

МІНІСТЕРСТВО ВНУТРІШНІХ СПРАВ УКРАЇНИ

Харківський національний університет внутрішніх справ

факультет № 4

кафедра кібербезпеки та DATA-технологій

МЕТОДИЧНІ МАТЕРІАЛИ

до практичних занять

з навчальної дисципліни

Розвідувально-аналітична робота у
кіберсфері

обов'язковий компонент освітньої програми другого рівня вищої освіти
125 Кібербезпека (безпека інформаційних та комунікаційних систем)

м. Харків
2023 рік

ЗАТВЕРДЖЕНО

Науково-методичною радою
Харківського національного
університету внутрішніх справ
Протокол від 30.08.2023 № 7

СХВАЛЕНО

Вченою радою факультету № 6
Протокол від 25.08.2023 № 7

ПОГОДЖЕНО

Секцією Науково-методичної ради
ХНУВС з технічних дисциплін
Протокол від 29.08.2023 № 7

Розглянуто на засіданні кафедри кібербезпеки та DATA-технологій (протокол від 15.08.2023 № 8)

Розробник:

Доцент кафедри кібербезпеки та DATA-технологій, к.ю.н., професор Манжай О.В.

Рецензенти:

Тулупов В.В., доцент кафедри кібербезпеки та DATA-технологій факультету № 6
Харківського національного університету внутрішніх справ к.т.н., доцент;

Павликівський В.І., перший проректор Харківського університету, д.ю.н., професор

1. Розподіл часу навчальної дисципліни за темами

Денна форма навчання

Номер та назва навчальної теми	Кількість годин, відведених на вивчення навчальної дисципліни						Вид контролю
	Всього	з них:					
		лекції	Семінарські заняття	Практичні заняття	Лабораторні заняття	Самостійна робота	
Семестр № 1							
Тема № 1 Теоретичні засади розвідувально-аналітичної роботи	38	6	2			30	Екзамен
Тема № 2 Методологія здійснення розвідувально-аналітичної роботи	46	8	4	4		30	
Тема № 3 Окремі інструменти накопичення та аналізу розвідувальних відомостей	20	6	6	8		0	
Тема № 4 Розвідувально-аналітична робота щодо груп злочинів	46	8	4	4		30	
Всього за семестр:	150	28	16	16		90	

Заочна форма навчання

Номер та назва навчальної теми	Кількість годин, відведених на вивчення навчальної дисципліни						Вид контролю
	Всього	з них:					
		лекції	Семінарські заняття	Практичні заняття	Лабораторні заняття	Самостійна робота	
Семестр № 1							
Тема № 1 Теоретичні засади розвідувально-аналітичної роботи	46	2	2			42	Екзамен
Тема № 2 Методологія здійснення розвідувально-аналітичної роботи	44	2	2			40	
Тема № 3 Окремі інструменти накопичення та аналізу розвідувальних відомостей	6	2	2	2		0	
Тема № 4 Розвідувально-аналітична робота щодо груп злочинів	44	2		2		40	
Всього за семестр:	150	8	6	4		132	

2. Методичні вказівки до практичного навчання

Практичне заняття. Упорядкування великих даних

Навчальна мета заняття: отримати навички упорядкування великих даних з використанням спеціалізованого програмного забезпечення та здійснення пошуку відповідної інформації серед таких даних.

Час проведення 4 год. Місце проведення: комп'ютерний клас.
(кількість годин) (полігон, комп'ютерний клас, лабораторія, лінгфонний кабінет)

Устаткування: персональний комп'ютер зі встановленою операційною системою Windows 2007 або вище; MS Access; MS Excel; застосунки СУБД; EmEditor; TextPipe, WindowsGrep.

Завдання, які потрібно виконати, підкреслено

В роботі правоохоронних органів нерідко доводиться мати справу з великими об'ємами даних, які не упорядковані належним чином. Ці дані можуть містити корисну інформацію, проте необхідність використання великої кількості застосунків для їх обробки та тривалий час самої обробки даних значно уповільнюють оперативно-службову діяльність. Враховуючи наведене, на декількох прикладах розглянемо, яким чином можна упорядкувати відповідні дані та як правильно організувати ефективний пошук.

Слід зазначити, що в органах поліції традиційно використовується велика кількість банків даних, створених під систему Cronos. Зважаючи на це, вбачаємо доцільним навести відповідні приклади у розрізі роботи даної системи.

Перегляд змісту великих текстових файлів

Якщо великі дані зберігаються у текстовому вигляді, то переглянути їх за допомогою неспеціалізованих програмних засобів є достатньо складним завданням. Алгоритм роботи стандартних засобів перегляду передбачає першочергове завантаження всього обсягу файлу до оперативної пам'яті. Якщо такий файл має об'єм декілька гігабайт, то його відкриття триватиме довго, тому з метою перегляду змісту таких документів слід користуватися спеціалізованими програмами. Однією з таких програм є редактор EmEditor. За його допомогою досить зручно переглядати великі текстові документи, здійснювати в них пошук, розділяти їх на частини, вносити інші зміни. У разі потреби перетворення текстових файлів у формат бази даних, може знадобитися їх попередня обробка для приведення до певної форми. В цьому випадку спеціалізовані редактори можуть бути використані для швидкого перегляду файлу та вилучення з нього фрагменту даних для відпрацювання процесу перетворення (рис. 1).



Рис. 1. Результат вилучення фрагменту даних

У подальшому вилучений фрагмент тексту може буде використаний для накладання відповідних фільтрів.

Приведення даних до потрібної форми

Для імпорту текстових даних до якоїсь СУБД вони нерідко мають бути перетворені у певну форму, вимоги до якої визначаються алгоритмом роботи СУБД. З метою швидкого внесення відповідних змін можуть бути застосовані спеціалізовані інструменти, як от TextPipe.

Порядок роботи з вказаною програмою є достатньо простим. У лівому полі обирається відповідний фільтр, який налаштовується, а потім розміщується в тому порядку, в якому його слід застосувати до відповідного файлу. Для ефективного створення фільтрів потрібно знати головні шаблони для перетворень. З відповідними прикладами можна ознайомитися, наприклад, за адресою datamystic.com/textpipe/manual/general_usage_easypatterns_reference.htm.

По суті створення фільтру нагадує процес написання простої програми (рис. 2).

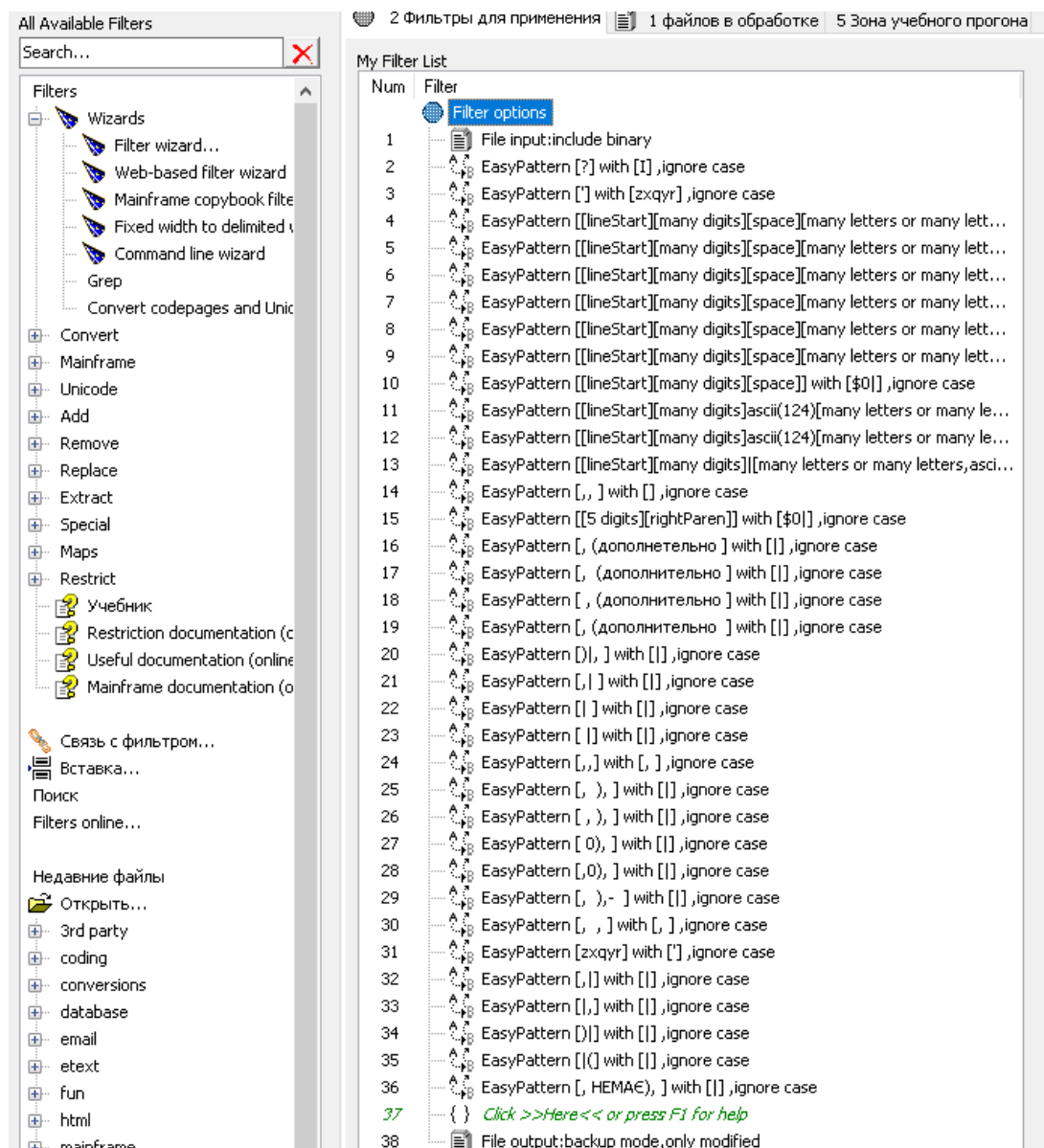


Рис. 2. Приклад фільтру

У програмі TextPipe відповідні фільтри поділено на категорії за призначенням, що значно спрощує процес знаходження потрібного елементу.

Після приведення до належного вигляду текстовий документ може бути імпортовано до СУБД. Це може бути зроблено декількома шляхами. Наприклад:

- 1) через вбудовану функцію імпорту з текстового файлу;
- 2) через попередній імпорт текстових документів до іншої СУБД (наприклад, MS Access);
- 3) з використанням таблиць відповідності.

Перший варіант є найбільш застосовним для імпорту невеликих текстових масивів, другий – для імпорту середніх за обсягом даних, третій – для імпорту великих текстових масивів (наприклад, декілька десятків гігабайт).

Існують випадки, коли імпортувати великі дані до СУБД є недоцільним, через суттєве зростання об'єму вихідних файлів банку. В такому випадку пошук можна здійснювати стандартними засобами або з використанням спеціалізованих утиліт.

В системі Windows, наприклад, для цього можна скористатися утилітою findstr з вказівкою потрібних параметрів. Наприклад,

```
findstr /s "що шукаємо" де_шукаємо
findstr /B "t.....31@yahoo.fr" rez_out.txt
```

Крім того, з цією метою можна використовувати утиліти Grep, Folder Find Text, DocFetcher.

Завдання

За завданням викладача:

- 1) привести фрагмент тексту до визначеної структури;
- 2) трьома способами імпортувати приведені дані до СУБД;
- 3) реалізувати в СУБД глобальний пошук у декількох базах даних;
- 4) здійснити пошук строки у текстовому файлі за допомогою декількох утиліт;
- 5) скласти звіт.

Практичне заняття. Перетворювачі Maltego

Навчальна мета заняття: навчитися додавати перетворювачі в Maltego декількома способами.

Час проведення 2 год. Місце проведення: комп'ютерний клас.
(кількість годин) (полігон, комп'ютерний клас, лабораторія, лінгфонний кабінет)

Одним з корисних рішень для системи Maltego є інтеграція пошукових можливостей системи Google для вирішення завдань OSINT в соціальних мережах. Для цього може бути застосований перетворювач Google Serches, робота з яким детально описана за адресою osintops.com/how-to-use-google-for-osint-on-maltego.

Для того щоб встановити цей перетворювач, потрібно перейти у домашнє полотно (Home) програми Maltego та у розділі Internal Hub Items натиснути кнопку «плюс» (рис. 1).

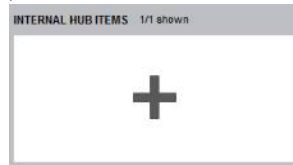


Рис. 1. Кнопка додавання перетворювача

У вікні, що з'явилося, потрібно додати Seed <https://cetas.paterva.com/TDS/runner/showseed/fastCSEs>, для чого ввести дані як на рис. 2.

ID	<input type="text" value="googleSearches"/>	
	The unique name of this Transform Seed	
Name	<input type="text" value="Google Searches"/>	
	The display name of this Transform Seed	
Seed URL	<input type="text" value="https://cetas.paterva.com/TDS/runner/showseed/fastCSEs"/>	<input type="button" value="Certificate"/>
	The URL of the Transform Seed from where to install the transforms	
Icon URL	<input type="text"/>	<input type="button" value="Browse"/>
	The web or file URL to a 48x48 icon	
Description	<input type="text"/>	
	A short description	
Details	<input type="text"/>	
	A long description	

Рис. 2. Вікно введення параметрів перетворювача

Після натискання кнопки «ОК» на домашньому полотні з'явиться новий перетворювач (рис. 3).

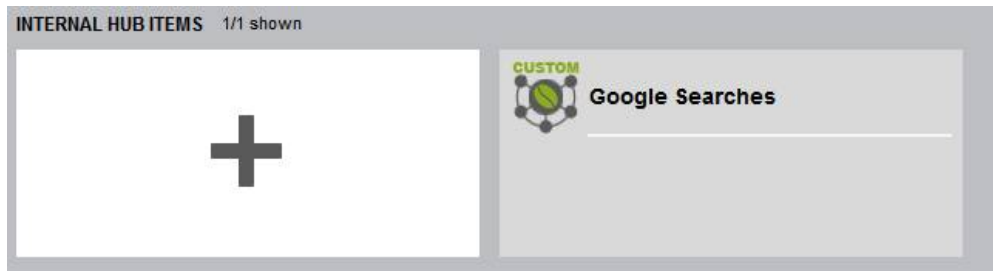


Рис. 3. Доданий перетворювач

Після наведення мишею на доданий перетворювач, потрібно його встановити натисканням «Install».

Після встановлення перетворювача можна приступати до його використання для вирішення завдань пошуку інформації. Для цього потрібно створити нове полотно (Create a new graph) та додати на нього сутність Фраза (Phrase). У текстове поле нової сутності слід записати фразу для пошуку, як зображено на рис. 4.

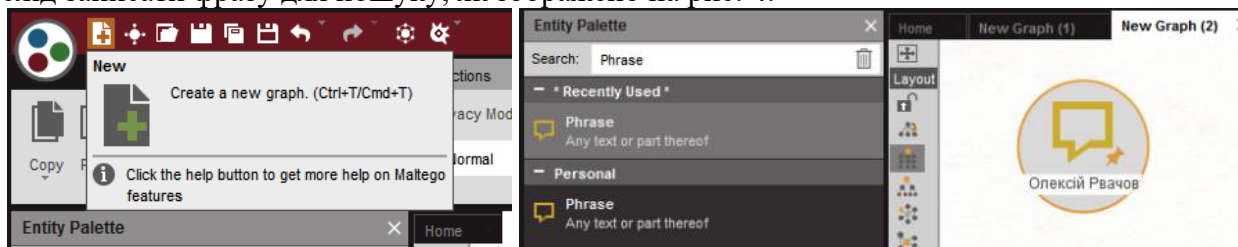


Рис. 4. Додавання сутності на новому полотні

Після виконання описаних дій при натисканні правою кнопкою миші на новоутворену сутність можна обрати перетворювач Google Serches та здійснити потрібний вид пошуку (рис. 5).

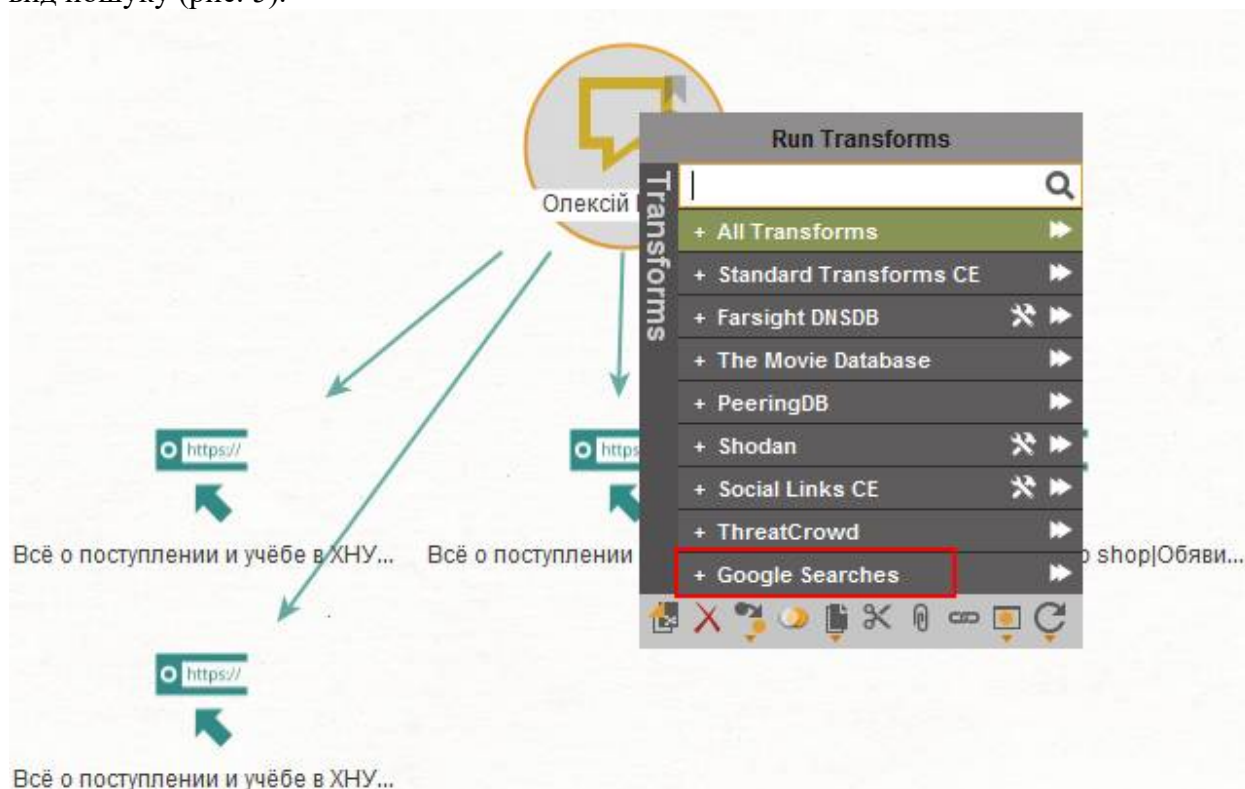


Рис. 5. Виконання пошуку

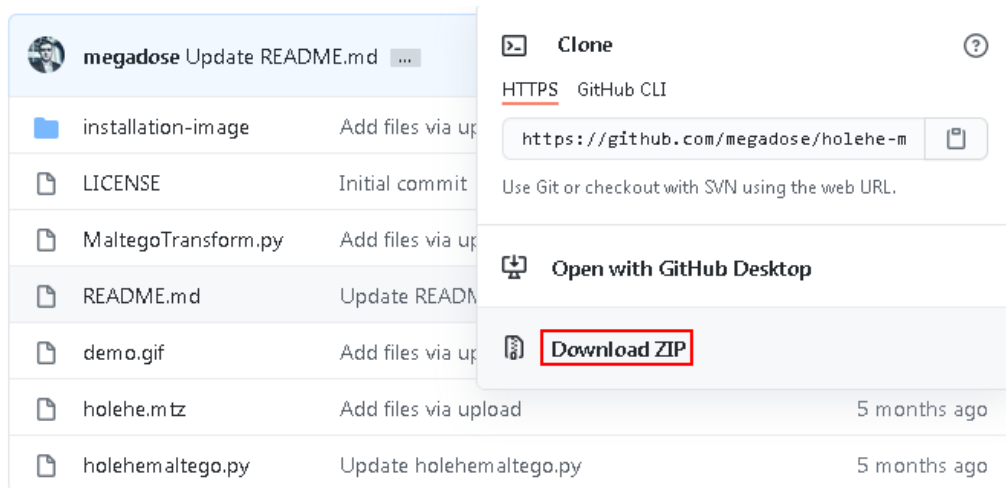
На даний момент перетворювач дозволяє здійснювати пошук у соціальних мережах: 4Chan, 8Chan, Dailymotion, Facebook, Gab, Instagram, LinkedIn, Reddit, Telegram, Twitter, VKontakte, YouTube. Для того, щоб здійснити якісний пошук потрібно заповнювати текстове поле сутності Фраза запитом з використанням мови Google (Google Dorks).

Завдання: спробуйте знайти згадування в соціальних мережах за назвою облікових записів або іменем і прізвищем. Самостійно опрацюйте перетворювачі з ресурсу osint.party

Додавання сутності та перетворювача в Maltego, написаного з використанням Python

Існує велика кількість рішень для системи Maltego, створених ентузіастами та доступних у відкритому доступі. Одним із таких рішень є перетворювач Holehe (github.com/megadose/holehe-maltego), який дозволяє за наявною назвою електронної пошти знаходити в мережі інформацію про те, на яких ресурсах вона використовувалася для реєстрації. Відповідно можна дізнатися, які облікові записи має користувач, якому належить така електронна пошта.

Для встановлення Holehe в Maltego потрібно спершу імпортувати відповідну сутність. Для цього в системі Linux у терміналі слід виконати команду `git clone github.com/megadose/holehe-maltego`, а в системі Windows – на сторінці github.com/megadose/holehe-maltego/tree натиснути кнопку Code та завантажити архів як на рис. 6.



Після завантаження та розпакування архіву потрібно додати сутність в систему (рис. 7).

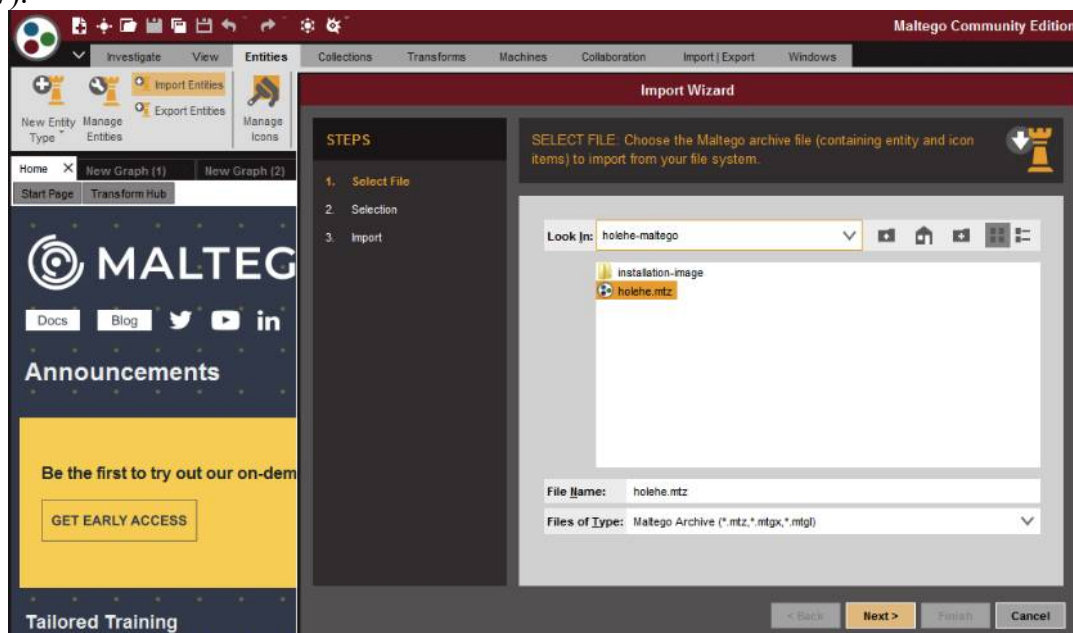


Рис. 7. Імпорт сутності Holehe

Тепер слід інсталиувати саму утиліту Holehe. Для виконання цього завдання в системі має бути встановлено Python. Для встановлення Holehe, необхідно в терміналі виконати команду `pip3 install holehe` (рис. 8).

```
C:\Users\User>pip3 install holehe
Collecting holehe
  Downloading https://files.pythonhosted.org/packages/a5/b2/c5d84582147f0d375e54082076729293998ac1686c15a7c56a210ed2731b/holehe-1.58.4.2-py3-none-any.whl (120kB)
100% |#####| 122kB 1.8MB/s
Collecting termcolor (from holehe)
  Downloading https://files.pythonhosted.org/packages/8a/48/a76be51647d0eb9f10e2a4511bf3ffb8cc1e6b14e9e4fab46173aa79f981/termcolor-1.1.0.tar.gz
Collecting trio (from holehe)
  Downloading https://files.pythonhosted.org/packages/5e/ce/1a6e875838058e9df989247ee339daa3d79cec599182a1a836ee1aa74750/trio-0.18.0-py3-none-any.whl (354kB)
100% |#####| 358kB 4.1MB/s
Collecting httpx (from holehe)
  Downloading https://files.pythonhosted.org/packages/2d/c6/59aa4188e7eddb9e89ec67a51598ca6bfc09f1b38c9b45f7ee45af7a4df4/httpx-0.16.1-py3-none-any.whl (65kB)
100% |#####| 71kB 5.1MB/s
Collecting bs4 (from holehe)
```

Рис. 8. Встановлення утиліти Holehe

У разі успішного встановлення утиліти Holehe, потрібно приступити до додавання відповідного перетворювача в Maltego (рис. 9).

