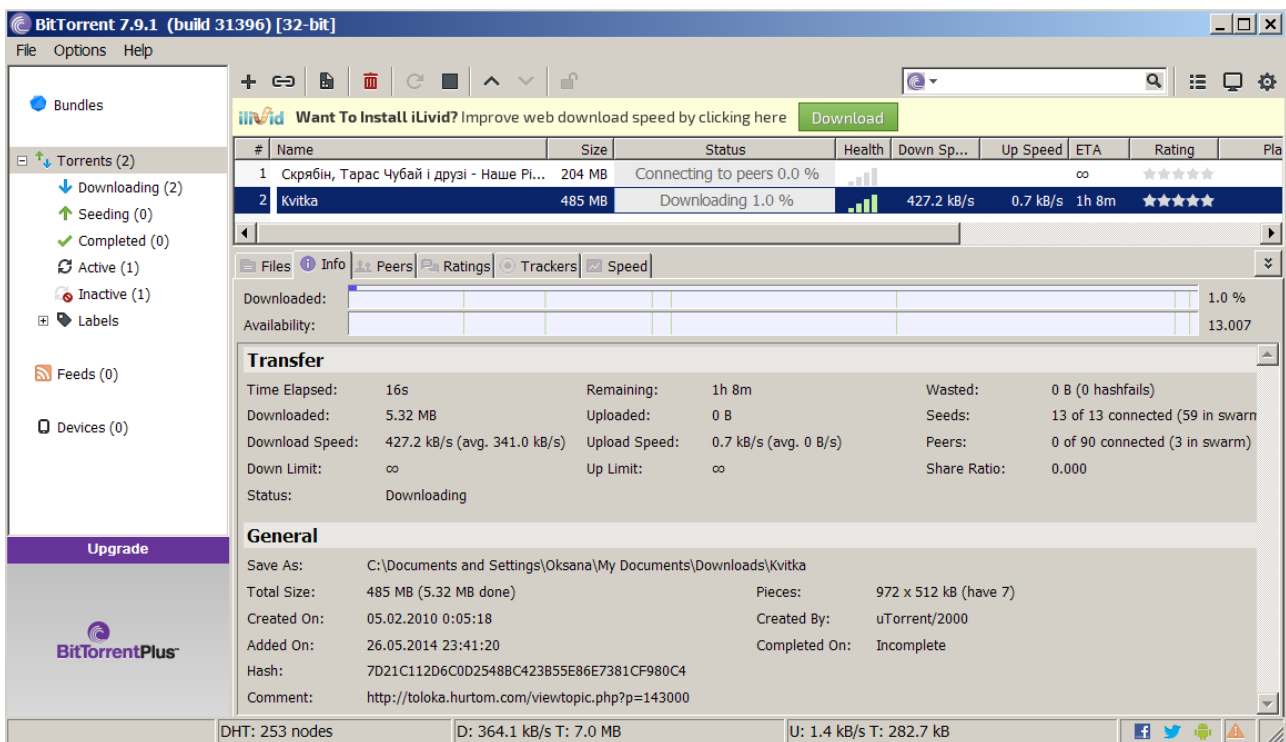


Рис.6. Інформація щодо активного завантаження

Google Trends

У Google є сервіс, за допомогою якого можна дивитися і аналізувати тренди пошукових запитів, по всіх регіонах, де є пошук Google. Це – Google Trends - <http://trends.google.com/trends/explore> (Рис.7).