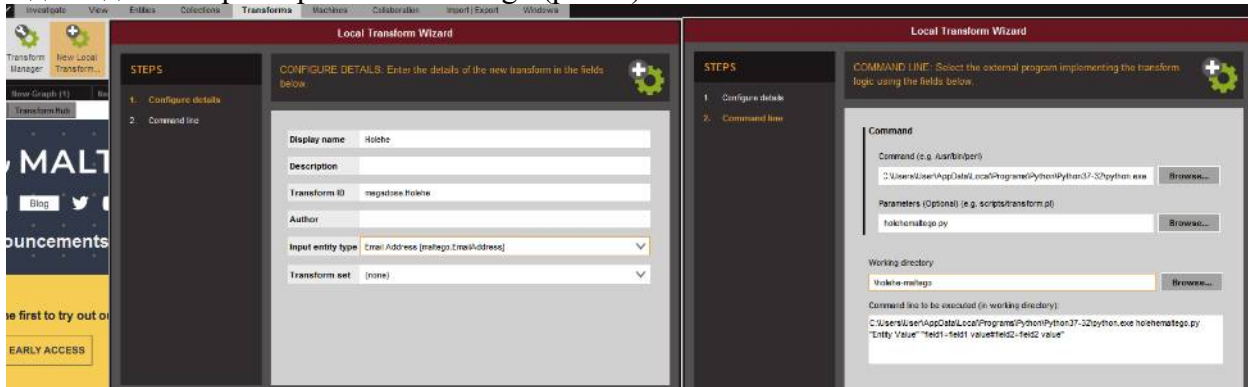


Рис. 9. Додавання перетворювача Holehe

Якщо перетворювач додано успішно, то його можна використовувати для встановлення факту реєстрації на різних ресурсах за наявними назвами електронної пошти (рис. 10).

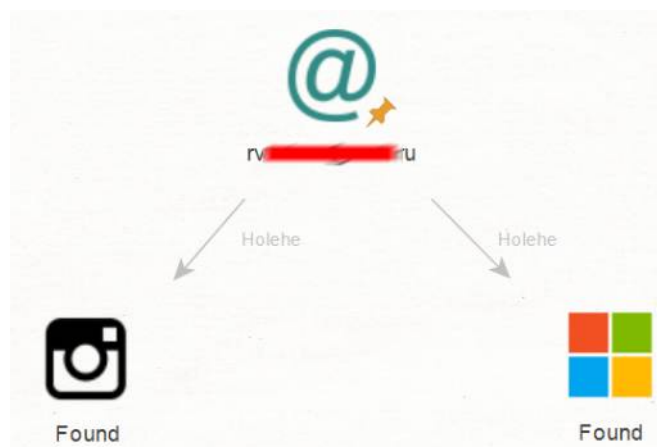


Рис. 10. Результат роботи перетворювача Holehe

Аналогічним чином можна встановлювати в системі й інші локальні перетворювачі.

Потрібно пам'ятати, що в окремих випадках для коректної роботи перетворювачів може знадобитися додаткове встановлення бібліотек Python. Так, наприклад, у випадку з

локальним перетворювачем для визначення назви облікового запису Skype за іменем (github.com/megadose/cqfd-maltego) може знадобитися додаткова інсталяція бібліотек cqfd та lxml (рис. 11).

```
c:\Program Files\Python38>pip install cqfd
Collecting cqfd
  Downloading cqfd-1.11-py3-none-any.whl (2.0 kB)
Collecting fake-useragent
  Downloading fake-useragent-0.1.11.tar.gz (13 kB)
Collecting argparse
  Downloading argparse-1.4.0-py2.py3-none-any.whl (23 kB)
Using legacy 'setup.py install' for fake-useragent, since package 'wheel' is not installed.
Installing collected packages: fake-useragent, argparse, cqfd
  Running setup.py install for fake-useragent ... done
Successfully installed argparse-1.4.0 cqfd-1.11 fake-useragent-0.1.11
WARNING: You are using pip version 20.2.3; however, version 20.3.3 is available.
You should consider upgrading via the 'c:\program files\python38\python.exe -m pip install --upgrade pip' command.

c:\Program Files\Python38>pip install lxml
Collecting lxml
  Downloading lxml-4.6.2-cp38-cp38-win_amd64.whl (3.5 MB)
    | 3.5 MB 2.2 MB/s
Installing collected packages: lxml
Successfully installed lxml-4.6.2
```

Рис. 11. Встановлення додаткових бібліотек

Як і у випадку зі встановленням Holec для додавання локального перетворювача Get Skype Account From Name потрібно натиснути кнопку New Local Transform... та ввести відповідні параметри (рис. 12).

Local Transform Wizard

STEPS

1. Configure details
2. Command line

CONFIGURE DETAILS: Enter the details of the new transform in the fields below.

Display name: Get Skype Account(cqfd)

Description: Get Skype Account From Name

Transform ID: om.GetSkypeAccount(cqfd)

Author: megadose

Input entity type: Person [maltego Person]

Transform set: (none)

Local Transform Wizard

STEPS

1. Configure details
2. Command line

COMMAND LINE: Select the external program implementing the transform logic using the fields below.

Command

Command (e.g. /usr/bin/perl): C:\Program Files\Python38\python.exe

Parameters (Optional) (e.g. scripts/transform.pl): cqfd-maltego.py

Working directory: D:\cqfd-maltego

Command line to be executed (in working directory): C:\Program Files\Python38\python.exe cqfd-maltego.py "Entity Value" "field1=field1 value#field2=field2 value"

Рис. 12. Додавання перетворювача Get Skype Account From Name

У випадку успішного встановлення за допомогою цього перетворювача можна здійснювати відповідний пошук як на рис. 13.

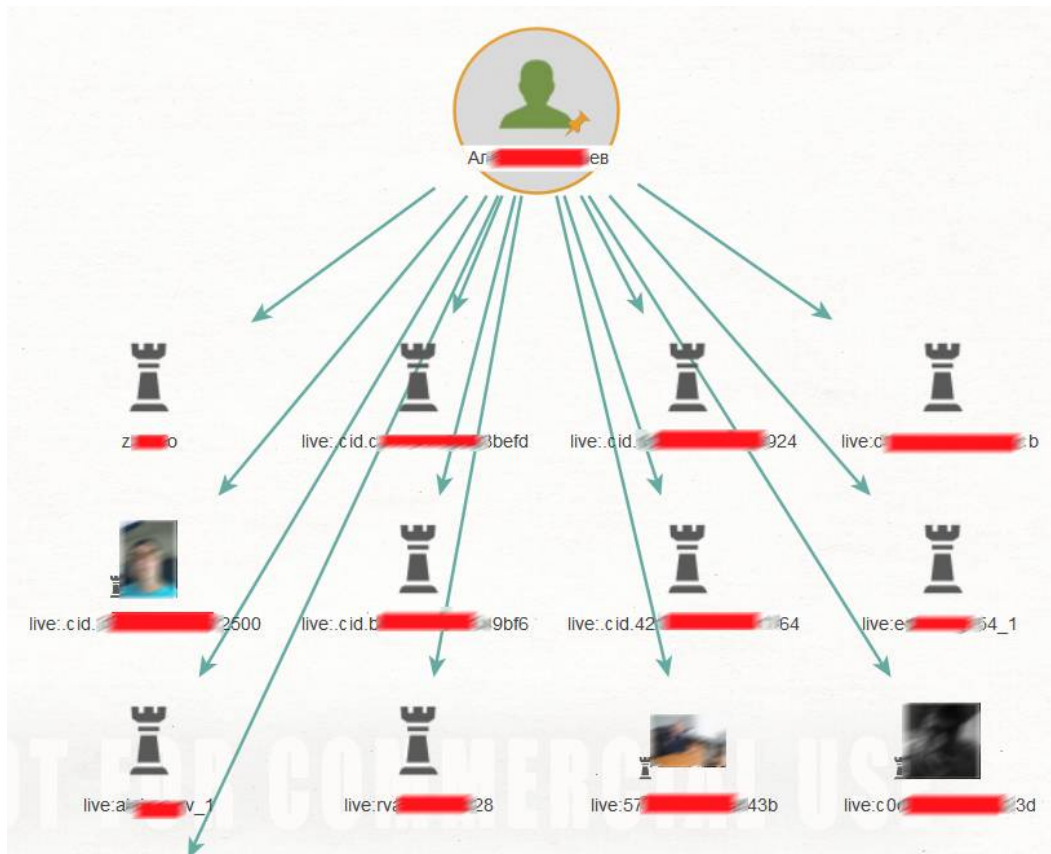


Рис. 13. Результат роботи перетворювача Get Skype Account From Name

Завдання: спробуйте самостійно додати один із локальних перетворювачів, наведених за адресами: github.com/HowToFind-bot/osint-tools, github.com/IntelligenceX/SDK/tree/master/Maltego%20Transform та протестувати його функціональність.

Практичне заняття. Автоматизація розрахунку матриці Sleipnir

Навчальна мета заняття: відпрацювати навички побудови системи оцінки ризиків.

Час проведення 2 год. Місце проведення: комп'ютерний клас.
(кількість годин) (полігон, комп'ютерний клас, лабораторія, лінгафонний кабінет)

Для автоматизації розрахунку матриці Sleipnir може бути застосовано програмне забезпечення Excel.

Спершу потрібно сформувати відповідну текстову частину таблиці та додати до неї кнопку через меню Розробник → Вставити (рис. 1).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1																				
2																				
3																				
4																				
5																				
6																				
7																				
8																				
9																				
10																				
11																				
12																				
13																				

Рис. 1. Підготовлена частина матриці

Після вставки кнопки, яку для прикладу назвемо «Розрахунок», відкриється середовище Visual Basic, де можна написати такий код:

```
Sub Розрахунок()
Dim InRange As Range
Dim R As Range
Set InRange = Selection
For Each R In InRange.Cells
R.Offset(0, 20) = R.Interior.Color
Next R
End Sub
```

Дана функція призначена для визначення числового коду кольору виділених комірок. Значення кодів виводитимуться для кожної комірки відповідно із зміщенням на 20 стовпців праворуч.

Тепер, коли відомі коди відповідних кольорів, їх можна визначити для кожної комірки (у комірках W2:AA2 введемо коди кольорів, що будемо використовувати, як на рис. 2).

	W	X	Y	Z	AA
1					
2	255	3243501	65535	5287936	15773696

Рис. 2. Коди кольорів, застосовуваних у підсумковій матриці

Також потрібно розрахувати вагові коефіцієнти. З цією метою у комірці C4 введемо формулу:

=ЕСЛИ(W4=\$W\$2;C\$2*\$P\$2;ЕСЛИ(W4=\$X\$2;C\$2*\$Q\$2;ЕСЛИ(W4=\$Y\$2;C\$2*\$R\$2;EC

ЛИ(W4=\$Z\$2;C\$2*\$SS\$2;ЕСЛИ(W4=\$AA\$2;C\$2*\$T\$2;0))))))

За допомогою цієї формули визначатимуться вагові коефіцієнти залежно від кольору комірки. Значення кожного коефіцієнту вводиться безпосередньо в кольорову комірку. Враховуючи це, можна підрахувати їх суми з використанням функції СУММ у стовпці В. Після проведення усіх розрахунків слід відсортувати рядки, які характеризують злочинні угруповання за ступенем загрози.

Підсумковий алгоритм роботи із створеним документом Excel можна представити так:

1. Створіть потрібну кількість рядків (один рядок відповідатиме одному організованому злочинному угрупованню) або видаліть непотрібні.
2. Зробіть копію формул з відповідних стовпців В:N у новостворені рядки.
3. Позначте кожен комірку у новостворених рядках кольором згідно зі ступенем загрози. Використовуйте кольори з комірок Р2:Т2.
4. Виділіть усі кольорові комірки, які характеризують організовані злочинні угруповання та натисніть кнопку "Розрахунок".
5. Відсортуйте рядки по стовпцю В, для чого виділіть відповідну кількість рядків з назвами злочинних угруповань та за допомогою Дані → Сортвання виконайте відповідний тип сортування (рис. 3).

Ідентифікатор групи, З - загальнонаціональна, Р - регіональна, М - місцева	Показник загрози	Критерії											Високий	Середній	Низький	Нульовий	Невідомий	
		12	11	10	9	8	7	6	5	4	3	2	1	4	2	1	0	2
		Корупція	Насилля	Впровадження	Відмивання коштів	Взаємодія	Відособлення	Монополізм	Масштаб	Використання розвідувальних відомостей	Диверсифікація	Дисциплінованість	Згуртованість	1. Створіть потрібну кількість рядків (один рядок відповідатиме одному ОЗУ) або видаліть непотрібні. 2. Зробіть копію формул з відповідних стовпців В:N у новостворені рядки. 3. Позначте кожен комірку у новостворених рядках кольором згідно зі ступенем загрози. 4. Виділіть усі кольорові комірки, які характеризують ОЗУ та натисніть кнопку "Розрахунок". 5. Відсортуйте рядки по стовпцю В, для чого виділіть відповідну кількість рядків з назвами ОЗУ та за допомогою Дані - Сортвання виконайте відповідний тип сортування.				
З-"Люті"	235	48	44	20	36	32	28	0	10	8	3	4	2					
З-"Старі"	192	24	22	40	9	16	14	24	10	16	12	4	1					
З-"Молоді"	188	12	44	20	18	32	28	6	10	0	6	8	4					
Р-"Скіфи"	152	24	22	10	18	16	14	12	10	8	12	4	2					
Р-"АРО"	141	24	44	0	18	32	0	6	5	0	0	8	4					
Р-"Резо"	139	24	22	0	9	16	28	12	0	4	12	8	4					
М-"Турист"	139	12	11	40	18	0	0	24	10	8	6	8	2					
М-"Іпмаш"	116	24	44	10	9	0	0	6	5	0	6	8	4					
М-"Більки"	76	12	22	0	0	16	14	6	0	0	3	2	1					
														Розрахунок				

Рис. 3. Результат автоматизації методу Sleipnir

Побудовану матрицю можна використовувати для планування роботи правоохоронного органу щодо протидії, як окремим організованим злочинним угрупованням, так і організованій злочинності в цілому.

Завдання: автоматизувати створення матриці Sleipnir з використанням можливостей табличного процесора.

Практичне заняття. Особливості накопичення та мережного аналізу електронних даних

Навчальна мета заняття: вивчити техніко-методологічні особливості виконання завдань накопичення та мережного аналізу електронних даних.

Час проведення 4 год. Місце проведення: комп'ютерний клас.

(кількість годин)

(полігон, комп'ютерний клас, лабораторія, лінгафонний кабінет)

Устаткування: персональний комп'ютер зі встановленою операційною системою Windows 7 або вище (Linux).

1. Накопичення та обробка даних з використанням програми Hunchly

Командний підхід в розвідувально-аналітичній роботі вимагає вміння автоматизувати процес накопичення та обробки даних та здійснення взаємного обміну відповідними відомостями з колегами та керівництвом.

Для вирішення цього завдання можуть застосовуватися спеціалізовані рішення, як от, наприклад, Hunchly (hunch.ly) та Kuiper (github.com/DFIRKuiper/Kuiper).

Для роботи з інструментом Hunchly потрібно встановити браузер Chrome та безпосередньо саму програму за адресою hunch.ly/downloads. Сам застосунок є платним, проте є можливість скористатися ним у пробний період після реєстрації. Перевагою Hunchly є те, що існують версії програми для різних операційних систем, у тому числі мобільних.

Після інсталяції версії для ОС Windows буде автоматично встановлено плагін в браузер Chrome, який виконуватиме роль комунікатора між браузером та головною частиною Hunchly.

Порядок роботи з програмою є інтуїтивно зрозумілим, а інтерфейс нескладним. При увімкненні плагіна Hunchly у браузері всі відкриті сторінки автоматично структуруватимуться та зберігатимуться в окремих справах, що призначаються користувачем-дослідником (рис. 1).



Рис. 1. Налаштування плагіна Hunchly

У самому ж плагіні можна викликати головну панель (Dashboard) та виконати відповідні налаштування (рис. 2).

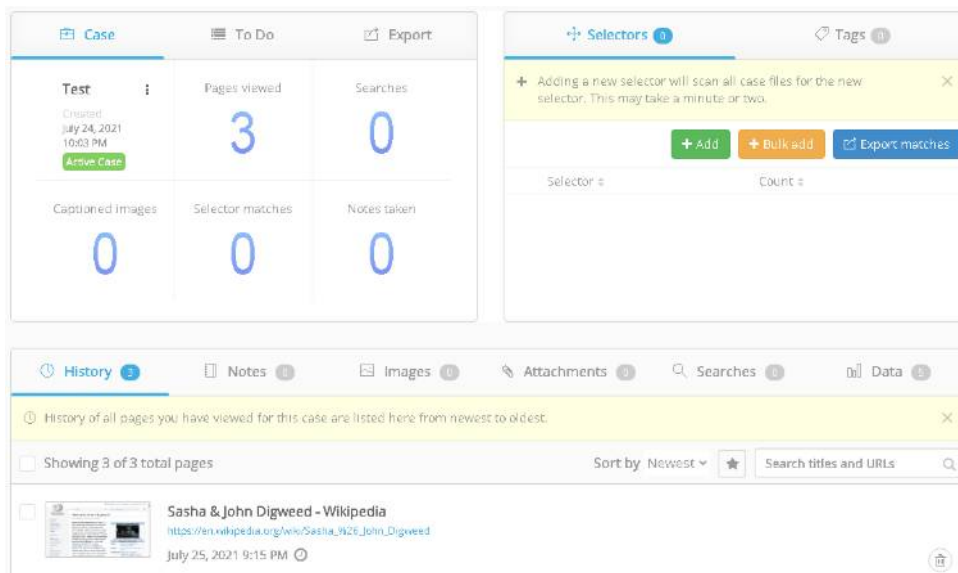


Рис. 2. Головна панель Hunchly

Програма Hunchly дозволяє вирішувати велику кількість різноманітних завдань аналітика, у тому числі синхронізувати роботу по накопиченню відомостей з колегами через Google-таблиці. Детально цей процес описано Дж. Сейтсом за посиланням threadreaderapp.com/thread/1398343406875598852.html, або twitter.com/jms_dot_py/status/1398343406875598852.

По суті синхронізація реалізується шляхом створення Google-таблиці з певними полями та аркушами (рис. 3).

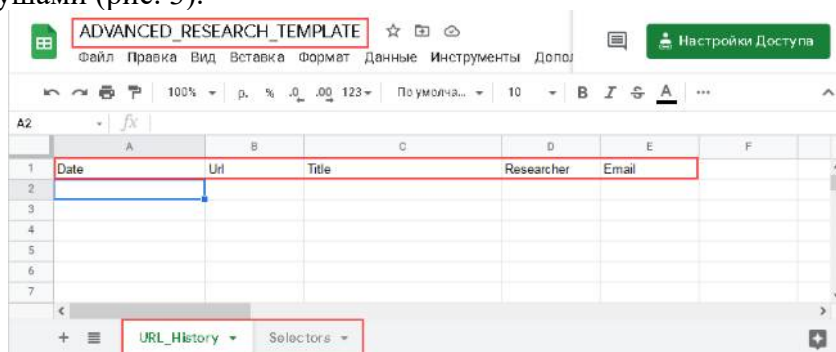


Рис. 3. Форма таблиці-шаблону

Після чого викликається редактор скриптів (рис. 4) та пишеться код, який синхронізуватиме роботу Hunchly зі створеною Google-таблицею (рис. 5).

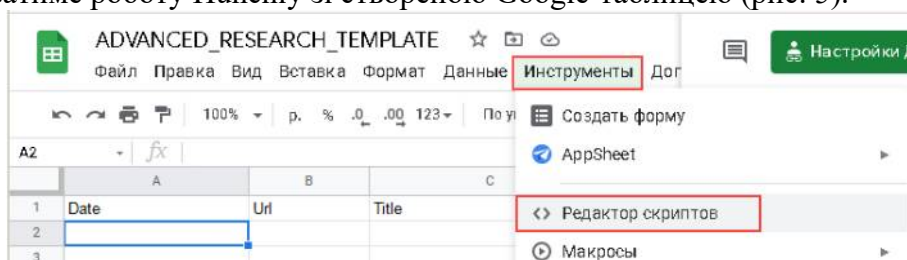


Рис. 4. Виклик редактора скриптів

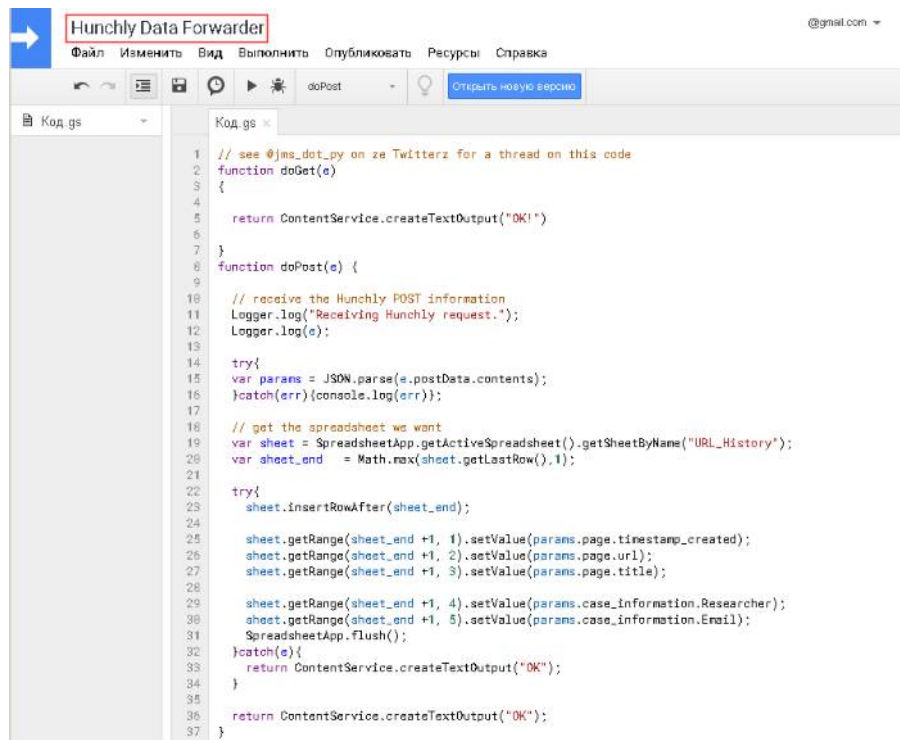



Рис. 5. Створення скрипта-обміну

При створенні коду для уникнення проблем із записом відомостей до таблиці бажано використовувати стару версію редактора, а також у меню «Виконати» потрібно вимкнути середовище виконання Apps Script (технології Chrome v8).


Після збереження скрипта  потрібно його опублікувати як web-застосунок, погодити надання відповідних привілеїв та розподілити його для всіх (рис. 6).

Deploy as web app

Current web app URL: <https://script.google.com/macros/s/AKfycbz23L1wD0pN1> [Disable web app](#)

Test web app for your latest code.

Project version:
10

Execute the app as:
Me ( @gmail.com)

You need to authorize the script before distributing the URL.

Who has access to the app:
Anyone, even anonymous

[Обновить](#) [Отмена](#) [Справка](#)

Рис. 6. Створення скрипта-обміну

Одержане посилання на скрипт потрібно внести до відповідних параметрів Головної панелі Hunchly, а також налаштувати параметри ідентифікації особи, яка внесла відповідний запис до таблиці (рис. 7).

Case Information

Data Forwarding

Enable ☒

HTTP(S) Address

Test connection

Case reference number

Email

Researcher

Рис. 7. Налаштування пересилання даних у Hunchly

Якщо усе виконано вірно, то після активації плагіна Hunchly та перегляду відповідної сторінки в Інтернеті, вона фіксуватиметься у Google-таблиці як на рис. 8.

ADVANCED_RESEARCH_TEMPLATE

Файл Правка Вид Вставка Формат Данные Инструменты Доп

100% р. % 0.00 123 По умолча... 10 B I S A

	A	B	C	D	E	F
1	Date	Url	Title	Researcher	Email	
2	2021-07-25T16:40:18Z	https://en.wikipedia.org/wiki/Communication	Communication - Wikipedia	Hose	1@1.us	
3	2021-07-25T16:40:19Z	https://en.wikipedia.org/wiki/Communicate_(disambiguation)	Communicate (disambiguation) - Wikipedia	Hose	1@1.us	
4	2021-07-25T16:41:43Z	https://en.wikipedia.org/wiki/Communicate_(Sasha_%26_John_Digweed_album)	Communicate (Sasha & John Digweed album) - Wikipedia	Hose	1@1.us	
5	2021-07-25T18:15:03Z	https://en.wikipedia.org/wiki/Communicate_(Sasha_%26_John_Digweed_album)	Communicate (Sasha & John Digweed album) - Wikipedia	Hose	1@1.us	
6	2021-07-25T18:15:20Z	https://en.wikipedia.org/wiki/Communicate_(Sasha_%26_John_Digweed_album)	Communicate (Sasha & John Digweed album) - Wikipedia	Hose	1@1.us	
7	2021-07-25T18:15:54Z	https://en.wikipedia.org/wiki/Sasha_%26_John_Digweed	Sasha & John Digweed - Wikipedia	Hose	1@1.us	
8	2021-07-25T18:49:35Z	https://en.wikipedia.org/wiki/Drum_machine	Drum machine - Wikipedia		1@1.us	
9	2021-07-25T18:50:29Z	https://en.wikipedia.org/wiki/Wikipedia:Verifiability	Wikipedia:Verifiability - Wikipedia	Somebody	1@1.ua	
10	2021-07-25T18:51:32Z	https://en.wikipedia.org/wiki/Wikipedia:Policies_and_guidelines	Wikipedia:Policies and guidelines - Wikipedia	Somebody	1@1.ua	
11	2021-07-25T18:53:20Z	https://en.wikipedia.org/wiki/Wikipedia,the_free_encyclopedia	Wikipedia, the free encyclopedia - Wikipedia	Somebody	1@1.ua	
12						

Рис. 8. Результуючі записи таблиці

Таким чином, уся робота команди аналітиків відображатиметься у таблиці, яка може бути доступною визначеному колу осіб.