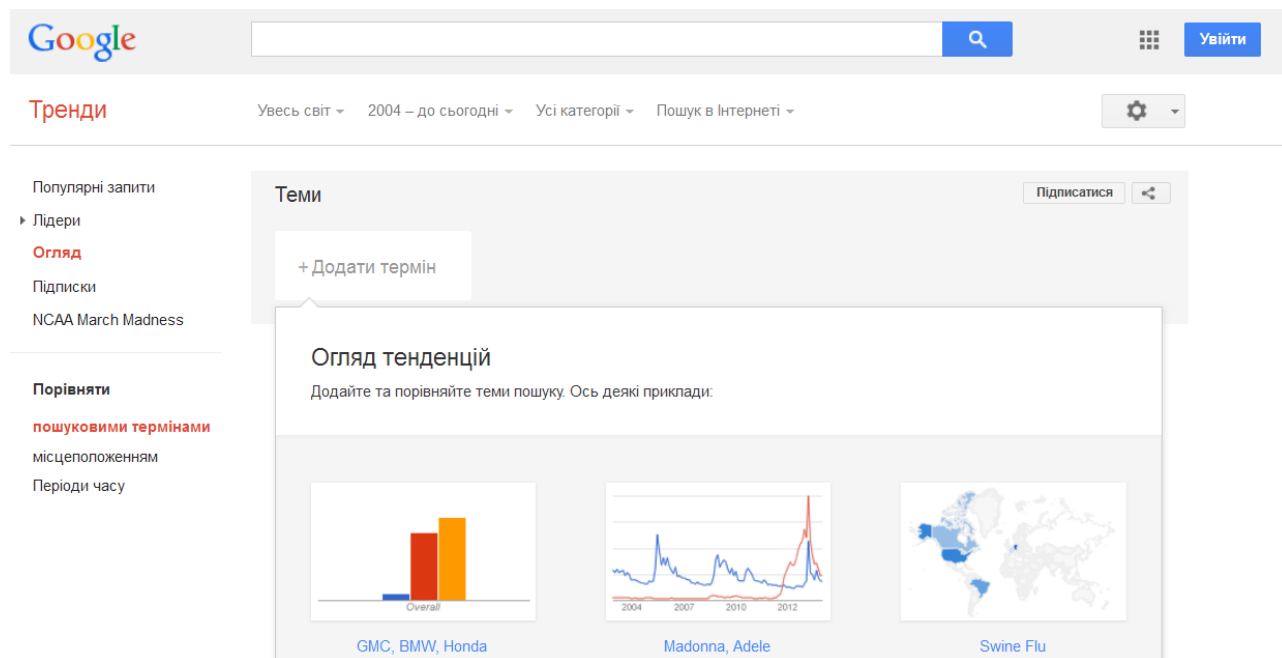


Рис. 7. Google Trends – головне вікно

Google Trends необхідний для того, щоб:

- Аналізувати сезонність запитів;
- Дивитися трендові пошукові слова;
- Дивитися розвиток ключових слів;
- Аналізувати попит.

Для відстежування тенденцій вводимо ключове слово і вибираємо регіон (Україна) (Рис.8).

Як ми бачимо, Google Trends відстежує динаміку з 2007 року. На графіках видно, що у тематики “хутро” є 2 сезони, зимовий і літній. Зимовий сезон більш активний, а літній менше. Подивимося, які області і міста є найактивнішими, і які популярні слова, пов'язані з “хутро” (Рис.9 і Рис.10).

Також можна переглянути швидко зростаючі слова в цій ніші, вибираючи вкладку “Зростаючі”(Rising).