Можливості Hunchly не обмежуються описаним. Наприклад, під час проведення досліджень можуть бути застосовані комбіновані інструменти автоматизації обробки великої кількості накопичених даних. Самі ці дані можуть бути експортовані у вигляді архіву, в якому вони структуровані і описані за допомогою звіту (рис. 9).

Test

Case created: July 24, 2021 22:08:22

Export created: July 26, 2021 1:43:11

ID	Title	Date	URL	Screenshot	Content hash
25	Communicate (Sasha & John Digweed album) - Wikipedia	July 25, 2021 21:15:03	https://en.wikipedia.org/wiki/Communicate_(Sasha_%26_John_Digweed_album)		af67d63b126735e0e90b5c10be950099e7b1111c2e121f816e52e1f915e1a
27	Communicate (Sasha & John Digweed album) Revision history - Wikipedia	July 25, 2021 21:15:20	https://en.wikipedia.org/w/index.php?title=Communicate_(Sasha_%26_John_Digweed_album)&action=history		561330929e30e9295e2e52030e2b2795d810e70e67b9711405f5021199eb
28	Sasha & John Digweed - Wikipedia	July 25, 2021 21:15:54	https://en.wikipedia.org/wiki/Sasha_%26_John_Digweed		97c19826883291571f9f8d0557f8fb8e0d8da9987c0e9761f47ebc7c945e

Рис. 9. Експортовані дані справи

У папці звіту дані розподілені у відповідних папках. Наприклад, фотографії представлені у каталозі photos. Їх можна одразу попередньо проаналізувати з

Для завантаження раніше одержаних даних у Gephi потрібно створити дві таблиці. Перша міститиме інформацію про вершини графа (Nodes), інша – про його ребра (Edges).

Слід звернути увагу, що програма є чутливою до регістру, тому значення «АбВ» не буде дорівнювати «абв».

Обов'язковими полями в таблиці Nodes є Id та Label, а у Edges – Source і Target.

У якості прикладу візьмемо відомості про учасників певної групи (рис. 11), де в таблиці Nodes буде розміщено Id користувачів Telegram (поле Id), відомості про них: ім'я, назва облікового запису, телефон (поле Label). Також окремою колонкою слід додати кількість повідомлень, розміщених користувачами в усіх досліджуваних групах (функція Excel СУММЕСЛИ). Наприкінці таблиці окремими рядками слід записати інформацію про самі групи (поле Id – назва групи, поле Label – посилання на нього, Messages count – кількість повідомлень в групі). Перенести відомості з декількох полів до одного можна за допомогою функції Excel СЦЕПИТЬ, наприклад, =СЦЕПИТЬ(B2;" ";C2;" ";D2;" ";E2).

	A	B	C
1	Id	Label	Messages count
2	10000021	На [REDACTED] 552	1
3	10000075	☺ @ [REDACTED] as	43
4	10000044	Ви [REDACTED] ма	108
5	10000047	Ми [REDACTED] 0710	8
6	10000000	Не [REDACTED] а	80

Рис. 11. Таблиця вершин графу

У таблиці Edges створимо чотири поля та занесемо до них такі відомості (рис. 12):

- поле Source – назва каналу
(=ПСТР(ЯЧЕЙКА("ИМЯФАЙЛА";A1);ПОИСК("");ЯЧЕЙКА("ИМЯФАЙЛА";A1))+1;255)

- поле Target – Id користувача Telegram, який є учасником відповідної групи;
- поле Label – відомості про учасника групи (ім'я, назва облікового запису, телефон);

- поле Messages Count – відомості про кількість повідомлень, залишених користувачем у групі (цей показник можна вирахувати за допомогою функції Excel СЧЁТЕСЛИ, наприклад, =СЧЁТЕСЛИ(Шлях_до_книги_на_диску\[Назва_файлу_з_повідомленнями.xlsx]Sheet1!\$D\$2:\$D\$13601;B2)).

	A	B	C	D	E	F	G	H
1	Source	Target	Label	first_name	last_name	username	phone	Messages count
2	e [REDACTED] Ukr	10000021	На [REDACTED] 552	Н [REDACTED]	Н [REDACTED] о		380 [REDACTED] 552	1
3	e [REDACTED] Ukr	14000075	☺ @Mi [REDACTED] as	☺		@M [REDACTED] as		14
4	e [REDACTED] Ukr	14000044	Ви [REDACTED] ма	Ви [REDACTED]	М [REDACTED] а			64
5	e [REDACTED] Ukr	10000047	Ми [REDACTED] 0710	М [REDACTED]	А [REDACTED] n	@ [REDACTED] 710		8
6	e [REDACTED] Ukr	12000090	Не [REDACTED] ица	Не [REDACTED] зя	[REDACTED] ица			6
7	e [REDACTED] Ukr	16000027	М [REDACTED]	М [REDACTED]				28
8	e [REDACTED] Ukr	66000085	.	.				169
9	e [REDACTED] Ukr	74000030	N @M [REDACTED] 222	N		@MIMIMI111222		108
10	e [REDACTED] Ukr	14000052	Раз [REDACTED] 5434	Р [REDACTED] я	А [REDACTED] а		380 [REDACTED] 54	12
11	e [REDACTED] Ukr	12000086	Ли [REDACTED] 7564	Л [REDACTED]			380 [REDACTED] 4	148

Рис. 12. Таблиця ребрів графу

Для завантаження підготовлених даних до програми Gerhi потрібно створити новий проект, перейти до вкладки «Лаборатория данных» та обрати відповідний файл для імпорту (рис. 13).

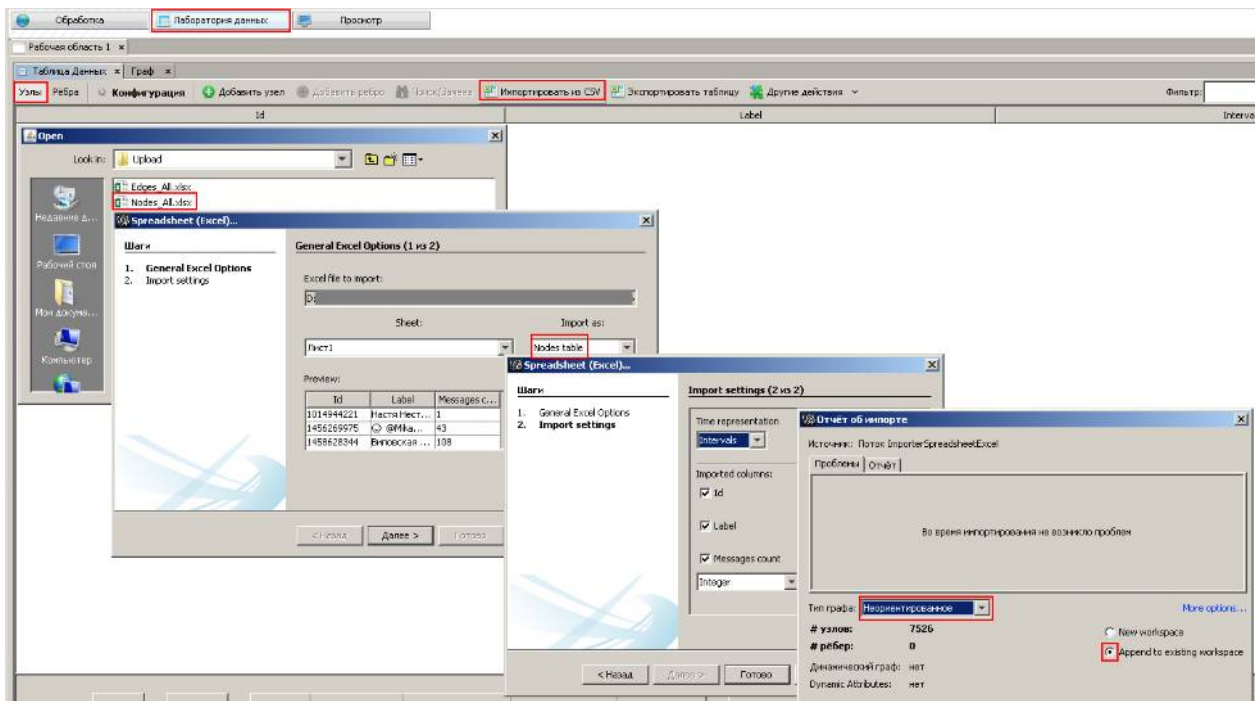


Рис. 13. Імпортування даних про вершини графу
Подібним чином слід імпортувати дані про ребра графів (рис. 14)

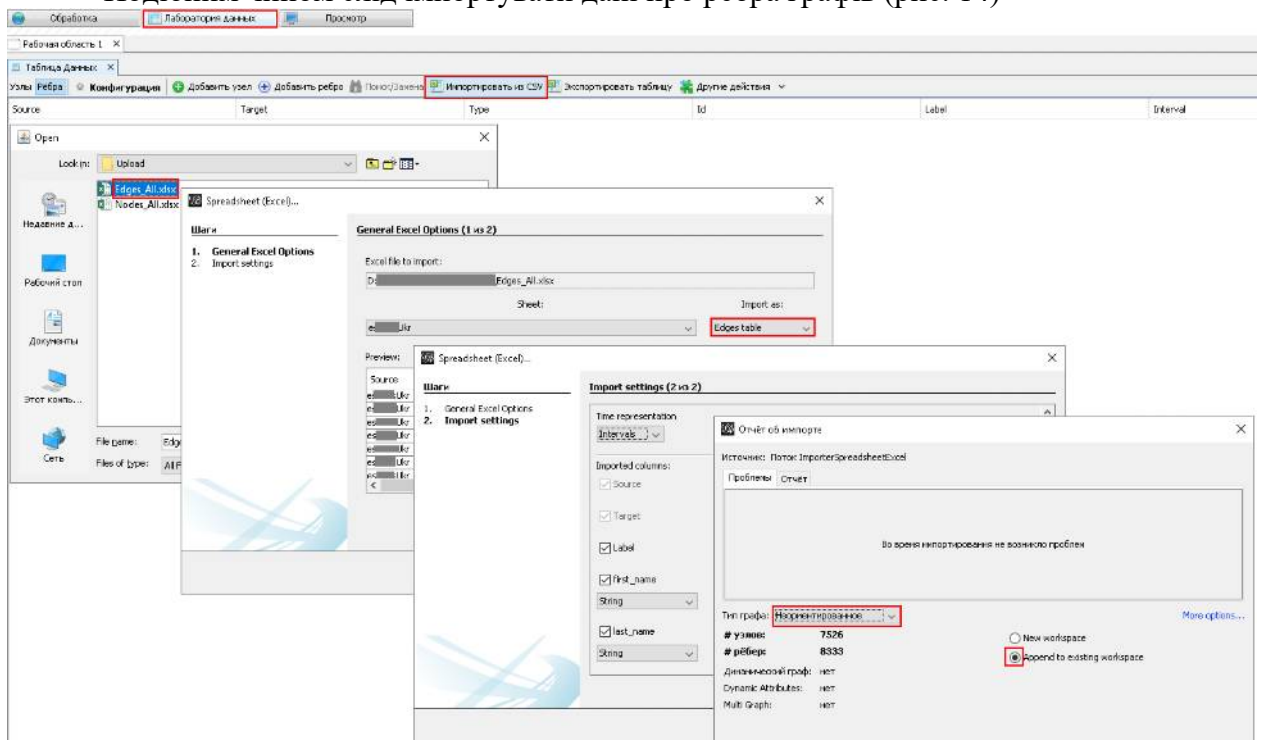


Рис. 14. Імпортування даних про ребра графу
Якщо усі дані імпортовано коректно, то можна побачити первинний граф у вкладці «Обработка» (рис. 15).

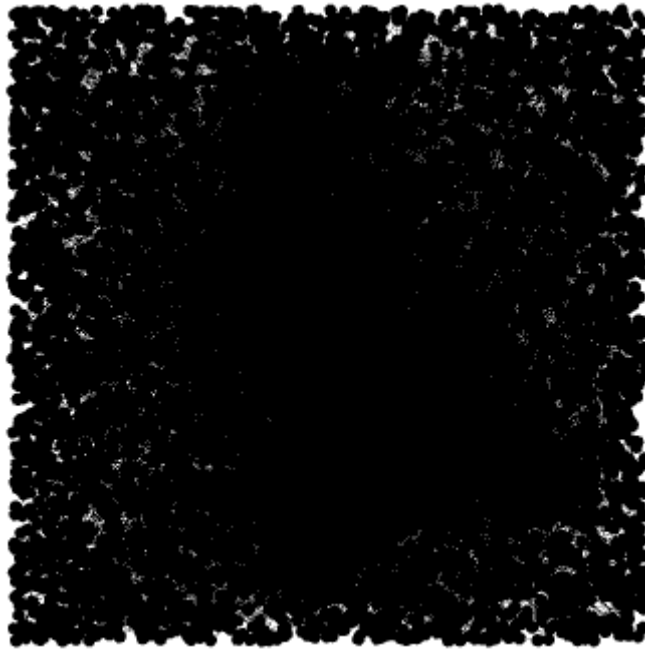


Рис. 15. Первинний граф

Як видно з рисунку, граф у такому вигляді є не дуже інформативним, тому його потрібно належним чином структурувати. Для цього, передусім, слід у вкладці «Укладка» обрати відповідний алгоритм упорядкування. У даному прикладі скористаємося алгоритмом ForceAtlas 2, у налаштуваннях якого слід обрати LinLog режим (більш щільні кластери), «Запрет перекрестия», Разреженность – 5.0 (рис. 16).

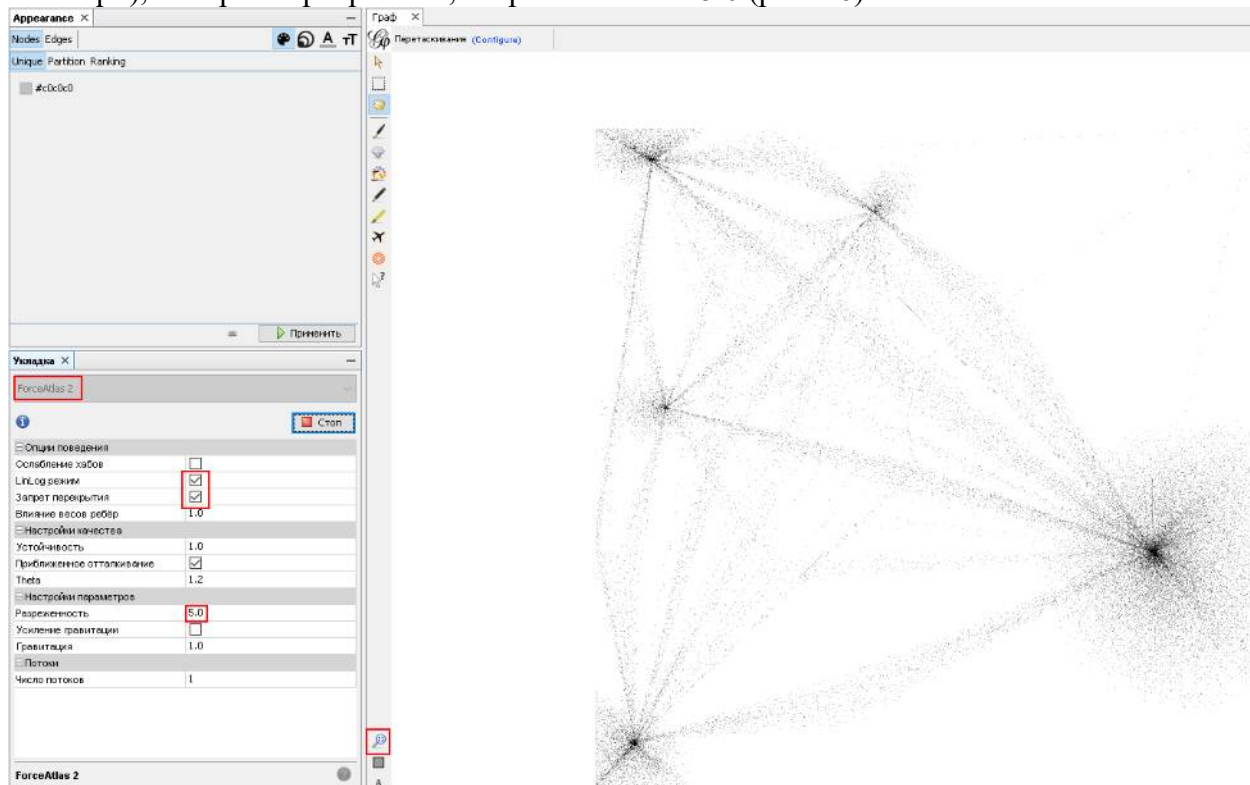


Рис. 16. Результати первинного упорядкування

Проводити маніпуляції щодо переміщення елементів графу можна за допомогою правої кнопки миші (рух схеми), колеса мишки (збільшення/зменшення масштабу), а також відповідних кнопок у нижній панелі, наприклад, кнопка із зображенням лупи центрує граф.

Перед тим, як переходити до наступного етапу аналізу слід розкрити значення окремих елементів теорії графів, які досить добре розтлумачено в матеріалі «Сетевой

анализ. Работа с Gephi» [Сетевой анализ. Работа с Gephi. URL: <https://ancatmara.gitbooks.io/digital-literacy-for-sfl/content/seminar-11.html> (дата звернення: 12.02.2021)]. Отже, потужність вузла (degree) визначається кількістю його зв'язків, зважений степінь (weighed degree) – кількістю зв'язків вузла, поділеною на загальну кількість зв'язків у графі.

Важливість вузла визначається різними способами, серед яких:

- центральність за степенем (degree centrality): важливішим є вузол, який має більше зв'язків;
- центральність за близькістю (closeness centrality): чим коротшим є шлях від вузла до решти вузлів, тим він важливіший;
- центральність за посередництвом (betweenness centrality) визначається кількістю найкоротших шляхів, що проходять через вузол;
- центральність за впливовістю (eigencentrality): чим більшою є кількість вузлів, пов'язаних з вузлами, зв'язаними з досліджуванним вузлом, тим він важливіший («чим більше друзів у ваших друзів, тим ви важливіші»).


Коефіцієнт асортативності (assortativity coefficient) визначає пов'язаність «важливих вузлів». Чим він більший, тим більше «важливих» вузлів пов'язані між собою, і навпаки.

Коефіцієнт кластеризації (clustering coefficient) визначає степінь взаємодії між собою найближчих сусідів вузла, тобто ймовірність, що найближчі сусіди вузла будуть пов'язані не тільки з ним, але і між собою.

Щільність графа (density) – відношення числа ребрів до максимально можливого. У співтовариствах високий коефіцієнт кластеризації та висока щільність.

Модулярність (modularity) показує, наскільки при заданому розбитті графа на групи щільність зв'язків усередині групи більше щільності зв'язків між групами. За допомогою цієї метрики граф розбивається на співтовариства.

У наведеному прикладі саме модулярність розрахуємо наступною. Для цього слід у вікні Контекст обрати меню Статистики та натиснути кнопку Пуск навпроти слова Модулярність (рис. 17). У даному прикладі під час розрахунку не буде враховуватися вага ребрів, оскільки вона не визначена.

Після розрахунку модулярності змінимо кольори для різних співтовариств. Для цього у вікні Appearance слід натиснути кнопки Nodes та Колір та у розділі Partition Modularity натиснути Применить. Для більш контрастного зображення слід змінити колір фону за допомогою кнопки .

Як видно з рисунку система виділила сім співтовариств та розрахувала частку кожного у загальному графі. Ці показники відповідають завантаженому раніше даним.

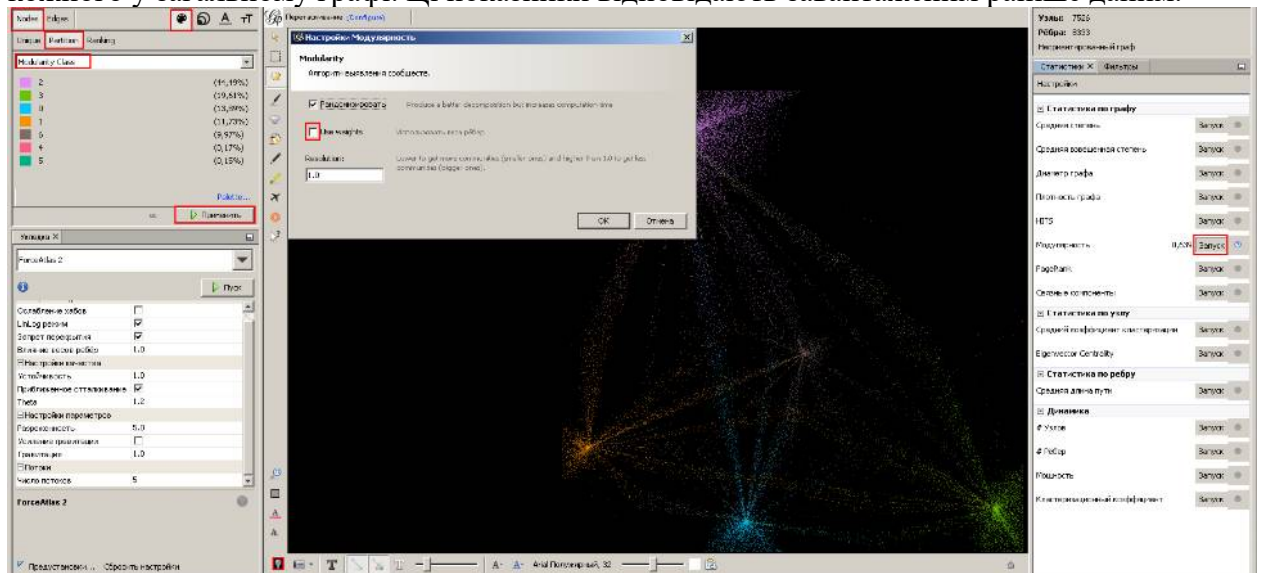


Рис. 17. Обчислення модулярності графу

Після побудови графу можна скорегувати розмір вузлів за визначеними параметрами. Спершу слід виокремити вузли, пов'язані з якомога більшою кількістю груп. З цією метою потрібно провести ранжування розміру вузлів за критерієм центральності за близькістю. Для цього попередньо слід розрахувати середню довжину шляху з нормуванням результатів у вікні Статистики. Після виконання вказаних дій ближче до середини графу можна побачити точки більшого розміру, які будуть вказувати на ключових осіб, присутніх у декількох групах одночасно (рис. 18).

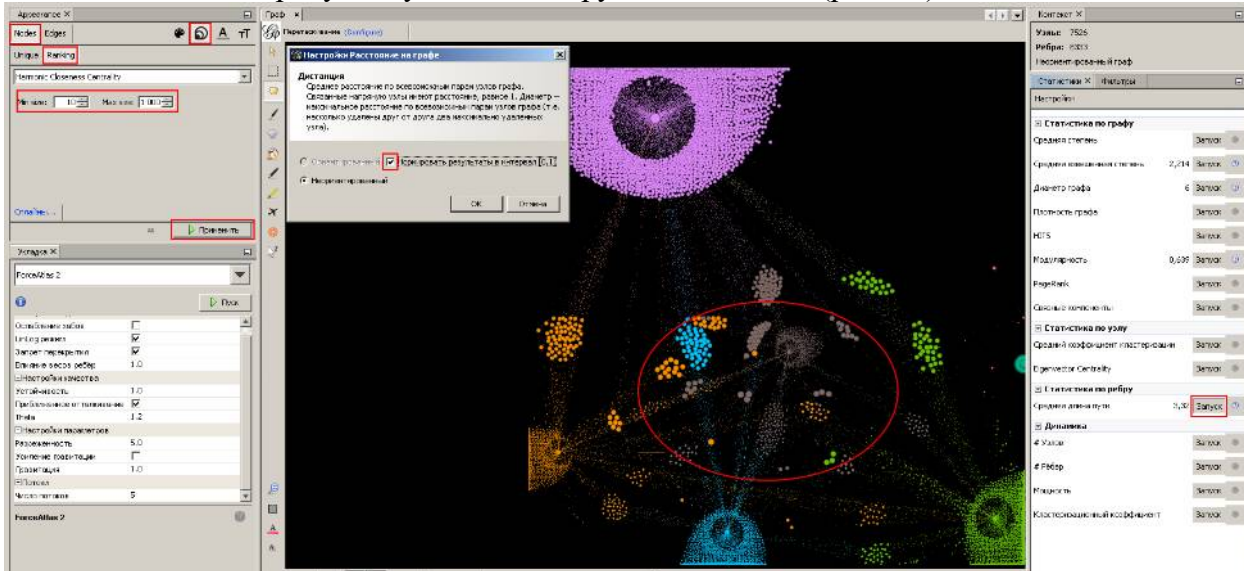


Рис. 18. Збільшення розмірів ключових вузлів за близькістю
Для більшої наочності результати слід відфільтрувати як на рис. 19.

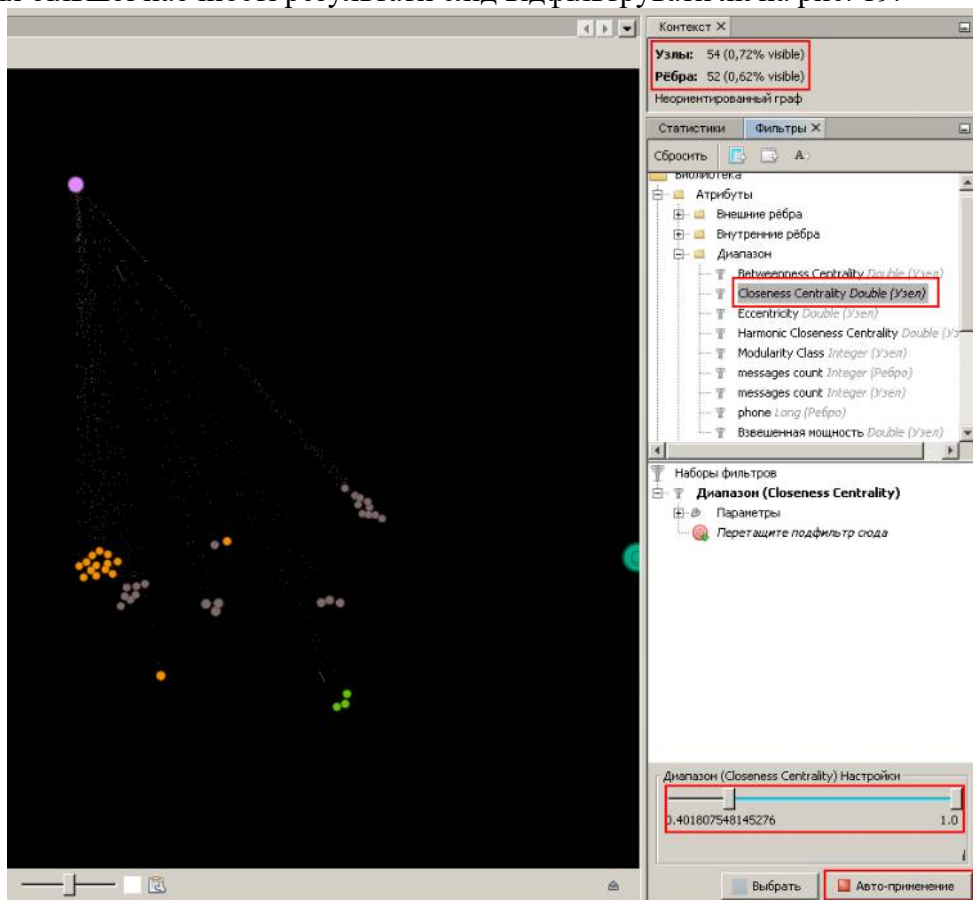


Рис. 19. Результати фільтрації

Унаслідок фільтрації як на рисунку можна лишити тільки вузли, що входять як мінімум до трьох груп одночасно.

Після зменшення величини графу та визначення ключових фігур можна вивести підписи для кожного вузла, натиснувши **T** у нижній частині вікна Граф. Для того, щоб мітки не накладалися одна на одну, слід обрати у вікні Укладка алгоритм упорядкування Укладка меток.

Після виокремлення вузлів з найбільшою кількістю зв'язків із різними групами можна змінити розмір наявних вузлів пропорційно кількості залишених повідомлень у групі, а також додати підфільтр для відсіювання за кількістю повідомлень (рис. 20).

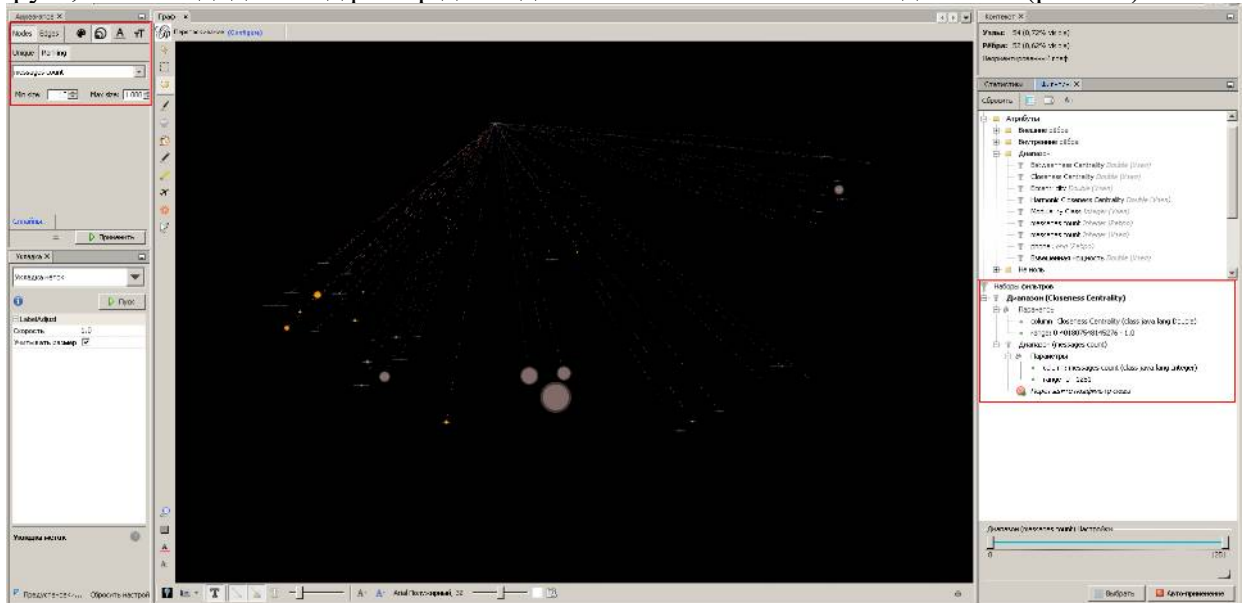


Рис. 20. Граф, який відображає найбільш активних учасників груп

У результаті одержимо граф з потрібною нам структурою. Для його перенесення у формат зображення потрібно перейти в меню Просмотр, обрати відповідні налаштування та натиснути кнопку Обновить (рис. 21).

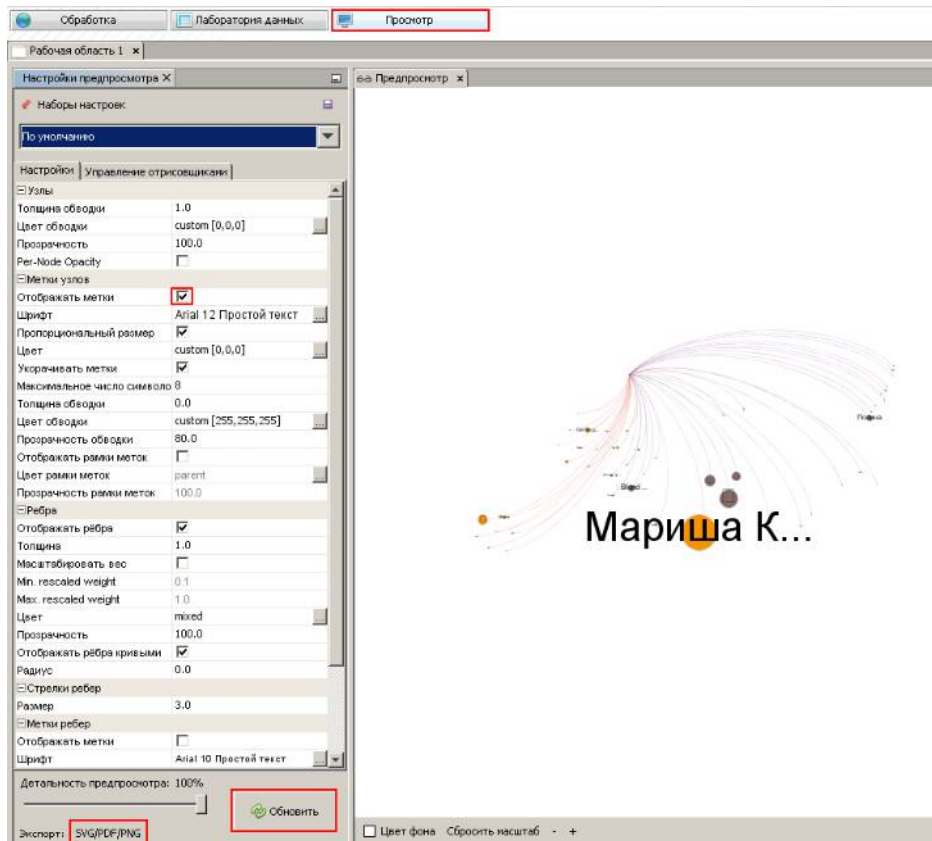


Рис. 21. Экспорт графа до файлу зображення

Під час експорту можна налаштувати якість підсумкового зображення. Для підвищення якості експортованих зображень слід змінити у файлі `etc/gephi.conf` параметр `default_options`. Наприклад, `default_options="--branding gephi -J-Xms256m -J-Xmx4096m -J-Xverify:none -J-Dsun.java2d.noddraw=true -J-Dsun.awt.noerasebackground=true -J-Dnetbeans.indexing.noFileRefresh=true -J-Dplugin.manager.check.interval=EVERY_DAY` (рис. 22).

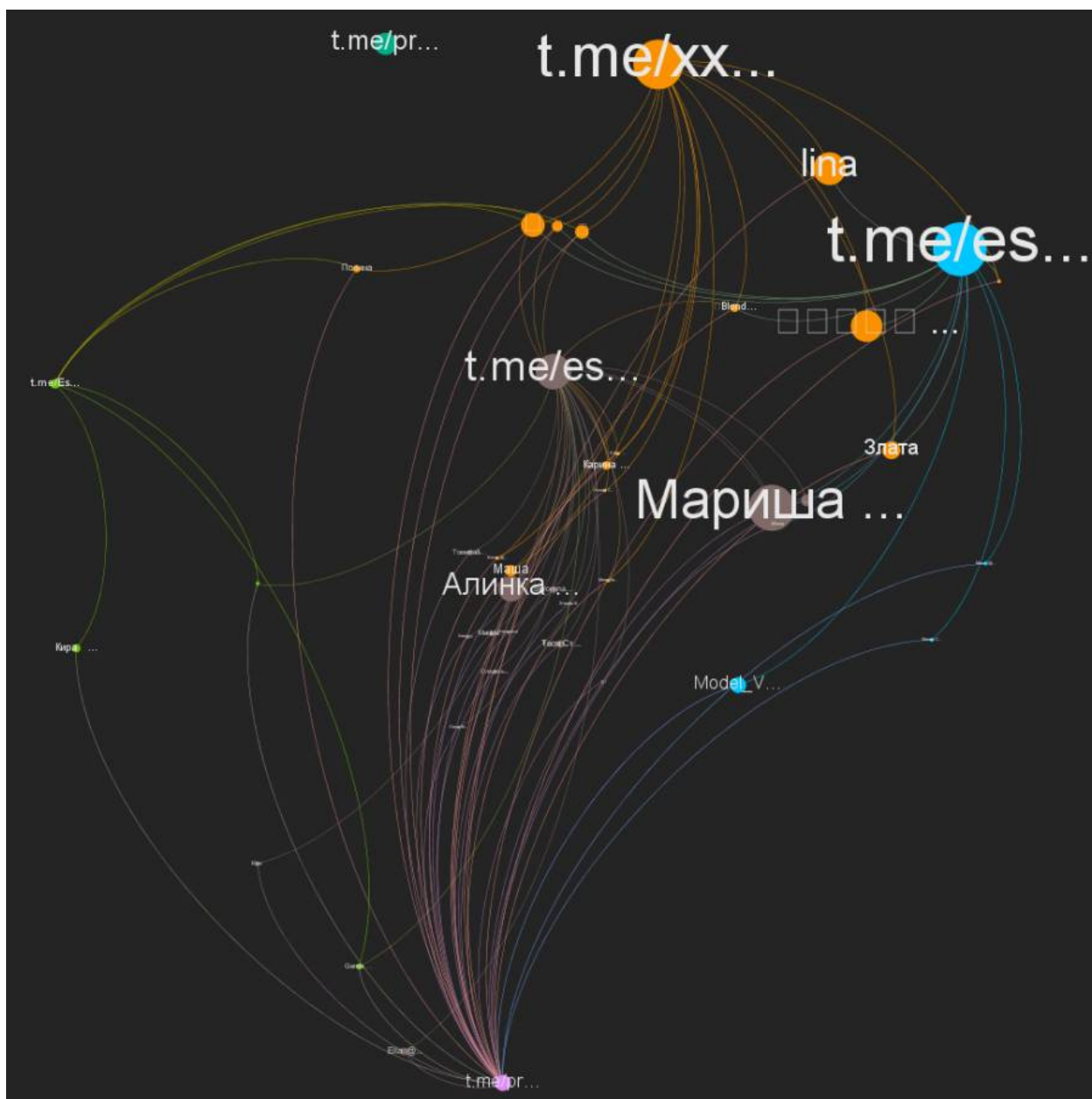


Рис. 22. Результуюче зображення

Для покращення сприймання зображення можна визначити в налаштуваннях вікна «Просмотр» відповідну товщину ребрів, наприклад, 17.0.

3. Збирання та аналіз інформації про веб-сайти

Головними завданнями, які виникають під час вивчення окремого або групи веб-сайтів є:

- встановлення IP-адреси за доменним іменем;
- визначення інших доменних імен, прив'язаних до однієї IP-адреси;
- визначення адміністратора та володільця сайту;
- побудова структури взаємозв'язків між різними мережними ресурсами;
- фіксація контенту;
- збирання та аналіз додаткової інформації.

Для вирішення описаних завдань можуть бути використані різні інструменти.

Однією з потужних платформ, яка накопичує інформацію про різні мережні ресурси та надає можливість її структурованої обробки і аналізу є Microsoft Defender Threat Intelligence (ti.defender.microsoft.com) /стапа назва RISKIQ (riskiq.com)/. Дане рішення має

декілька інтерфейсів. Найбільш простим з яких є доступ через веб. Ресурс потребує реєстрації, для чого бажано використовувати корпоративну електронну пошту. Реєстрація з використанням такого облікового запису надає більше можливостей для пошуку.

Після реєстрації достатньо ввести шуканий ресурс у вікно пошуку (рис. 23).



Рис. 23. Вікно введення даних для пошуку

Знайдені результати представляються у зручному вигляді як на рис. 24.

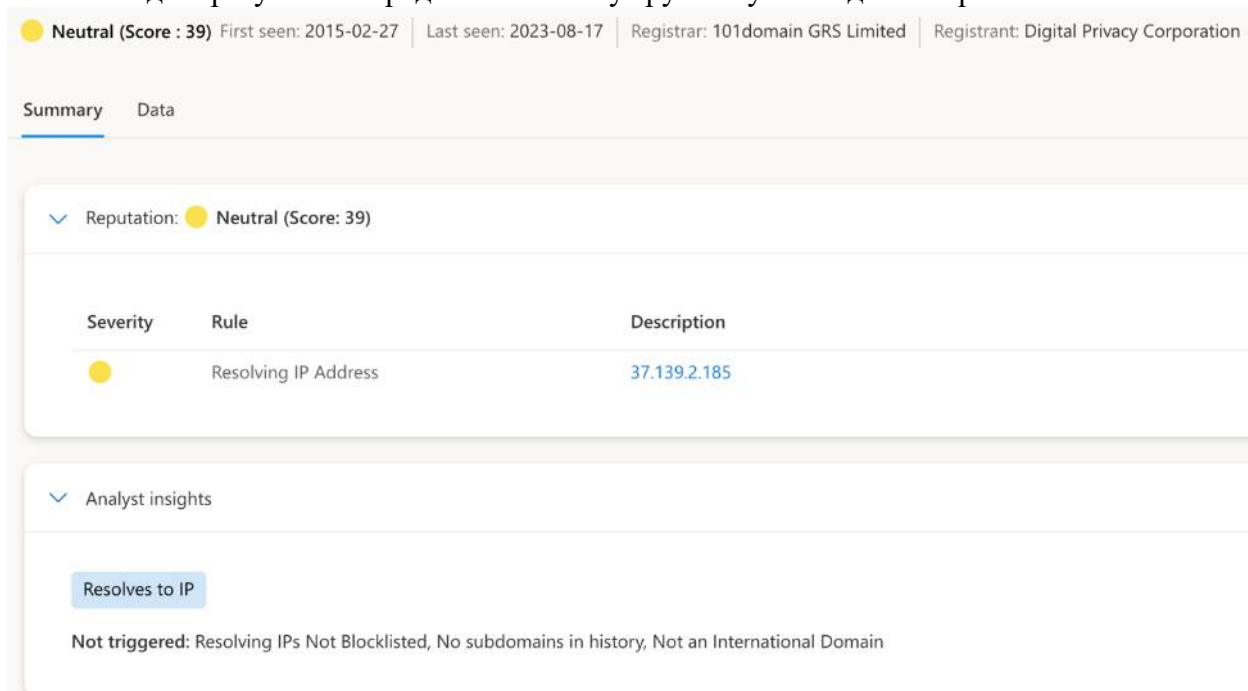


Рис. 24. Результати пошуку

Знайдені результати дозволяють вивчити застосовувані сертифікати, існуючі субдомени, історію функціонування сайту (як правило, від початку його створення із вказівкою використовуваних IP-адрес, реєстраційних даних) тощо. Цікавою інтегрованою можливістю є дослідження аналітичних метрик, які можуть бути присутніми у вихідному коді сайту. За їх допомогою можна зробити попередній висновок про сайти, які адмініструють з одного облікового запису (рис. 25).

Trackers

< > 1 - 7 of 7

Type Hostname

HOSTNAME	First seen	Last seen ↓	Type	Value
sexo.185.67.1.153.com	2023-05-27	2023-08-09	YandexMetricId	2998
sexo.185.67.1.153.com	2021-03-20	2023-08-09	IntercomAppId	pdv
sexo.185.67.1.153.com	2023-05-27	2023-08-09	YandexMetricaCounterId	2998
sexo.185.67.1.153.com	2021-03-20	2021-11-26	GoogleAnalyticsAccountNumber	ua-150
sexo.185.67.1.153.com	2021-03-20	2021-11-26	GoogleAnalyticsTrackingId	ua-150
sexo.185.67.1.153.com	2021-03-20	2021-03-20	GoogleAnalyticsAccountNumber	ua-158
sexo.185.67.1.153.com	2021-03-20	2021-03-20	GoogleAnalyticsTrackingId	ua-158

Рис. 25. Вивчення результатів

Крім спеціалізованих платформ для пошуку інформації за IP-адресою можуть бути застосовані і класичні інформаційно-пошукові системи. Так, наприклад, виконавши на сайті [bing.com](https://www.bing.com) запит `IP:***.***.***.***`, де замість зірочок вказується відповідна IP-адреса, можна дізнатися, які доменні імена прив'язано до цієї адреси. У випадку, якщо результат виконання запиту не буде виводитись на екран, слід видалити частину результуючої URL, як на рис. 26.

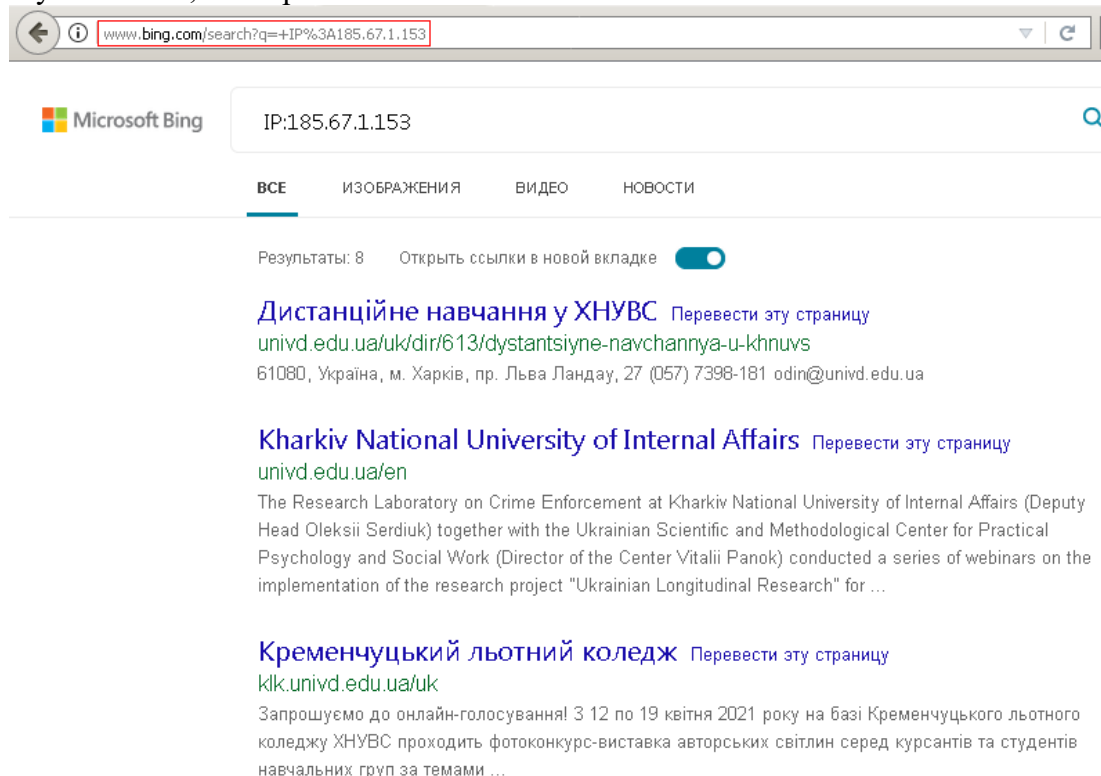


Рис. 26. Встановлення доменних імен, які прив'язані до однієї IP-адреси

Слід відзначити, що ринок реклами в мережі Інтернет обумовлює необхідність оцінки певних ресурсів. Для цього серед іншого використовуються спеціальні аналітичні метрики (коди відстеження), які розміщуються у вихідному коді певних сторінок сайту.

Вказані метрики можуть бути використані під час аналізу зв'язків між сайтами та визначення тих з них, які адмініструються однією або декількома пов'язаними особами.

Визначити метрики для кожного сайту можна:

- вручну, для чого слід натиснути правою кнопкою миші в браузері на потрібній сторінці, обрати відкриття вихідного коду та дослідити його зміст;

- автоматизовано, коли відповідне вилучення трекерів виконується з використанням спеціалізованих сервісів або програм. Програма може бути підготовлена самостійно, як от, наприклад, описано за вказаною адресою: bellingcat.com/resources/2017/07/31/automatically-discover-website-connections-tracking-codes/. Водночас для вилучення метрик можливо застосовувати вже готові рішення. Серед них потрібно виділити сервіси: RISKIQ (riskiq.com), Spyonweb (spyonweb.com) та Tracker Tracker (tools.digitalmethods.net/beta/trackerTracker/#). При використанні останнього вказаного інструменту (рис. 27) слід пам'ятати, що завантажені до нього дані можуть бути доступними іншим особам.

Рис. 27. Вікно сервісу Tracker Tracker

Підготовлені для аналізу дані (рис. 28) можуть бути завантажені як до мережних сервісів побудови графів, як от, , так і до спеціалізованих застосунків, як от Gephi.

A			B		
1	Id	Label	1	Source	Target
2	aze	vs.com	2	az	Google UA-382
3	b.bel	info	3	b.b	Yandex 69384
4	bing	slugi	4	b.b	Google UA-18022
5	bishke	om	5	bing	Google UA-10518
6	bor	et	6	bisl	Google UA-1273
7	cary	com	7	borc	Google UA-54695032
8	cherr	com	8	car	Google UA-12696
9	elite	com/uslugi	9	cher	Google UA-19202
10	e	ge	10	elit	Yandex 504

Рис. 28. Сформовані дані для аналізу

В останньому випадку для візуалізації кластерів певного розміру можна попередньо застосувати метод укладки Force Atlas 2 або інший, розрахувати модулярність та скористатися фільтром «Modularity Class (Node)» з меню «Разбиение (Partition)». Цей фільтр допоможе залишити на схемі вузли та ребра, які належать вибраній множині класів розбиття (рис. 29).