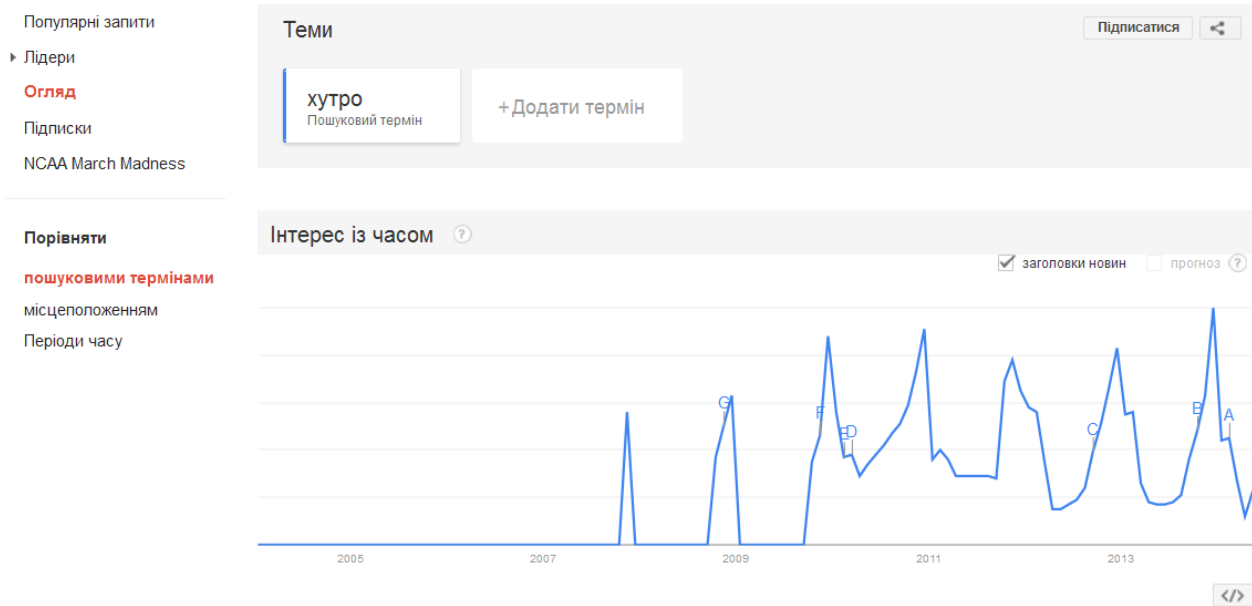


Рис.8. Відстеження тенденції за одним ключовим словом

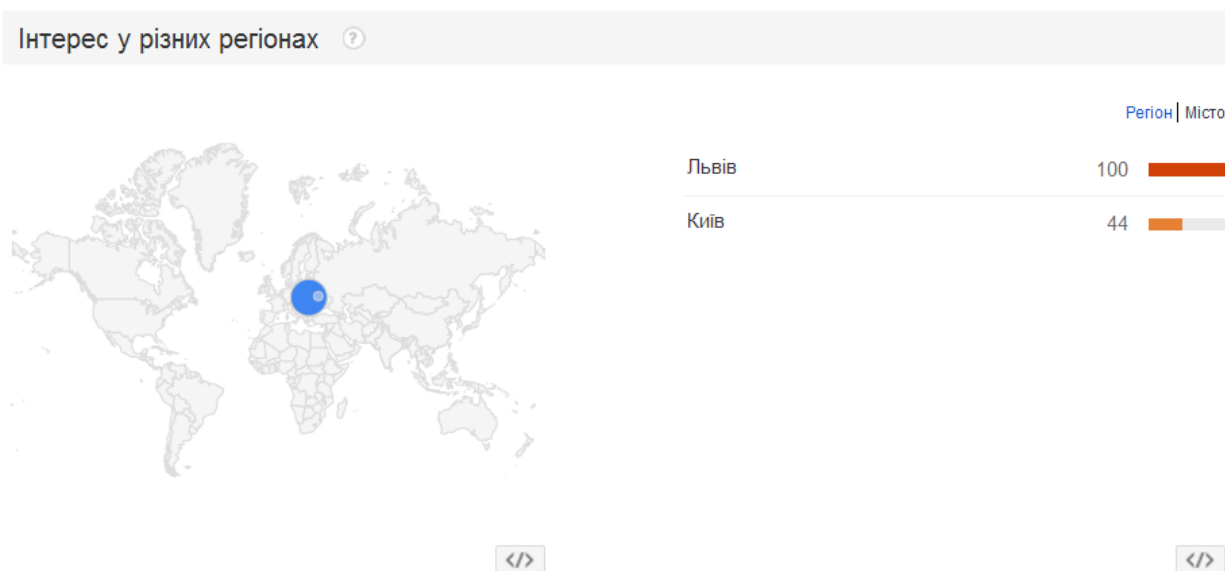


Рис.9. Рівень інтересу в різних регіонах

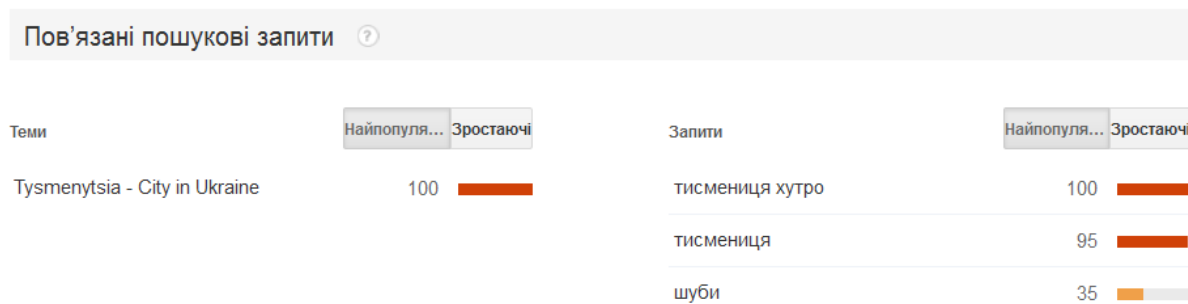


Рис.10. Пов'язані пошукові запити

Так само можна порівняти пошукові запити (Рис.11).