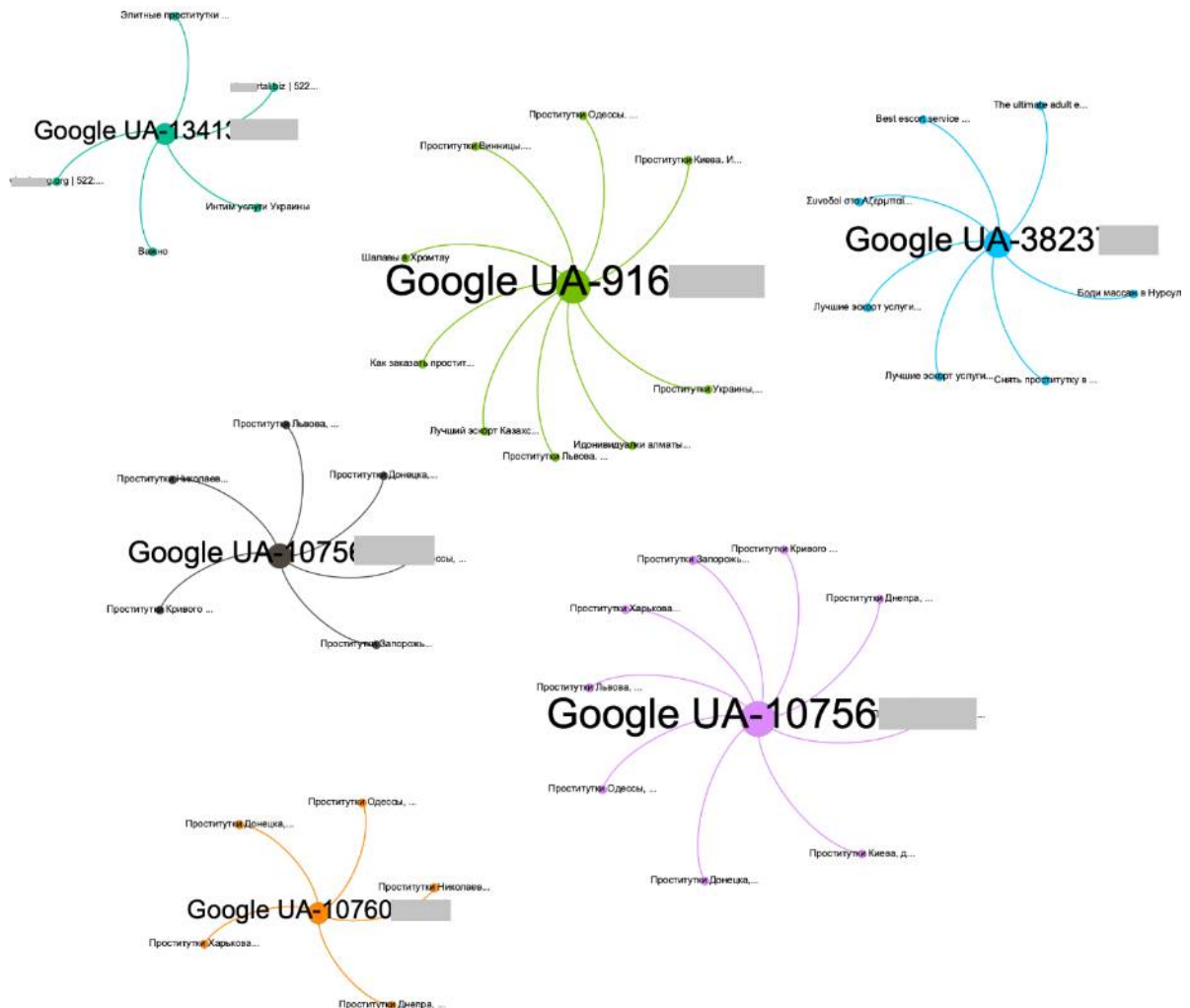


Рис. 29. Результат візуалізації зв'язків сайтів

Такі нескладні маніпуляції дозволять швидко визначити та наочно переглянути зв'язки між конкретними сайтами та спростити аналітику підготовку аналітичного висновку.

4. Завантаження та аналіз даних з соціальних мереж

У роботі аналітика часто виникають завдання, пов'язані з опрацюванням даних із соціальних мереж. Різні підходи до організації та представлення таких даних певним чином ускладнюють аналітичний процес. Враховуючи викладене, представляється доцільним навести один із методів, який може бути використано для вирішення завдань:

- 1) накопичення;
- 2) обробки;
- 3) та аналізу даних із соціальних мереж.

Накопичення даних

На теперішній час існує багато різноманітних програм і сервісів, орієнтованих:

1) на цільове збирання даних з окремих ресурсів із можливістю їх подальшого аналізу;

2) на комплексне вирішення завдань завантаження з ресурсів в мережі Інтернет.

З точки зору універсальності більш вигідно виглядає друга категорія засобів. Зважаючи на це, в якості прикладу продемонструємо окремі техніки накопичення даних за допомогою програми Octoparse (www.octoparse.com), призначеної для вилучення веб-даних.

Після авторизації у встановленому застосунку слід обрати спосіб вилучення даних (рис. 30). У наведеному прикладі будемо використовувати розширений режим.

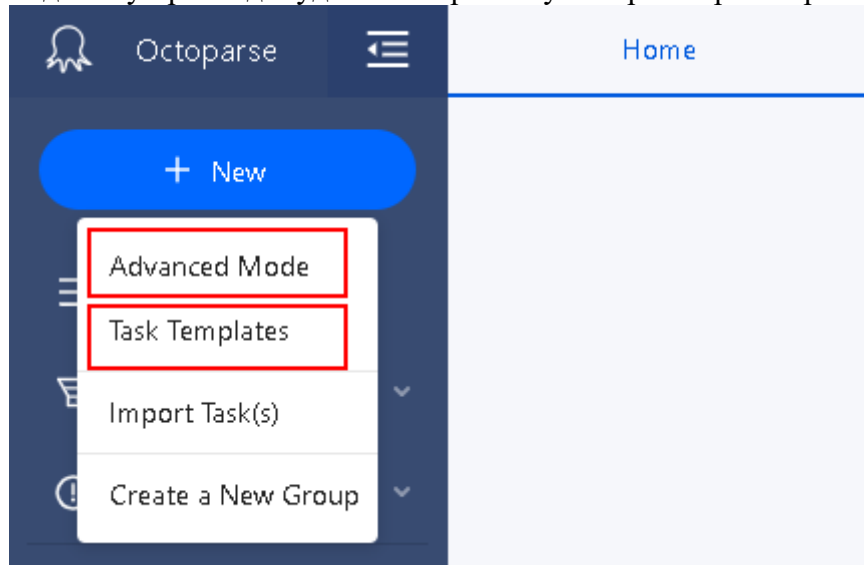


Рис. 30. Обрання способу вилучення даних

У вікні, що відкрилося, слід ввести адресу сайту, з якого потрібно завантажити дані (рис. 31).

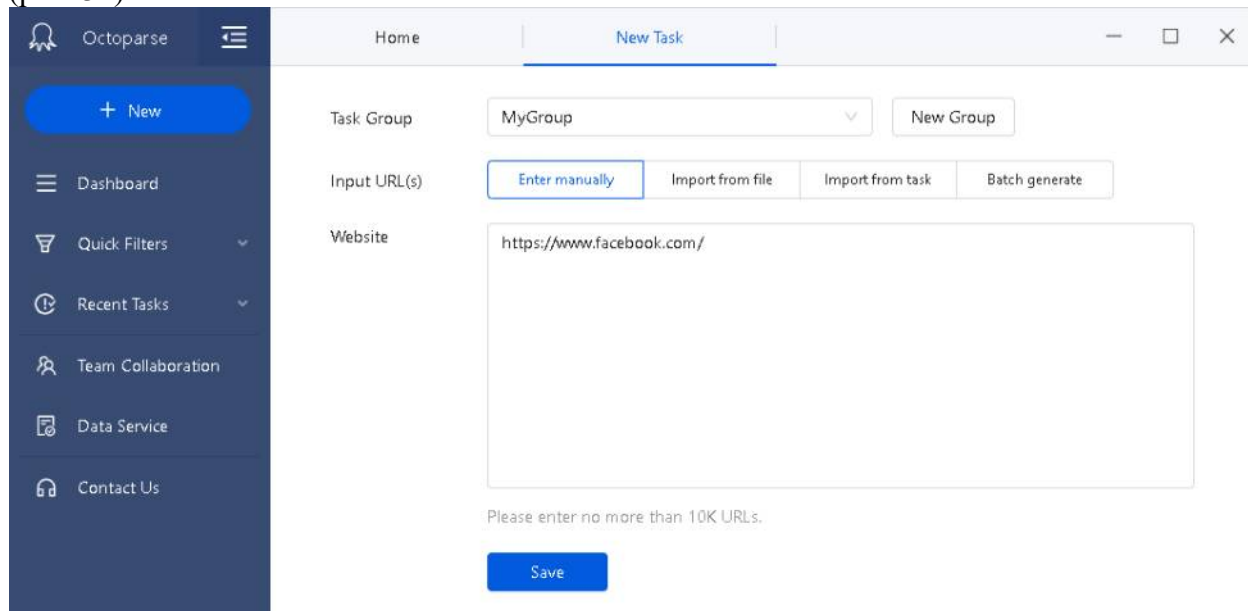


Рис. 31. Визначення сайту, з якого планується завантажити дані

Якщо назву сайту вказано вірно, то після натискання кнопки Save цей сайт відкриється у новій вкладці, після чого для застосунку потрібно визначити послідовність дій з авторизації (рис. 32).

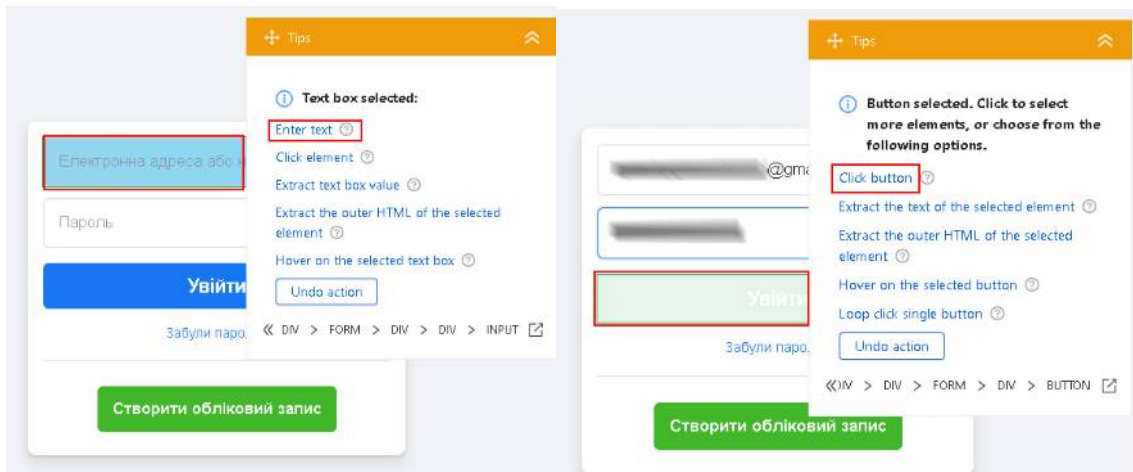


Рис. 32. Налаштування процесу авторизації

Визначена послідовність дій представляється в одному з фреймів програми в якості блок-схеми. Саме у цій частині слід додати блок, за допомогою якого відкриватиметься потрібна сторінка (рис. 3333).

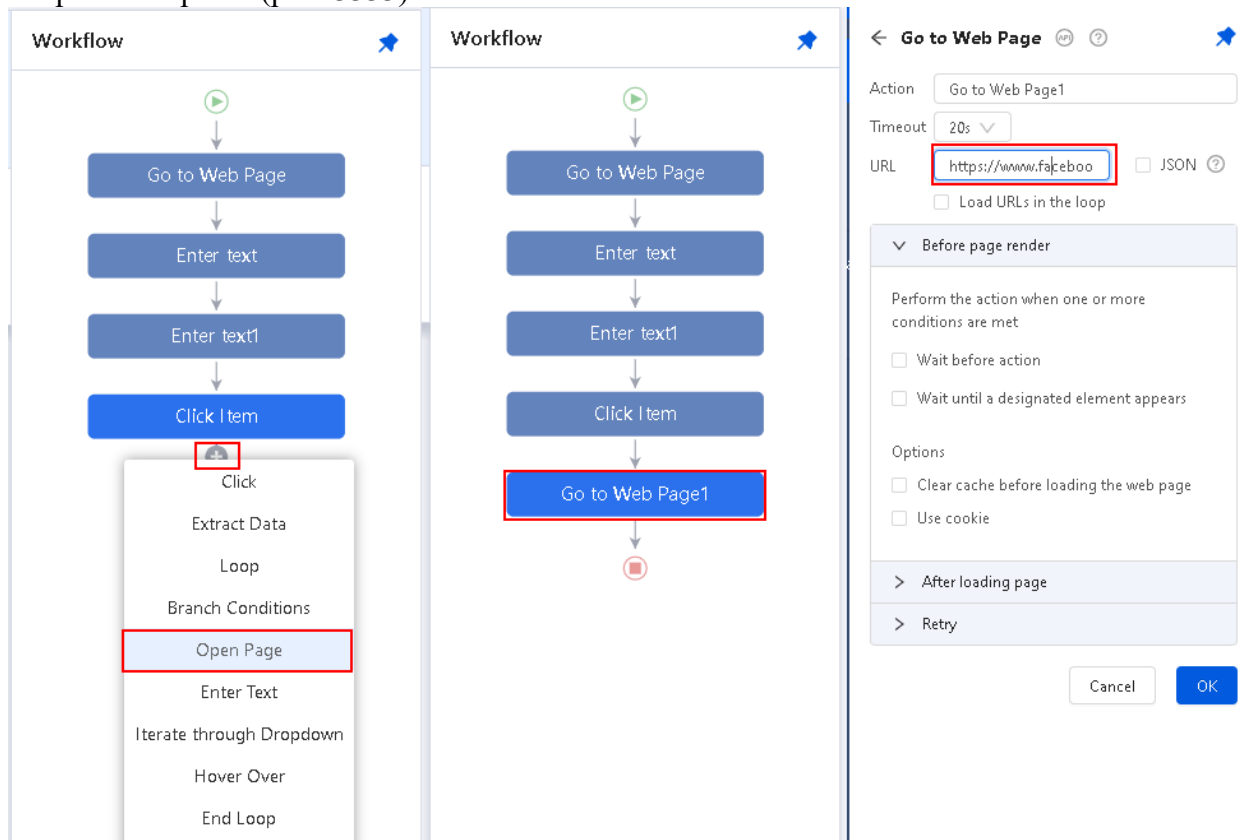


Рис. 33. Налаштування відкриття потрібної сторінки

У даному випадку спробуємо завантажити інформацію про друзів особи, яка становить інтерес, у форматі:

- ім'я особи, яка становить інтерес;
- ім'я друга;
- посилання на його профіль;
- посилання на фотографію профіля друга;
- загальні відомості про друга.

Для виконання цього завдання найпростіше натиснути «Auto-detect web page data». Після проведеного аналізу слід скорегувати структуру відповідного масиву даних для завантаження та натиснути кнопку «Create workflow» (рис. 34).

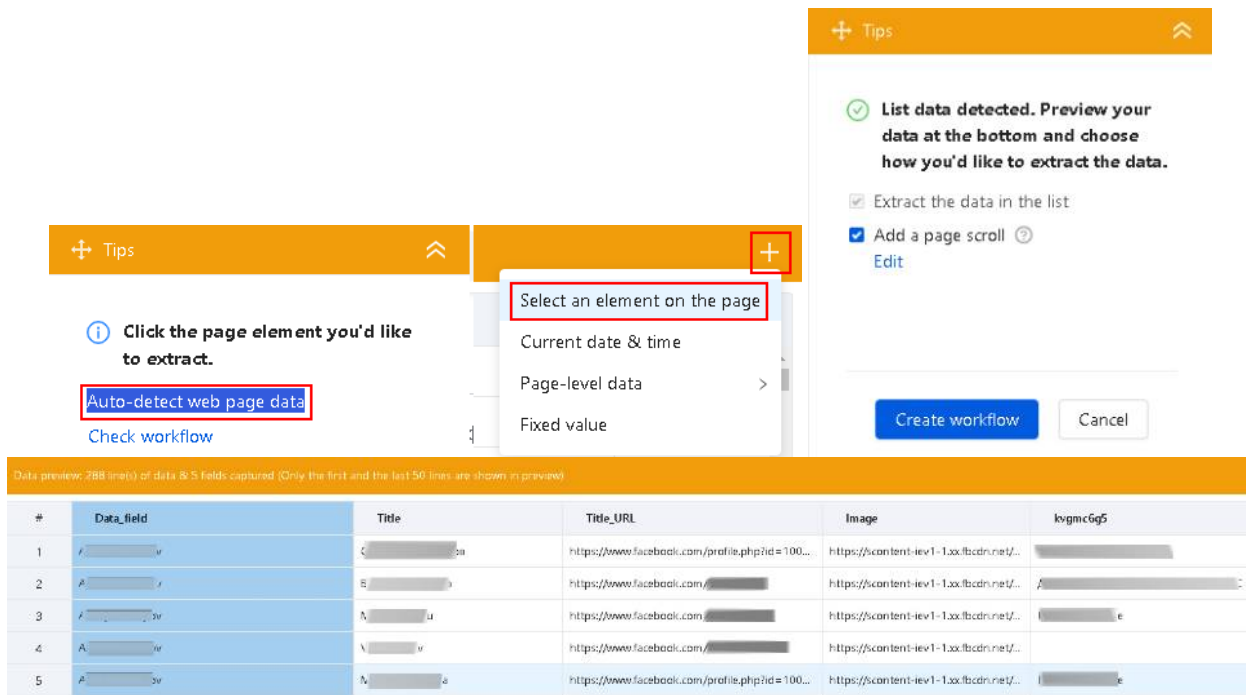


Рис. 34. Визначення структури даних для завантаження

Після того, як сформовано послідовність завантаження даних, в окремих випадках може знадобитися внести зміни до процесу виконання окремих блоків з використанням налаштувань (рис. 35).

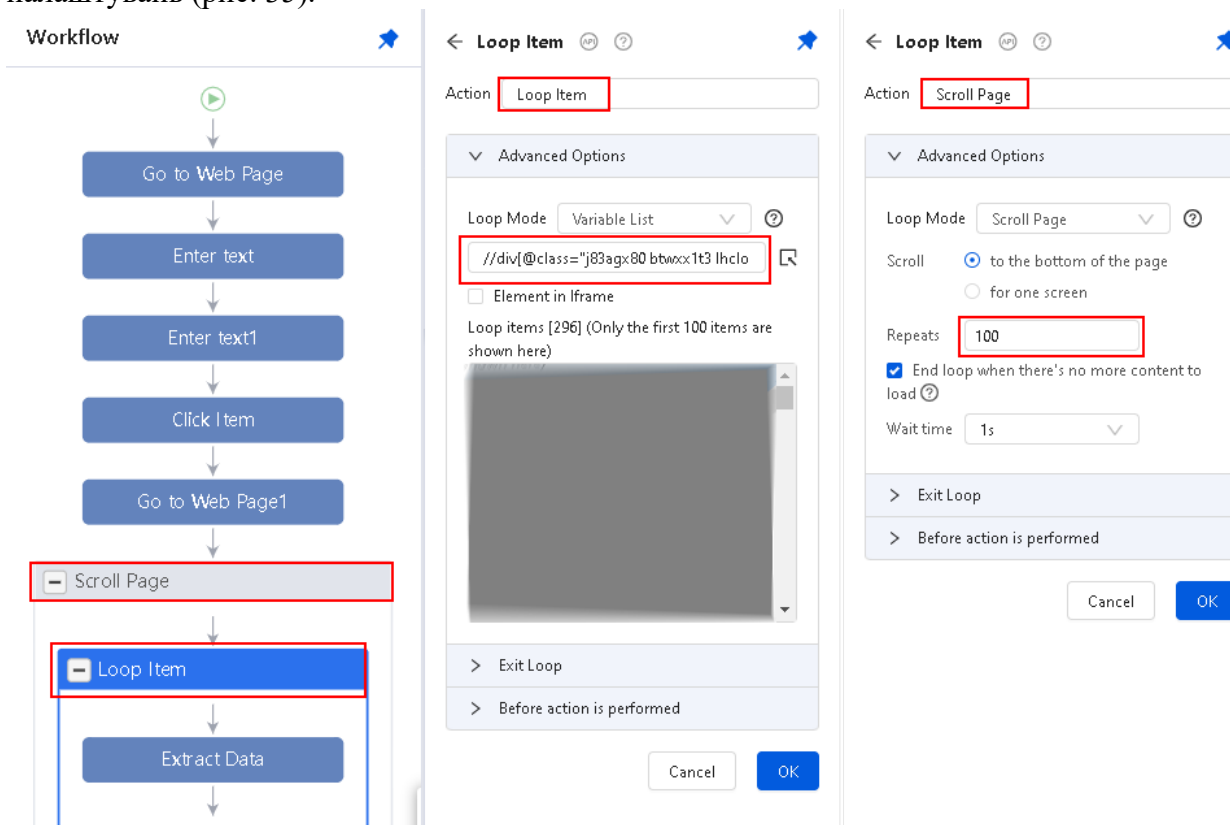


Рис. 35. Налаштування виконання окремих блоків

Після проведення відповідних налаштувань слід зберегти зміни та виконати послідовність шляхом натискання кнопки **Run**. Завантажені дані можна зберегти в різних форматах як на рис. 36.

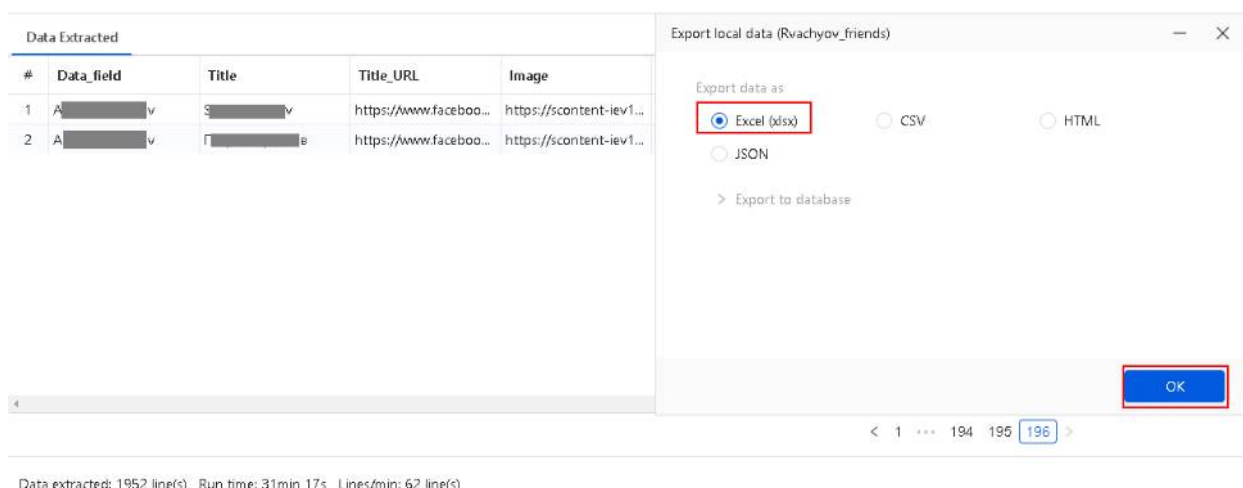


Рис. 36. Експорт завантажених даних

Після того як одержано перший набір даних про друзів особи, яка становить інтерес, його потрібно обробити таким чином, щоб сформувати перелік посилань на друзів інших осіб, яких потрібно також дослідити на етапі аналізу. При цьому потрібно враховувати різницю у формуванні посилань для облікових записів, які містять ID у посиланні профілю та які містять літерні назви профілю, наприклад, www.facebook.com/example. Найбільш вірним у даному випадку представляється одержання ID для таких користувачів, аби у подальшому уніфікувати перелік посилань, наприклад, так:

https://www.facebook.com/profile.php?id=1000*****&sk=friends

Сформований перелік посилань можна зберегти в табличному або текстовому форматі.

Після виконання описаних дій можна циклічно завантажити відомості про друзів осіб, які становлять інтерес, для подальшого аналізу. Для цього в програмі Octoparse потрібно вибудувати цикл, в рамках якого будуть послідовно відкриватися визначені сторінки, і з них завантажуватимуться дані.

Для спрощення виконання цього процесу можна створити нове завдання у розширеному режимі та обрати імпортування адрес з файлу (рис. 37).

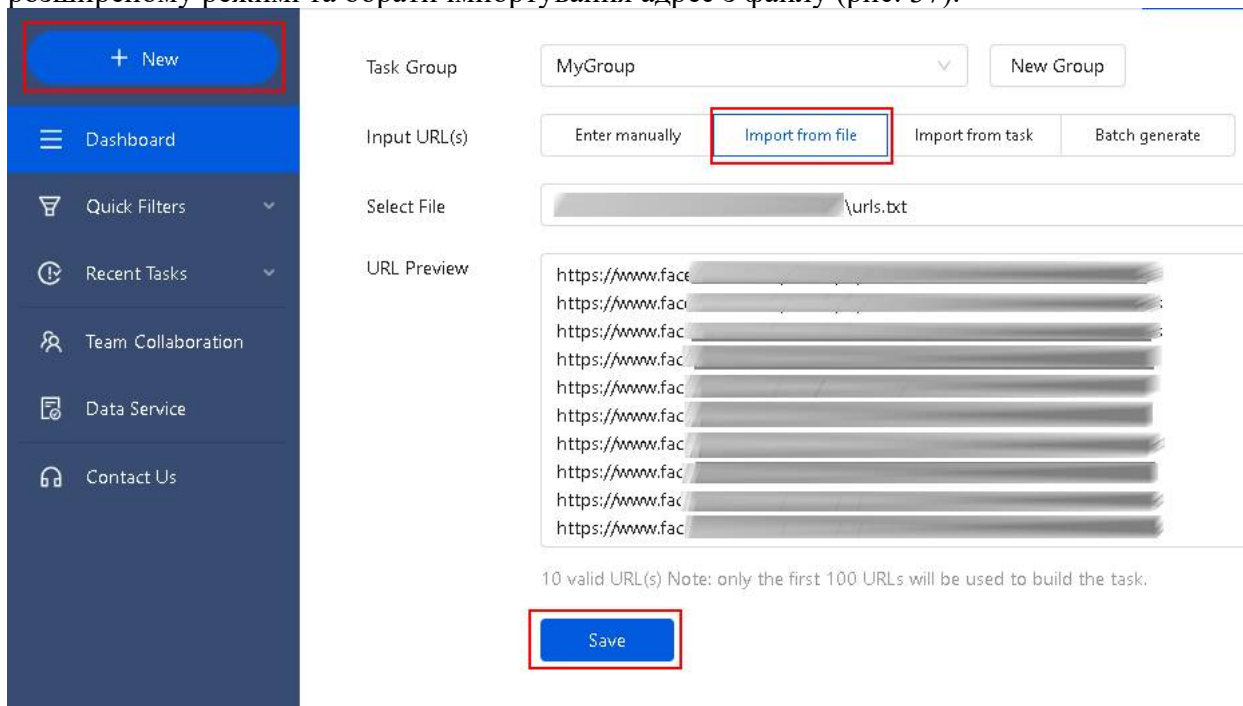


Рис. 37. Налаштування відкриття декількох адресних рядків у циклі

У подальшому підготовлений цикл можна перенести до вже існуючого завдання або додати інші потрібні блоки, як у раніше наведеному випадку (рис. 38).

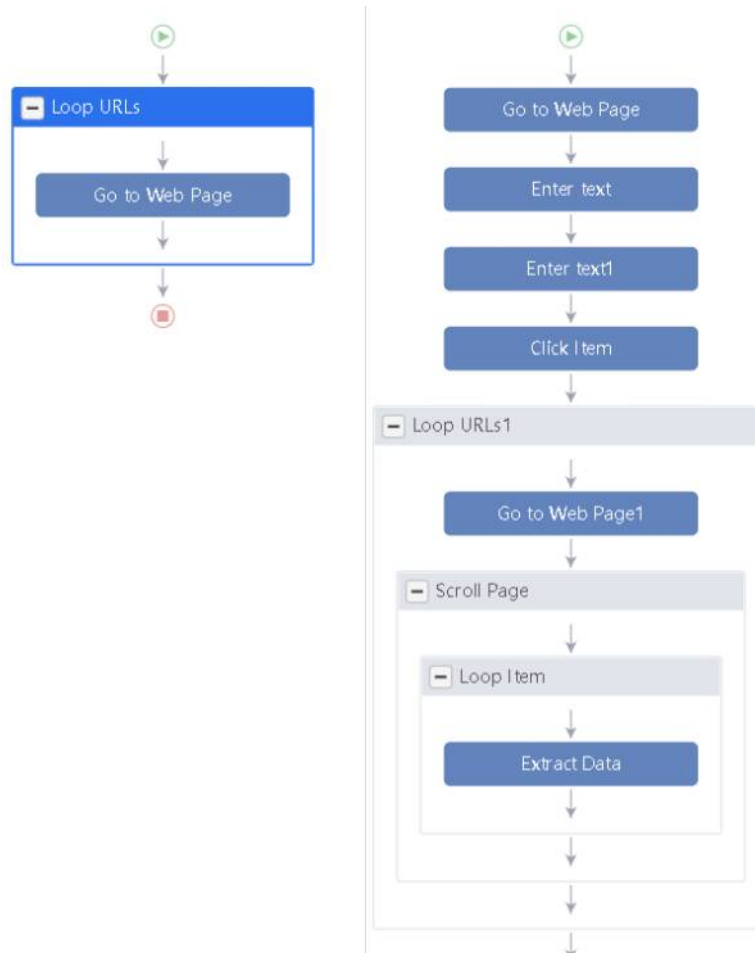


Рис. 38. Формування циклічних послідовностей
За потреби, після виконання послідовності слід видалити дублікати записів (рис. 39).

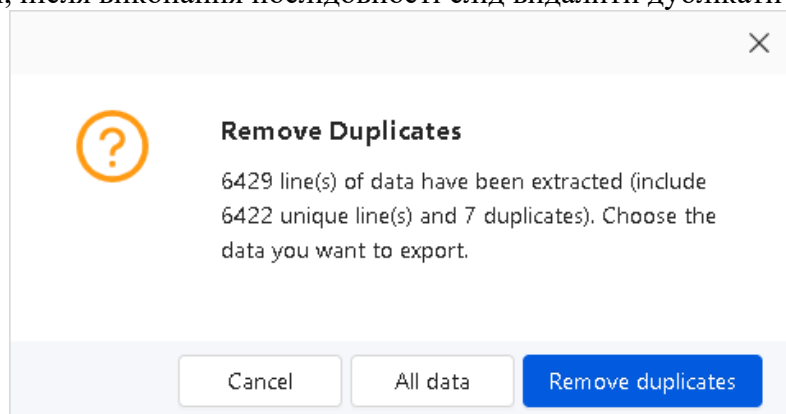


Рис. 39. Налаштування видалення дублікатів


Після завантаження усіх потрібних даних, їх слід належним чином обробити, для того щоб імпортувати до програмного забезпечення, призначеного для аналізу.

Обробка даних

У якості середовища для аналізу може бути використана платформа Gephi. Враховуючи це, потрібно обробити та представити накопичені дані в форматі, придатному для імпорту до згаданого середовища.

Спершу, у сформованому файлі слід видалити інформацію, яка могла туди потрапити помилково. Як правило, такі дані розташовуються наприкінці відомостей про друзів конкретної особи.

Після проведення попереднього очищення для зручності потрібно призначити унікальний ідентифікатор кожному користувачу. Це можна зробити, наприклад, так:

1) скопіювати на окремий аркуш документу Excel посилання на завантажені профілі і видалити дублікати з використанням кнопки  в меню «Дані». Після цього потрібно пронумерувати записи, що залишились, це і буде унікальним числовим ідентифікатором;

2) за допомогою формули

=ИНДЕКС(Лист2!\$B\$2:\$B\$5604;ПОИСКПОЗ(ИСТИНА;ИНДЕКС(Лист2!\$A\$2:\$A\$5604=\$D2;0);0))

сформуати колонку ID на головному аркуші таблиці (рис. 40).

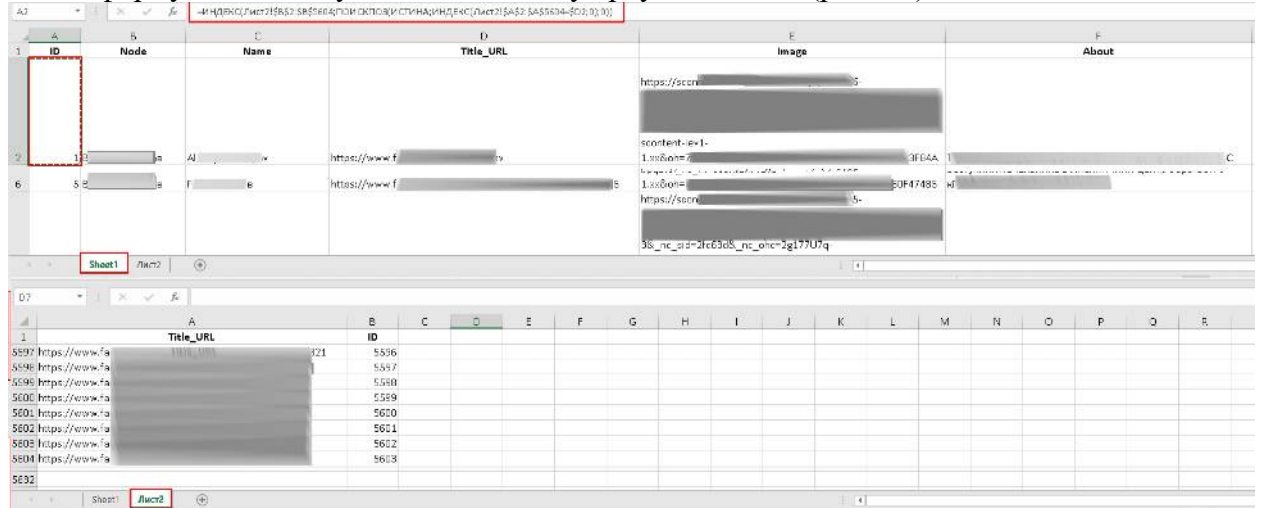


Рис. 40. Налаштування видалення дублікатів

3) перевести формули у сформованій колонці у значення шляхом копіювання самої колонки та спеціальної вставки тільки значень у те саме поле таблиці.

Далі, за потреби, можна в окрему папку завантажити зображення для кожного профілю. Одним із варіантів такого завантаження може бути використання макросу (наприклад, як за цим посиланням excelvba.ru/code/DownloadPictures2).

Для того, щоб цей макрос запрацював у 64-розрядних системах слід змінити його заголовок так:

```
#If VBA7 Then
Private Declare PtrSafe Sub Sleep Lib "kernel32" (ByVal ms As LongPtr)
#Else
Private Declare Sub Sleep Lib "kernel32" (ByVal ms As Long)
#End If
Private Declare PtrSafe Function URLDownloadToFile Lib "urlmon" Alias
"URLDownloadToFileA" _
    (ByVal pCaller As Long, ByVal szURL As String, _
    ByVal szFileName As String, ByVal dwReserved As Long, _
    ByVal lpfnCB As Long) As Long
```

Після формування папки із зображеннями вузлів майбутнього графу потрібно створити два підсумкових аркуші із вершинами та ребрами графу. При цьому в аркуші із вершинами слід передбачити колонку Images із назвами файлів зображень для відображення в якості вузлів графу. Вигляд підсумкових таблиць може бути як на рис. 41.

	A	B	C	D	E	F	G	H
1	Source	Target						
271		275	270					
272		275	271					
273		19	272					
274		19	273					
275		19	274					
276		19	275					
277		19	276					
278		19	277					
279		165	278					

Готово	All	Nodes	Edges	+
--------	-----	-------	--------------	---

D2

=СЦЕПИТЬ(\$A2,".jpg")

	A	B	C	D	E	F
1	Id	Label	Image	Image_formula		
2	1	A v	1.jpg	1.jpg		
3	2	E B	2.jpg	2.jpg		
4	3	D K	3.jpg	3.jpg		
5	4	V iv	4.jpg	4.jpg		
6	5	P бв	5.jpg	5.jpg		
7	6	T r	6.jpg	6.jpg		
8	7	K et	7.jpg	7.jpg		

	All	Nodes	Edges	+
--	-----	--------------	-------	---

Рис. 41. Зміст таблиць Nodes та Edges

Підготовлені таблиці можуть бути імпортовані до середовища Gephi.

Аналіз даних

Для того, щоб використовувати додаткові техніки аналізу та виводити зображення в якості вершин графу у Gephi можна встановити відповідні модулі (рис. 42).

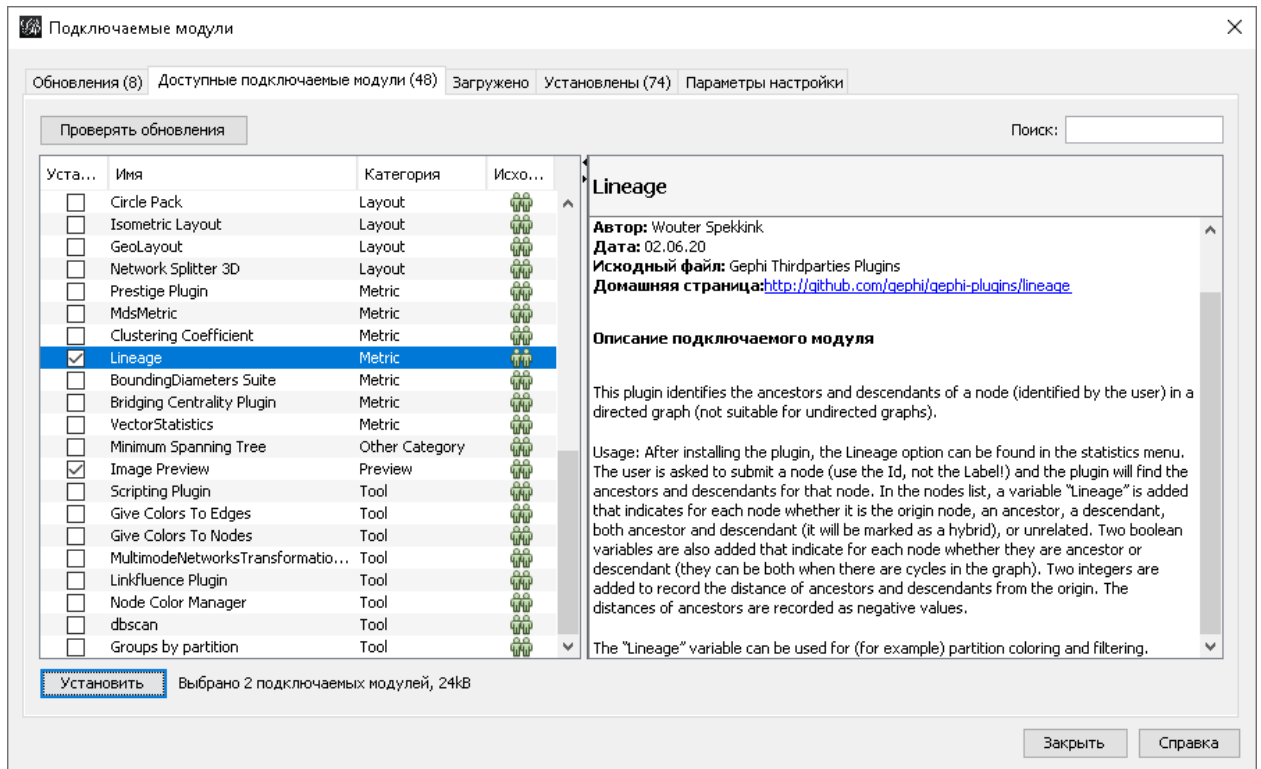


Рис. 42. Встановлення додаткових модулів

Після завантаження підготовлених даних у Gephi, можна проводити їх аналіз за різними критеріями та виводити підсумкові результати у вигляді зображення (рис. 43).

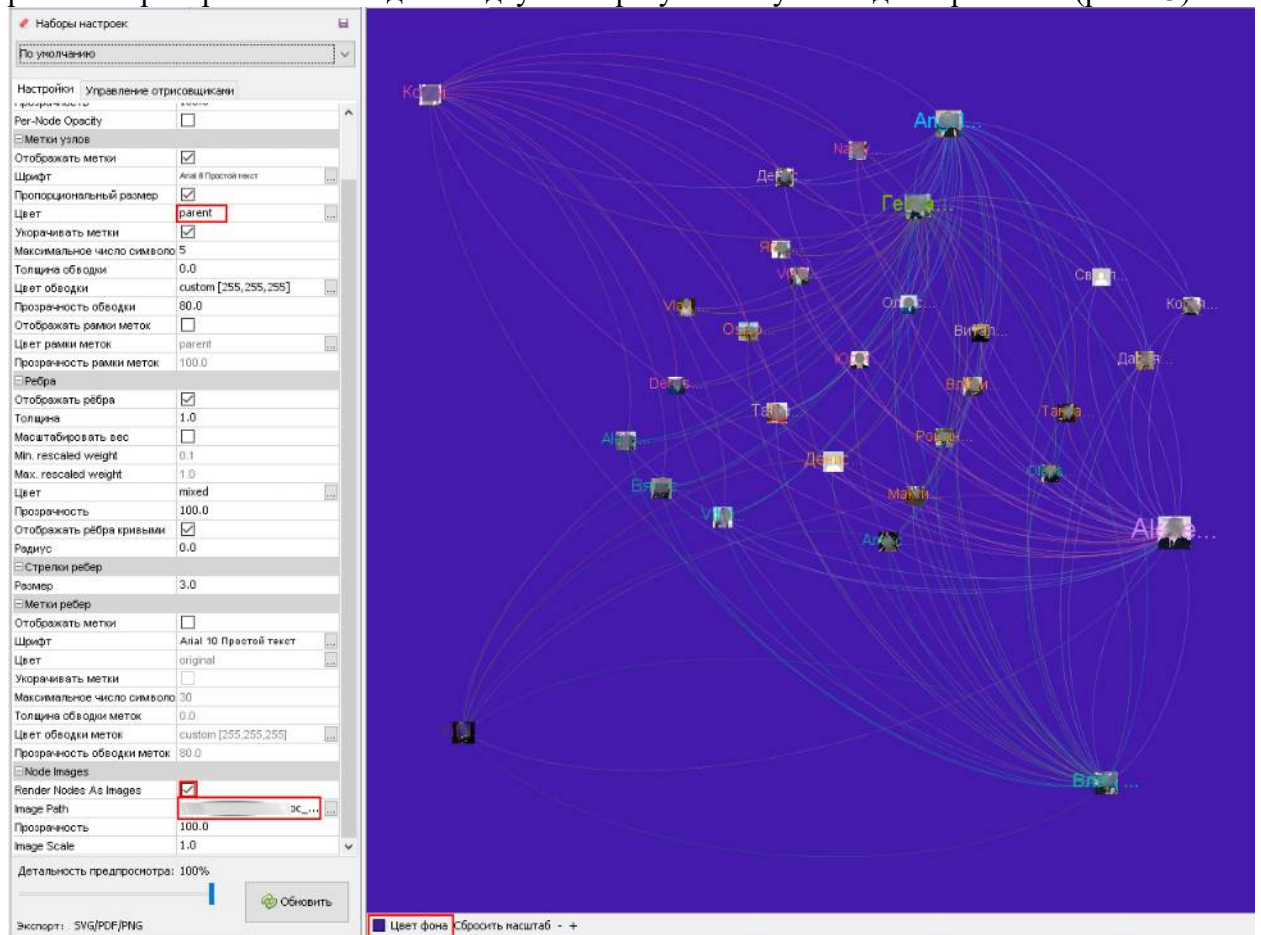


Рис. 43. Підготовка до збереження візуальних результатів аналізу

Аналогічним чином можна здійснювати накопичення, обробку та мережний аналіз даних із інших електронних ресурсів.

5. Аналіз локальних таблиць з даними про рух фінансових цінностей

У роботі правоохоронних органів часто виникають завдання обробки та аналізу великих масивів даних, що надходять від різних фінансових установ, та містять інформацію про рух цінностей осіб, які становлять інтерес. Інколи результати аналізу таких даних виводять слідчого, аналітика або оперативного працівника на ключових фігурантів розслідування та дозволяють зібрати важливі докази їх протиправної діяльності. Крім наведеного, такий аналіз є вкрай корисним при роботі над відшкодуванням збитків, завданих протиправною діяльністю.

У якості інструментарію для роботи з наведеними даними у більшості випадків достатньо використовувати табличний процесор та спеціальний застосунок візуалізації та мережного аналізу накопичених відомостей. У якості прикладу наведемо невеличкий фрагмент аналітичної роботи над файлами, що містять інформацію про рух коштів на банківських рахунках декількох осіб.

У даному випадку маємо низку навчальних таблиць, що імітують відповідь банківської установи правоохоронним органам про рух коштів по банківських платіжних картках восьми осіб за період 1,5 року та містять такі поля:

- | | | |
|----------------------|--------------------|-------------------|
| - основна карта; | - термінал; | - карта / рахунок |
| - поточна карта; | - код авторизації; | одержувача; |
| - дата транзакції; | - ID транзакції; | - ПІН / ОКПО |
| - DB/CR; | - карта / рахунок | отримувача; |
| - сума валюти карти; | відправника; | - ПІБ одержувача; |
| - комісія; | - стан рахунку; | - тип транзакції. |
| - код валюти; | - ПІН / ЄДРПОУ | |
| - баланс; | відправника; | |
| - MCC; | - ПІБ відправника; | |

Обрання відповідних інструментів, методів і технік аналізу буде залежати від завдань, які потрібно вирішити. У даному випадку потрібно визначити ключових осіб, які найбільше пов'язані між собою за кількістю переказів, з метою встановлення кола осіб не охоплених увагою на попередніх етапах розслідування.

Для цього потрібно спершу переконаватися, що вихідні дані не містять дублювання. Воно може проявлятися у наявності файлів, що містять інформацію по різних рахунках, проте по однакових картках, однакових транзакціях в різних картках. Наведене може бути випадково враховано декілька разів в одному протоколі або аналітичній довідці, тощо.

Після виконання попередньої перевірки слід за потреби видалити з файлів даних записи, які не становлять інтересу. Це можуть бути, наприклад, відомості про комісійні виплати по операціях, поточні покупки тощо. Вказані дані можна видалити і на етапі візуалізації руху коштів, проте у відповідних застосунках це зробити, як правило, складніше, аніж у табличних процесорах, наприклад, із застосуванням фільтрів.

Підготовлені таблиці з даними потрібно імпортувати в середовище візуалізації та аналізу. У якості прикладу такого середовища можна згадати систему IBM i2 Analyst's Notebook.

З метою завантаження даних у вказане середовище слід обрати відповідні опції імпорту (рис. 45) та провести низку налаштувань специфікації, яку можна буде у подальшому застосовувати до інших подібних файлів з даними.

Выберите рабочий лист | Выделить строки | Действия со столбцами | Выбрать структуру | Назначить столбцы | Подроб

На этой странице можно исключить строки заголовка и другие помеченные строки данных.

Включить выбранную строку (строки) | Исключить выбранную строку (строки) | Задать строку заголовка

Строка	Основ...	Поточ...	Дата т...	DB/CR	Сума ...	Сума ...	Комісія	Код В...	Баланс	MCC	Термін...	Код а...	ID тра...	Карта ...
1	Виплат...													
2	Основ...	Поточ...	Дата т...	DB/CR	Сума ...	Сума ...	Комісія	Код В...	Баланс	MCC	Термін...	Код а...	ID тра...	Карта ...
3	877799...	777799...	2020-0...	DB	204	204	0	980	1267,86	4814	ZZ2407...	112442	234567...	777799...
4	877799...	777799...	2020-0...	DB	102	102	0	980	1165,86	4814	ZZ2407...	112443	234567...	777799...
5	877799...	777799...	2020-0...	DB	204	204	0	980	961,86	4814	ZZ2407...	112444	234567...	777799...
6	877799...	777799...	2020-0...	CR	794,01	798	3,99	980	1755,87	6010	ZZ2408...	112445	234567...	777799...
7	877799...	777799...	2020-0...	CR	794,01	798	3,99	980	1755,87	6010	ZZ2408...	112445	234567...	444444...
8	877799...	777799...	2020-0...	DB	200	200	0	980	1555,87	4829	ZZ2232...	112446	234567...	777799...
9	877799...	777799...	2020-0...	DB	113	113	0	980	1442,87	4814	ZZ2407...	112447	234567...	777799...
10	877799...	777799...	2020-0...	DB	200	200	0	980	1242,87	4829	ZZ2232...	112448	234567...	777799...
11	877799...	777799...	2020-0...	DB	100	100	0	980	1142,87	4829	ZZ2232...	112449	234567...	777799...
12	877799...	777799...	2020-0...	CR	1170	1175,88	5,88	980	2312,87	6010	ZZ2409...	112450	234567...	444444...
13	877799...	777799...	2020-0...	CR	1170	1175,88	5,88	980	2312,87	6010	ZZ2409...	112450	234567...	777799...
14	877799...	777799...	2020-0...	DB	2300	2300	0	980	12,87	6011	ZZ2410...	112451	234567...	777799...
15	877799...	777799...	2020-0...	CR	349,74	2300	1,76	980	362,61	6536	ZZ2361...	112452	234567...	777799...
16	877799...	777799...	2020-0...	CR	349,74	350	1,76	980	362,61	6536	ZZ2361...	112452	234567...	444444...
17	877799...	777799...	2020-0...	DB	350	350	0	980	12,61	4829	ZZ2232...	112453	234567...	777799...
18	877799...	777799...	2020-0...	CR	995	1000	5	980	1007,61	6536	'	112454	234567...	777799...
19	877799...	777799...	2020-0...	CR	995	1000	5	980	1007,61	6536	'	112454	234567...	444444...
20	877799...	777799...	2020-0...	DB	150	150	0	980	857,61	4829	ZZ2232...	112455	234567...	777799...
21	877799...	777799...	2020-0...	DB	143,21	143,21	0	980	714,4	4829	ZZ2232...	112456	234567...	777799...
22	877799...	777799...	2020-0...	DB	200	200	0	980	514,4	6011	ZZ2411...	112457	234567...	777799...
23	877799...	777799...	2020-0...	DB	101,57	101,57	0	980	412,83	5411	ZZ2412...	112458	234567...	777799...
24	877799...	777799...	2020-0...	DB	184,9	184,9	0	980	777,93	5411	ZZ2413...	112459	234567...	777799...

Если какая-либо строка содержит заголовки столбцов данных, введите номер этой строки ниже.

☒ Извлечь заголовки столбцов из строки: 2

Если строки, содержащие комментарии или данные, которые следует игнорировать, помечены специальным символом, введите этот символ ниже.

☐ Игнорировать строки, начинающиеся с: %

Сохранить | Закрыть | < Назад | Далее > | Импорт | Справка

Рис. 45. Визначення заголовків

та налаштування неврахування окремих рядків

У разі потреби в незначному редагуванні імпортованих даних можна використати спеціальні інструменти як на рис. 46.