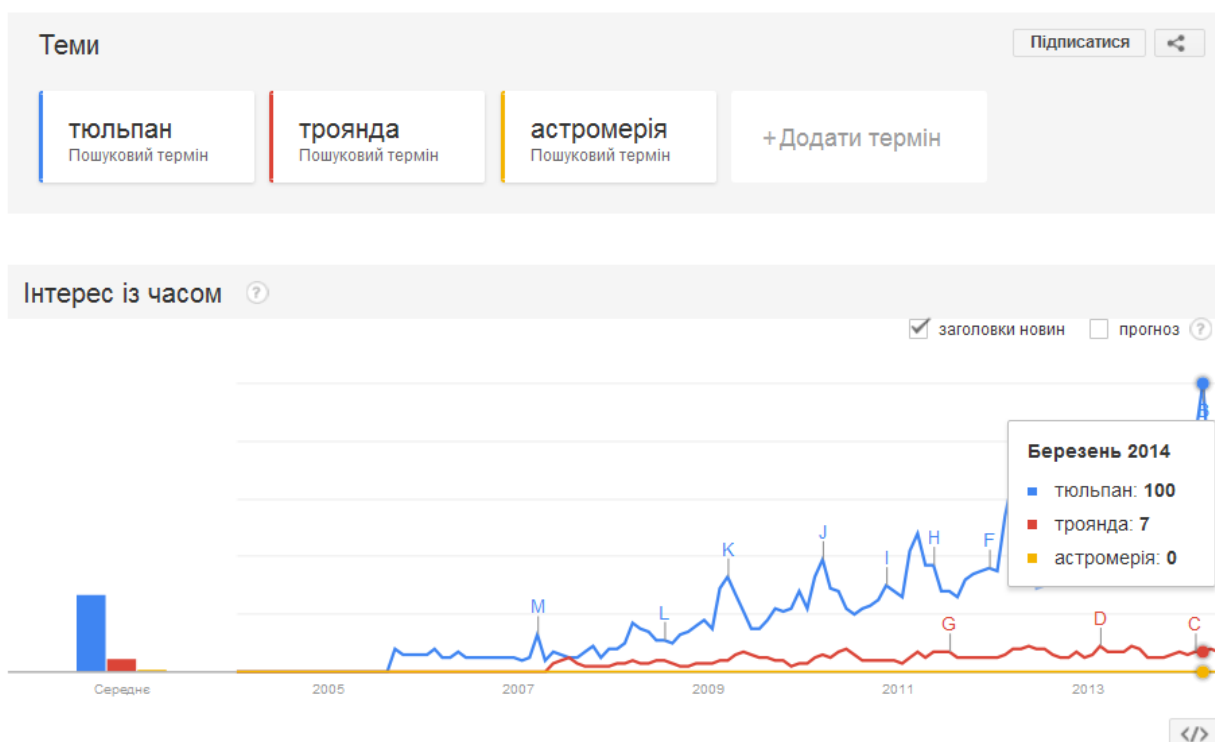


Рис.11. Порівняння частоти запитів

У Google Trends є різні фільтри, за якими можна відсіювати результати. Можна переглядати тренди по новинах, зображеннях, покупках, в різних регіонах, у різних сферах або у всіх відразу тощо (Рис.12).