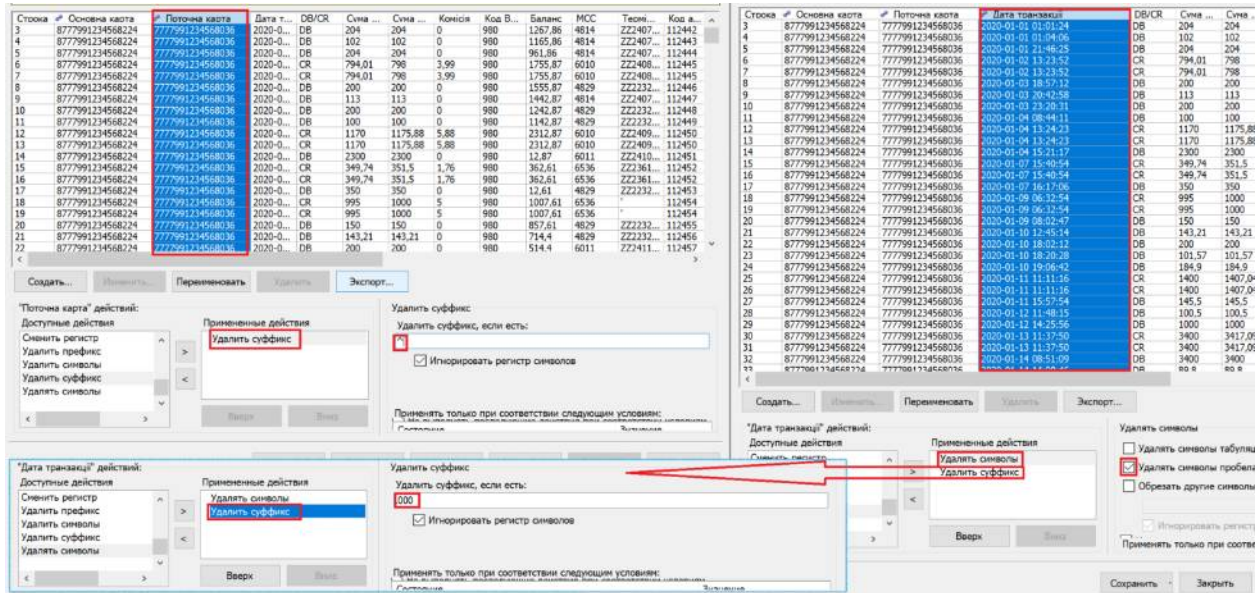


Рис. 46. Попереднє редагування даних для імпорту

Після проведення описаних процедур потрібно обрати вигляд схеми для виведення та налаштувати її параметри (рис. 47).

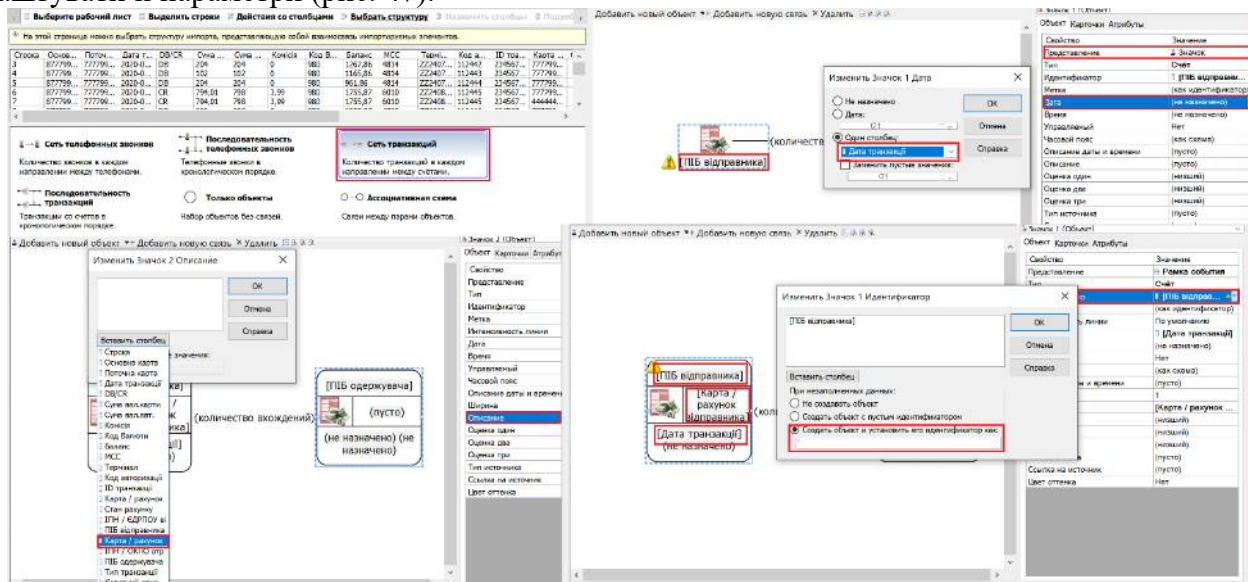


Рис. 47. Налаштування схеми з даними

Результуючу специфікацію потрібно зберегти та за її допомогою додати до схеми усі потрібні таблиці з даними. Надалі імпортовані дані можна упорядкувати та розрахувати для них деякі показники в рамках проведення мережного аналізу (рис. 48).

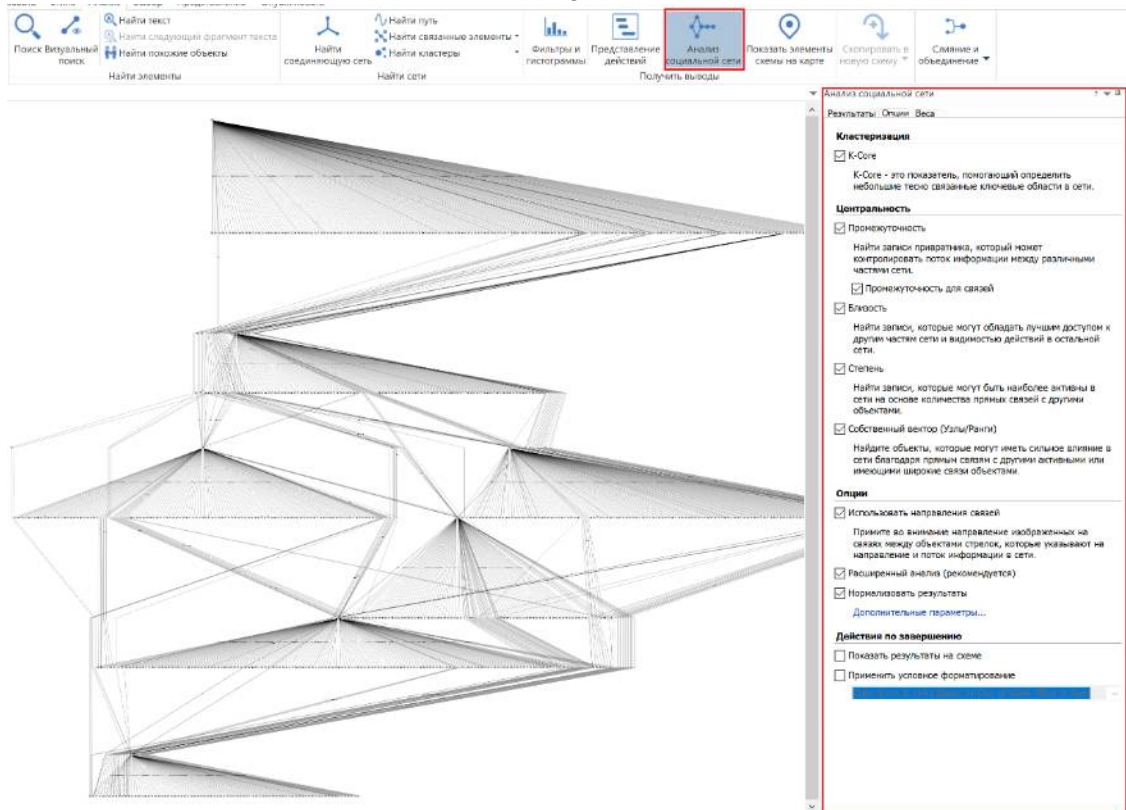


Рис. 48. Здійснення попереднього мережного аналізу організованих даних

Наведена схема не може зручно сприйматися як аналітиком, так і кінцевим споживачем аналітичного продукту. Враховуючи це, потрібно видалити нерелевантні дані зі схеми або об'єднати їх в окремі блоки. Для кращого сприйняття відомостей бажано прив'язати до відповідних вузлів зображення осіб та провести кластеризацію нової схеми (рис. 49).

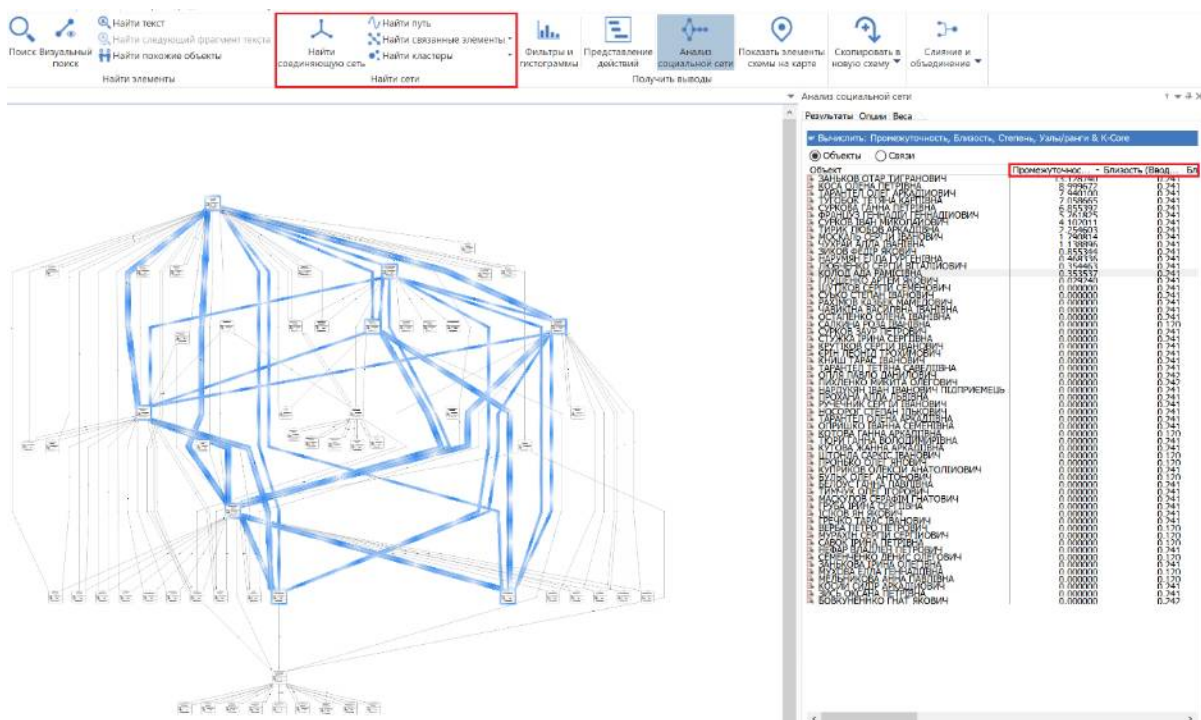


Рис. 49. Проведення кластеризації зменшеної кількості вузлів

У подальшому схема може удосконалюватися відповідно до поставлених завдань, а також використовуватися під час підготовки аналітичних висновків.

Завдання

1. Відпрацюйте роботу програми Hunchly. Сформуйте звіт та синхронізуйте результати в Google-таблиці. Виконуйте це завдання в парі.
2. Ознайомтесь з роботою програми Kuiper, підготуйте звіт.
3. Вище було описано приклад аналізу даних з груп Telegram. Подібним чином можна проводити аналіз каналів Telegram, де зв'язки описуються пересланими повідомленнями з одного каналу до іншого або відповідними згадуваннями. Оскільки дані каналів доступні для попереднього огляду через веб-доступ, можна без авторизації в Telegram дізнатися відомості про кількість учасників каналу, повідомлення в ньому, час їх розміщення тощо. Враховуючи викладене, ознайомтесь із статтею «Telegram Network Visualization — Tracing Forwards and Mentions» (<https://medium.com/dataseries/telegram-network-visualization-tracing-forwards-and-mentions-f75746712fcf>) та проаналізуйте кілька Telegram каналів за завданням викладача.
4. За допомогою ресурсу <https://combot.org/telegram/top/groups> обрати декілька груп з відповідною мовою за завданням викладача. Проаналізувати їх. Підготувати аналітичний висновок.
5. За визначеним критерієм сформуйте таблицю з даними про сайти, які становлять інтерес. Спробуйте проаналізувати підготовлені дані з використанням онлайн-сервісів та встановленого на комп'ютері програмного забезпечення. Порівняйте одержані результати.
6. Завантажте та проаналізуйте інформацію про друзів (підписників) 10 осіб в одній із соціальних мереж. Під час виконання завдання спробуйте також накопичити відомості про кількість друзів (підписників) для кожної особи та розмістити їх в окремому полі таблиці з назвою Size. Проведіть аналіз із урахуванням поля Size.
7. Ознайомтесь із матеріалом за адресою: <https://infosecwriteups.com/data-exploration-and-visualization-on-leaked-clubhouse-data-25408b03664>. Спробуйте провести аналогічну обробку та візуалізацію даних соціальної мережі Clubhouse.
8. Ознайомтесь з утилітою TikTok Scraper (github.com/drawrowfly/tiktok-scraper). Завантажте відомості про 10 користувачів сервісу TikTok. Побудуйте граф та проведіть відповідний мережний аналіз набутих даних.
9. Для наданих таблиць даних визначити осіб, які системно здійснювали трансакції між рахунками одна одної, із виведенням кількості переказів на ребрах графу. Побудувати граф. Здійснити мережний аналіз завантажених даних за допомогою застосунку.
10. Побудувати схему, де вузлами графу будуть відомості про осіб, які переказали сукупно більше 1000 грн. Розрахувати мережні показники та визначити кластери. На ребрах графу відобразити загальну суму переказів за відповідним напрямком.
11. Порівняти одержані схеми та підготувати фрагмент аналітичного висновку за результатами опрацювання вказаних даних з урахуванням системи 5W+H.

Практичне заняття. Картографування об'єктів за допомогою різних програмних інструментів

Навчальна мета заняття: відпрацювати різні технології картографування.

Час проведення 2 год. Місце проведення: комп'ютерний клас.
(кількість годин) (полігон, комп'ютерний клас, лабораторія, лінгафонний кабінет)

Устаткування: персональний комп'ютер (ПК) зі встановленою операційною системою Windows 2000 або вище та доступом до мережі Інтернет.

У сучасних версіях Excel є декілька інструментів, які дозволяють працювати з геоданими та візуалізувати їх на карті.

Одним із способів перенесення даних на карту є використання звітів Power View. Для активації цієї надбудови слід активувати відповідні вкладки (Параметри → Налаштувати стрічку). Потім у меню розробника потрібно натиснути кнопку Надбудови COM та встановити позначку навпроти Microsoft Power View for Excel. Після цього у меню Вставка з'явиться кнопка Power View. Якщо кнопка не з'явилася, то додатково слід встановити SilverLight та створити нову налаштовану групу в стрічці й додати туди з «Усі команди» об'єкт Power View.

Для відображення на карті відповідних даних слід виділити їх у таблиці та натиснути кнопку Power View. У результаті на новому аркуші відкриється виділена таблиця як на рис. 1.

Щелкните здесь, чтобы добавить заголовок

Широт	Довго	Назва	Кільк	result_num	osm_id
50,00	36,23	Дзеркальний струмінь	13	0	702879881
49,96	36,26	Обласний аптечний склад № 335	16	0	140693264
50,03	36,23	Парковка	9	0	277838813
50,01	36,22	Сквер Воїнів-Інтернаціоналістів	8	0	117394631
49,98	36,26	Стадіон «Металіст»	6	0	148932854
50,01	36,24	Стадіон «Піонер»	8	0	155114400
49,97	36,28	Харків - Балашовський	12	0	177289293
49,98	36,24	Харків Левада	7	0	5445256349
49,93	36,27	Харківський національний університет внутрішніх справ	6	0	135268798
50,01	36,23	Харківський національний університет ім. Каразіна	7	0	10383471
499,89	362,48				

Рис. 1. Перенесення даних на вкладку Power View

Надалі у меню Конструювання слід натиснути Карта, в результаті чого можна налаштувати відповідне відображення на полотні занесених до таблиці геоданих (рис. 2).

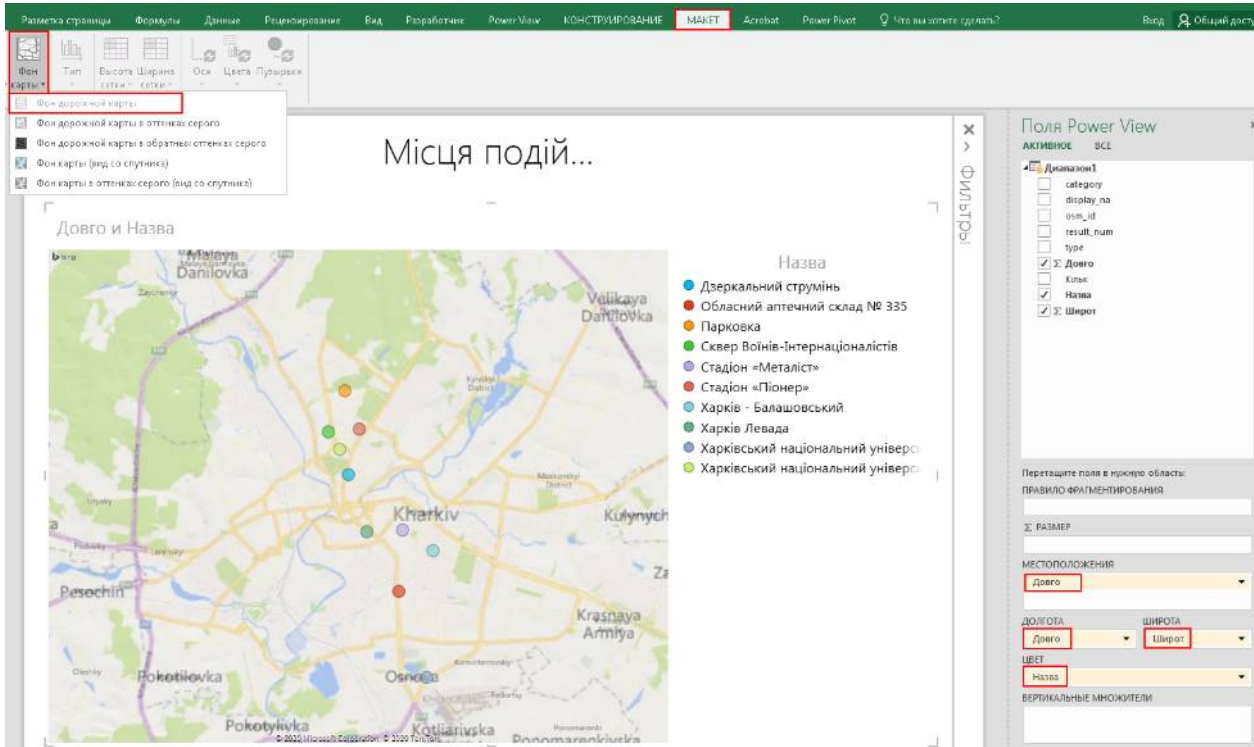


Рис. 2. Налаштування відображення геоданих

У вкладці фільтри за потреби можна налаштувати, які саме дані слід відображати з таблиці.

Ще один спосіб візуалізації геоданих полягає у використанні надбудови Power Map. Вона розташовується у меню Вставка (3D-карта), а якщо вона там відсутня слід вручну додати її по аналогії з надбудовою Power View.

Після виділення відповідних даних в таблиці та натискання кнопки 3D-карта можна досить гнучко налаштувати параметри відображення даних на карті (рис. 3).

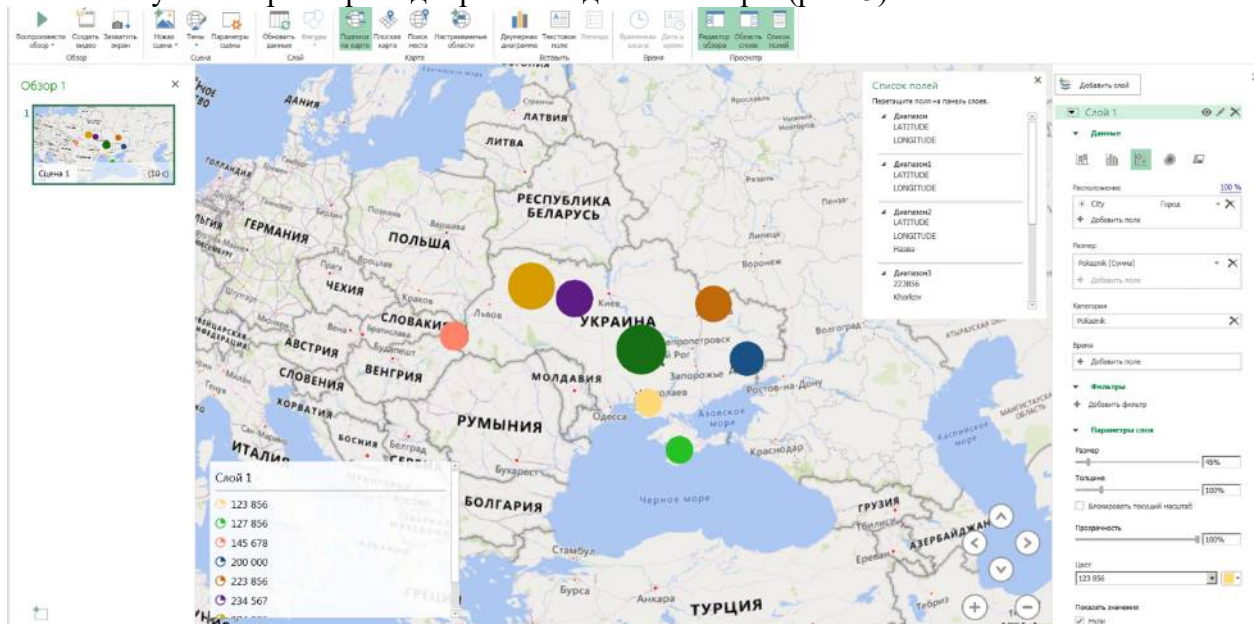


Рис. 3. Робота з Power Map

Корисною особливістю досліджуваної надбудови є можливість побудови динамічних графіків, як от, наприклад зміна показників на певній території з плином часу.

Ще одним інструментом, який дозволяє працювати з інтерактивними мапами, є Google My Maps. Даний сервіс надає можливість інтерактивного позначення об'єктів на карті, а також завантажувати вже підготовлені дані про конкретні місця. Для цього потрібно заздалегідь підготувати таблицю, яка міститиме стовпці з широтою, довготою та назвою конкретних об'єктів, які потрібно позначити на карті. Після цього потрібно авторизуватися в обліковому записі Google та перейти за адресою google.com/maps/d/. Надалі потрібно натиснути кнопку **+ СТВОРИТИ КАРТУ** та імпортувати раніше підготовлені дані (рис. 4).

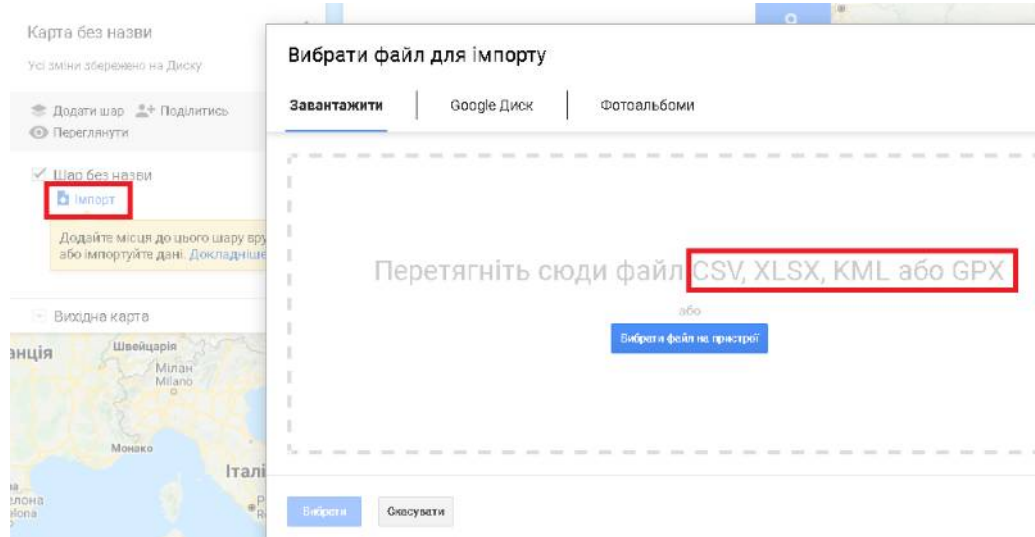


Рис. 4. Імпорт даних для нової карти

Під час здійснення імпорту потрібно вказати сервісу, який саме стовпчик таблиці міститиме які дані (рис. 5).

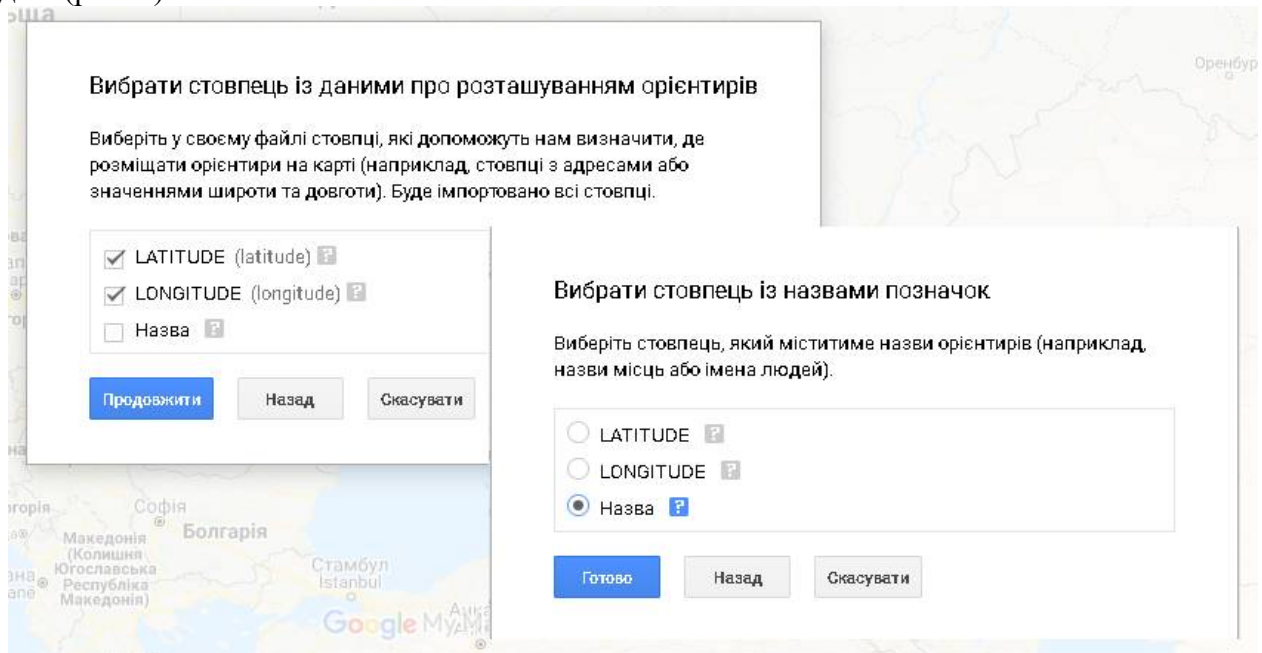


Рис. 5. Зіставлення стовпців з даними та параметрів мапи

За результатами імпорту даних одержимо карту як на рис. 6.