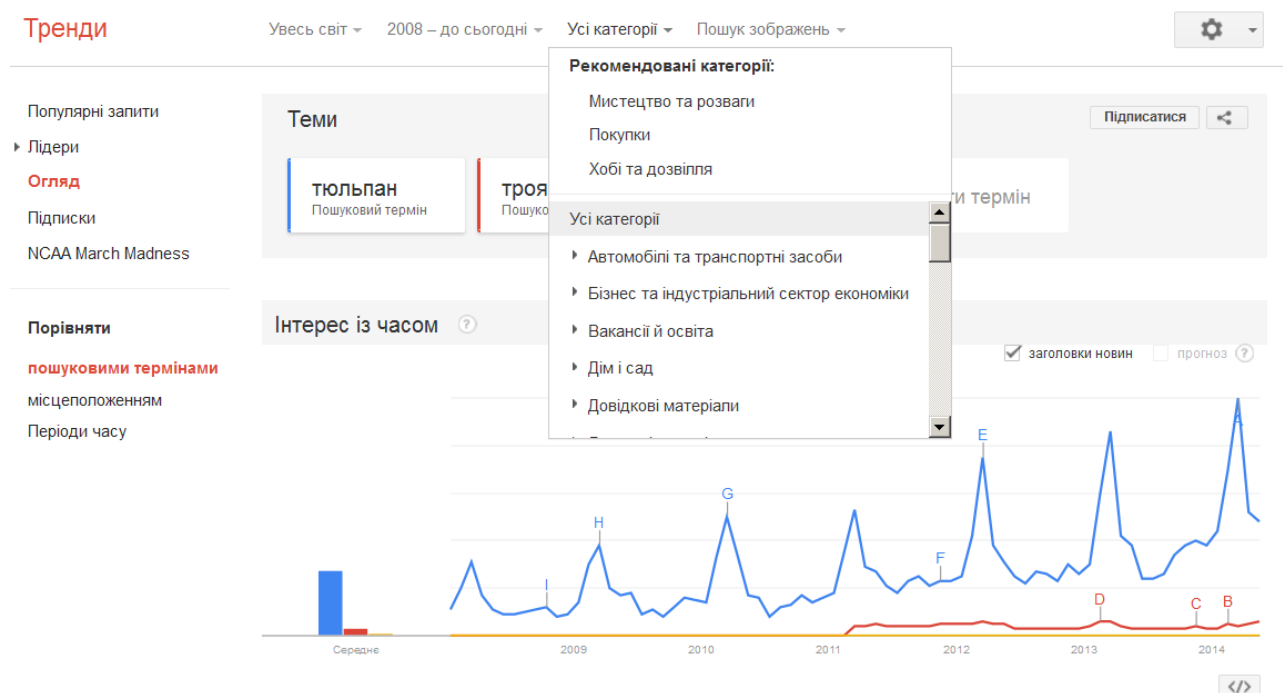


Рис.12. Використання фільтрів

Отже, Google Trends можна використовувати як для аналізу, так і для підбору ключових слів. Це відмінний інструмент, який дає потрібну інформацію щодо тенденцій пошуку в Google.

Хід роботи:

1. Завантажте інсталяційний пакет торрент-клієнта BitTorrent з офіційного сайту (<http://www.bittorrent.com/>), вибравши безкоштовну програму (Free Download).
2. Встановіть BitTorrent на свій ПК. Встановлення всіх додатків, які пропонує інсталяційний пакет, не є обов'язковим.
3. Після встановлення, відкрийте головне вікно програми. Ознайомтеся з усіма можливостями, які пропонує основне вікно програми.
4. Зареєструйтеся на торрент-трекері Гуртом за адресою <http://toloka.hurtom.com>.
5. Ознайомтеся з темами, які є на цьому торрент-трекері. Виберіть один з розділів, опишіть його структуру. Зробіть відповідну копію з екрану.
6. В одній із тематик виберіть файл, який необхідно завантажити.
7. Опишіть структуру опису файлу – яку інформацію можна отримати з опису. Зробіть відповідну копію з екрану.
8. Використайте лінк-магніт для початку завантаження файлу, простежте, скільки пірів беруть участь у роздачі та яка швидкість завантаження файлу. Зробіть відповідну копію з екрану.
9. Залежно від порядкового номера студента у списку підгрупи, виконайте пошук зазначеної книги із списку, а тоді кроки 6-7.

№	Файл для опрацювання
1.	Червона книга України (2009)
2.	Самовчитель польської мови (2006)
3.	Закарпатські народні страви (1990)
4.	Міфи давньої Греції (1991)
5.	З історії львівських вулиць (1990)
6.	Етикет і культура спілкування (2001)
7.	Замки Західної України (2004)
8.	Фразеологія сучасної української мови (2005)
9.	Книга “Дива України” (2009)
10.	Словник давньоукраїнської міфології (1993)
11.	Великодній сувенір. Українська народна писанка (2005)
12.	Італійська кухня (2013)
13.	Антисуржик (1994)
14.	Книга рекордів України (2000-2002)

15.	20 кращих екскурсій Україною (2009)
-----	-------------------------------------

10. Відкрийте сервіс Google Trends за адресою: <http://trends.google.com/trends/explore>. Ознайомтеся з його можливостями.
11. Доберіть три пари ключових слів для порівняння зміни їхньої популярності. Проаналізуйте інтерес до цих ключових слів у різних регіонах України. Перегляньте пов'язані пошукові запити. За результатами досліджень, заповніть подану таблицю:

№	Термін	Період найвищої зацікавленості	Період найнижчої зацікавленості	Зацікавленість по регіонах/містах	Пов'язані пошукові запити	Пояснення щодо коливань у тенденціях

12. Оформіть звіт.