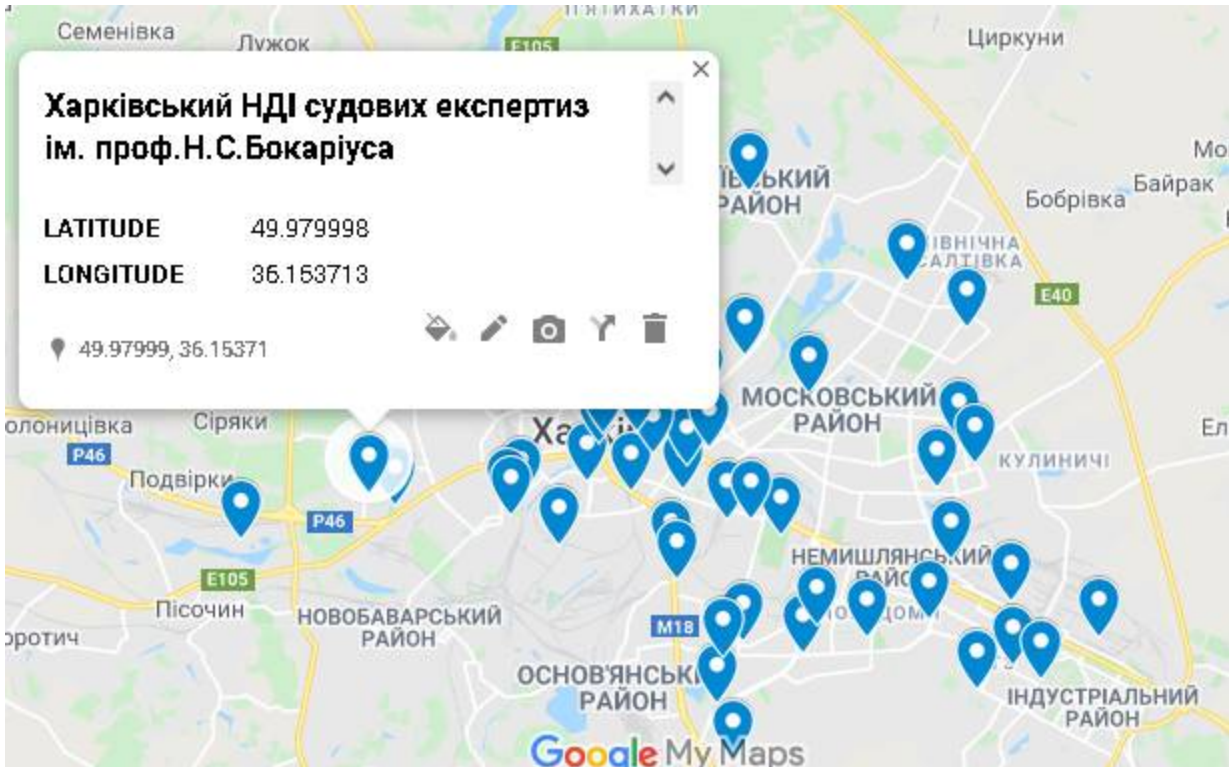


Рис. 6. Зіставлення стовпців з даними та параметрів мапи

У подальшому до створеної карти можна додавати інші дані до вже створеного шару або створювати інші шари.


Завдання

1. На основі статистичних даних з Єдиного реєстру досудових розслідувань сформуєте таблицю, яка міститиме інформацію про кількість зареєстрованих грабежів в областях за період з 2013 по поточний рік:

Рік	Назва обласного центру	Кількість грабежів
-----	------------------------	--------------------

З використанням інструментів Power View, Power Map та Google My Maps перенесіть дані таблиці на карту.

2. Складіть звіт, який в якості висновку має містити порівняння одержаних результатів.

Тепер, коли відповідні вхідні дані збережено в універсальному форматі їх можна імпортувати до системи QGIS. Для цього у меню Шар → Додати шар слід обрати пункт Додати текстовий з роздільниками шар  та у вікні, що з'явилося обрати параметри як на рис. 3.

Назва файлу E:\Video\QGIS\Kharkiv.csv

Ім'я шару Kharkiv Кодування windows-1251

Формат файлу

☐ CSV (значення, розділені комами)
 ☒ Табуляція
 ☐ Двокрапка
 ☐ Пробіл
☐ Роздільник регулярних виразів
 ☐ Крапка з комою
 ☐ Кома
 Інші
☒ Користувальницькі роздільники
 Лапки
 Вихід

Параметри запису та полів

Кількість рядків заголовка для видалення 0
 ☒ Десятковий роздільник є комою
☒ Перший запис має назви полів
 ☐ Обрізати поля
☒ Виявити типів полів
 ☐ Відхилити порожні поля

Визначення геометрії

☒ Координати точок
 Поле X Довгота
 Поле Z
 Поле Y Широта
 M поле
☐ Відомий текст (WKT)
 ☐ DMS координати
☐ Не містить геометрії (атрибут тільки таблиця)
 Система координат геометрії EPSG:4326 - WGS 84

Налаштування шарів

Зразок даних

	Широта	Довгота	Назва об'єкту	Кількість
1	49.921038	36.279044	Міжнародний аеропорт Харків	1

Рис. 3. Імпорт текстових даних

При виконанні імпорту вкрай важливо вірно обрати стандарт представлення координат. Найбільш поширеним стандартом, який застосовується у багатьох картографічних застосунках є WGS 1984, який власне і було обрано, як видно з рисунку.

Після натискання кнопки «Додати» на полотні QGIS з'являться точки, розміщені згідно з заданими в імпортованому текстовому файлі координатами (рис. 4). Як видно із рисунку, точки розташовані просто на білому полотні. Для того, щоб прив'язати їх до конкретної мапи в програмі QGIS, слід встановити необхідні модулі. У даному випадку для цього буде достатньо модуля QuickMapServices, який інсталюється через меню Плагіни → Управління та встановлення плагінів. У вікні, що з'явиться, слід задати пошук QuickMapServices, обрати модуль та натиснути відповідну кнопку установки.

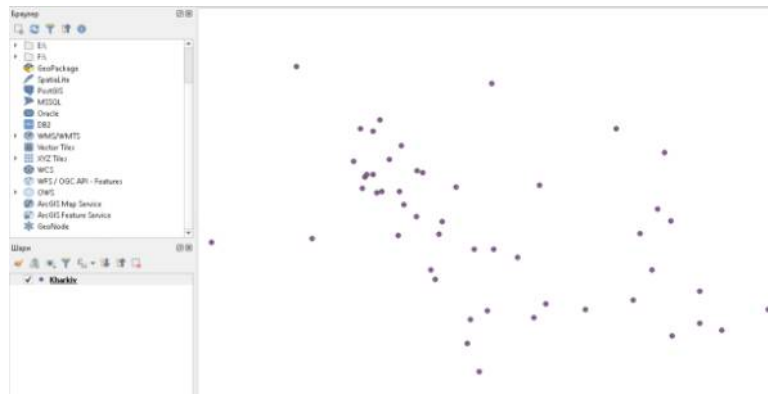


Рис. 4. Точки, розміщені в системі координат

Після встановлення QuickMapServices в панелі інструментів з'явиться кнопка, за допомогою якої можна обрати різні види карт, які додаватимуться в проект як окремі шари (рис. 5).

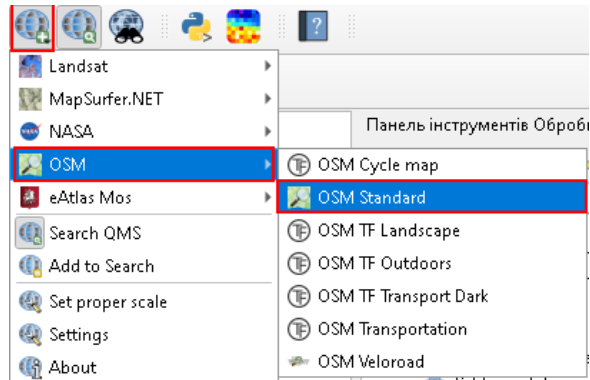


Рис. 5. Додавання нового шару-карти

У результаті додавання нового шару-карти можна наочно побачити візуалізацію певних подій з прив'язкою до карти (рис. 6).

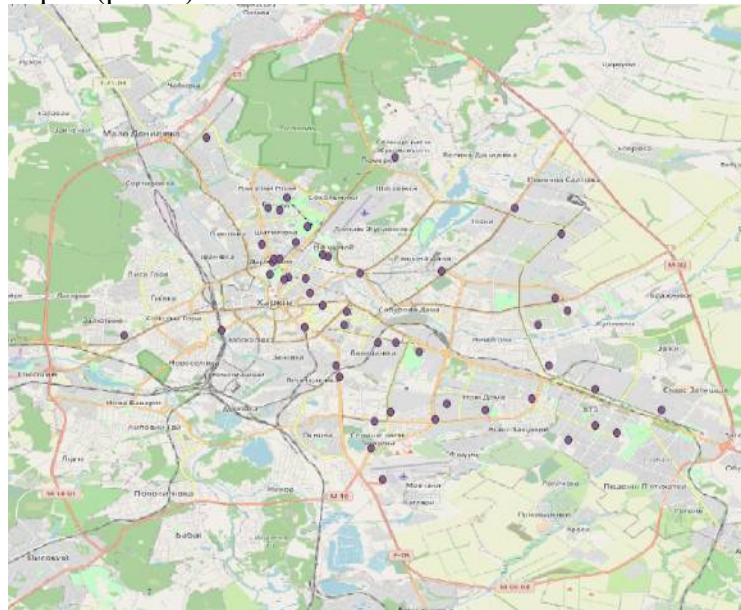


Рис. 6. Точки подій на карті

Із завантаженими даними тепер можна проводити різні розрахунки та аналізувати. Так, у найпростішому варіанті, можна виміряти відстані між точками на карті, які відображають маршрут руху певної особи (рис. 7).

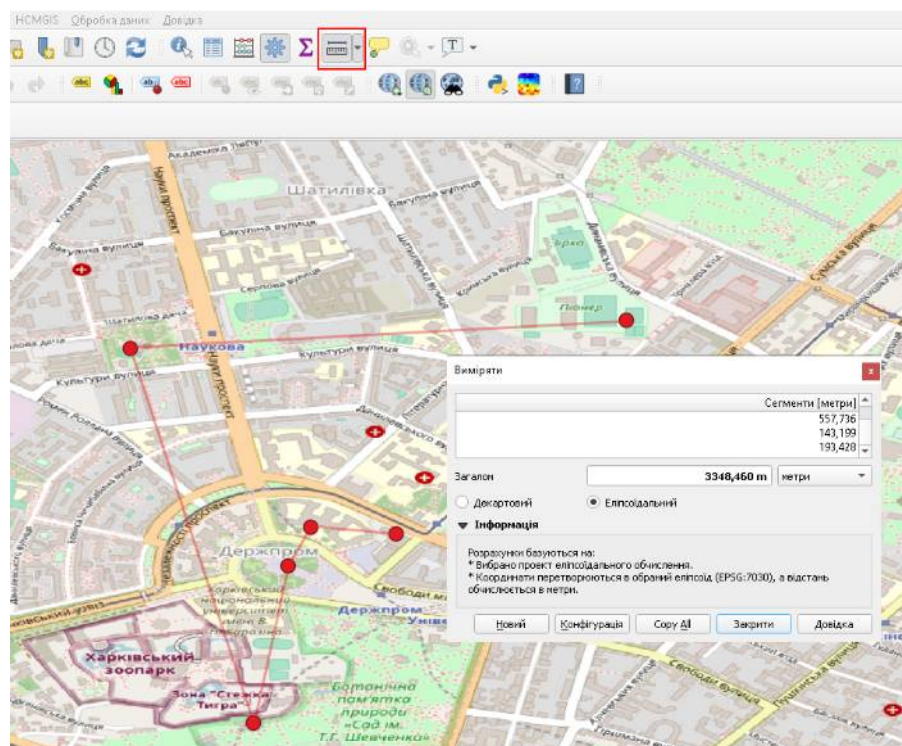


Рис. 7. Вимірювання відстані між точками на карті

З імпортованими даними також можна робити інші операції, наприклад, виводити тільки певні дані за заданими критеріями. Вказану процедуру можна виконати за допомогою Конструктора запитів, доступного у властивостях шару на вкладці Джерело (рис. 8).

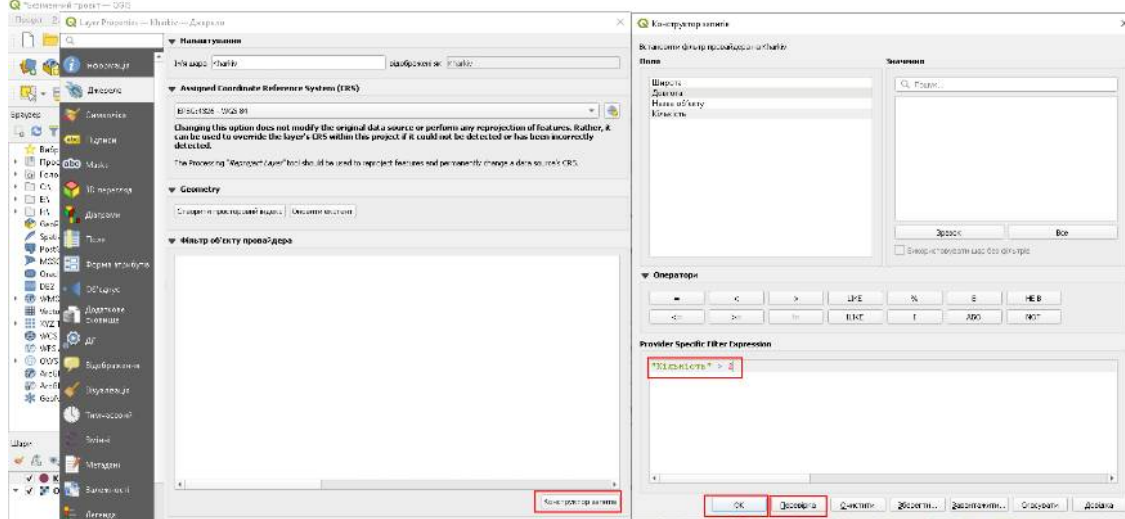


Рис. 8. Запуск конструктора запитів

У випадку застосування фільтра як на рисунку на карту буде виведено тільки ті точки, для яких у стовпчику Кількість вказано число > 2.

Імпортовані дані також можна представити у вигляді теплової мапи. Для цього слід у властивостях шару обрати вкладку Символіка, у якій обрати замість звичайного знаку позицію Теплокарта. Далі можна обрати тип градієнта (YlOrRd), радіус (чим більше, тим більше тепловий контур). Також у налаштуваннях градієнта слід налаштувати непрозорість, для того, щоб один шар не перекривав інший. У результаті можна побачити теплову мапу, за допомогою якої можна, наприклад, відслідковувати найбільш криміногенні осередки (рис. 9).

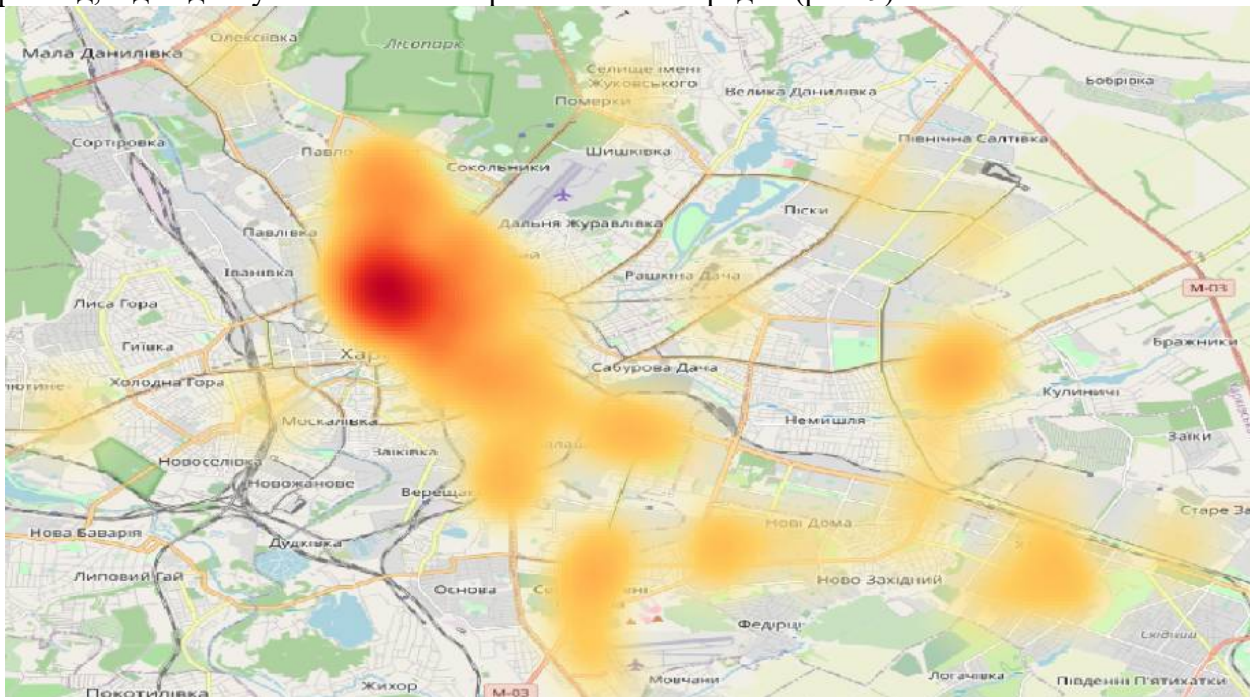


Рис. 9. Теплова карта

Якщо потрібно створити більш інформативну картину візуалізації, то у властивостях відповідного шару можна обрати вкладку Діаграма, у якій обрати вид Текстова діаграма. Далі слід обрати доступні атрибути з імпортованих раніше текстових даних (рис. 10).

Рис. 10. Деталізоване відображення подій

Описані вище процедури стосуються ситуації, коли відомі координати певних об'єктів. Разом з тим у поліцейській практиці частіше трапляються ситуації, коли наявні великі масиви інформації, де в якості просторового орієнтиру використовуються адреси об'єктів. Розглянемо, яким чином можна працювати в програмі QGIS з визначеними адресами.

У якості прикладу можна взяти довільний перелік організацій (<https://raw.githubusercontent.com/pssguy/fortune500/master/fortune500.csv>), де в якості місця розміщення вказано просто назву міста. Для побудови геокоординат за відомими адресами в програмі QGIS можна використати модуль MMQGIS (рис. 11).

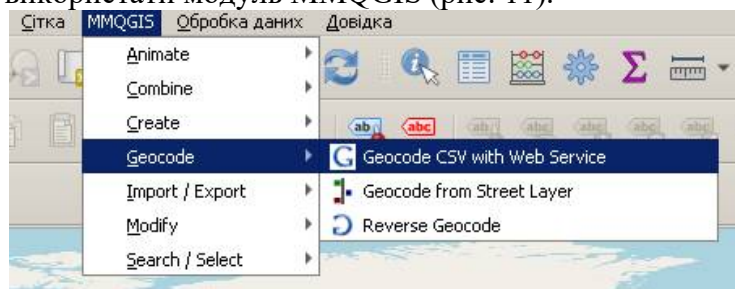


Рис. 11. Меню модуля MMQGIS

У подальшому слід обрати параметри як на рис. 12.

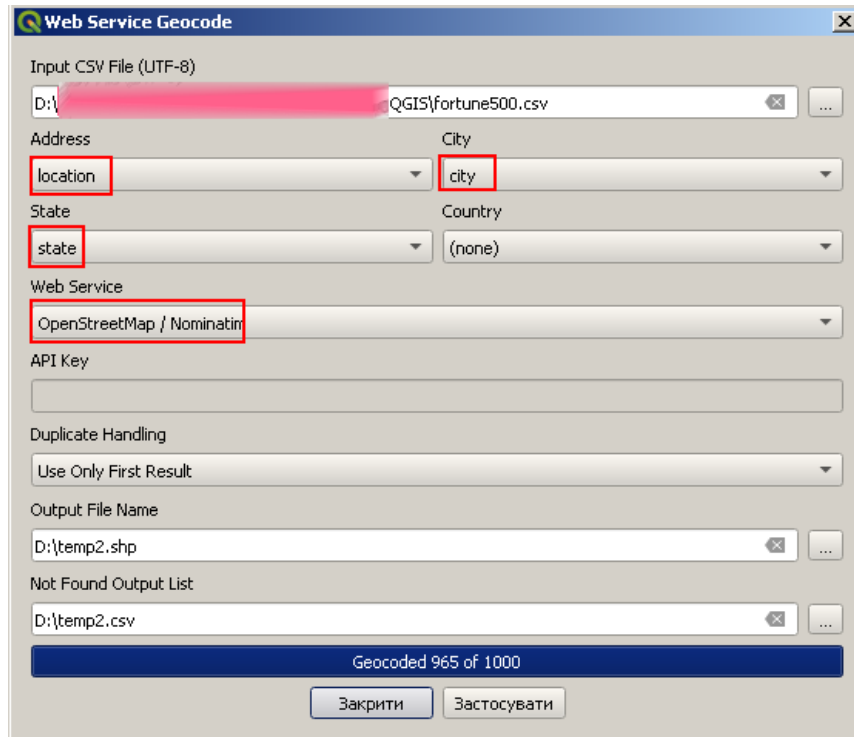


Рис. 12. Параметри імпорту

У результаті одержимо точки розміщення відповідних адрес на карті (рис. 13).

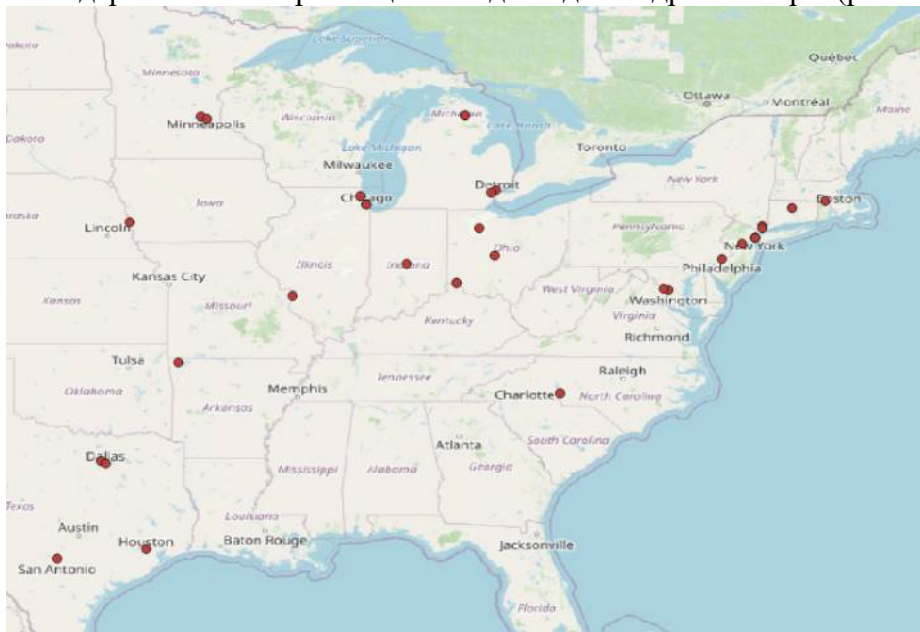


Рис. 13. Параметри імпорту

Так само за допомогою інструмента Reverse Geocode модуля MMQGIS можна здійснити зворотний пошук адрес за координатами (рис. 14).

Об'єкт	Значення
temp3	
osm_id	177289293
(Вив...	
(Дії)	
Широт	49,974209000000002
Довго	36,276784999999997
Назва	Харків - Балашовський
Кільк	12
resul...	0
osm_id	177289293
displ...	Вокзал "Балашівка", 114, Балашовський проїзд, Завод ім. Малишева...
cate...	building
type	yes
latlong	49.973997749999995,36.27672188265159

Рис. 14. Параметри імпорту

Для збереження новоутворених та відредагованих шарів у структурованому вигляді слід натиснути правою кнопкою миші на назві шару та обрати Експорт Зберегти об'єкти як... У вікні, що з'явиться слід обрати відповідний формат експорту даних, наприклад, CSV.

Для визначення точок на карті за адресами можна використовувати модуль GeoCoding, після встановлення якого відповідний інструмент з'являється в меню Плагіни (рис. 15).

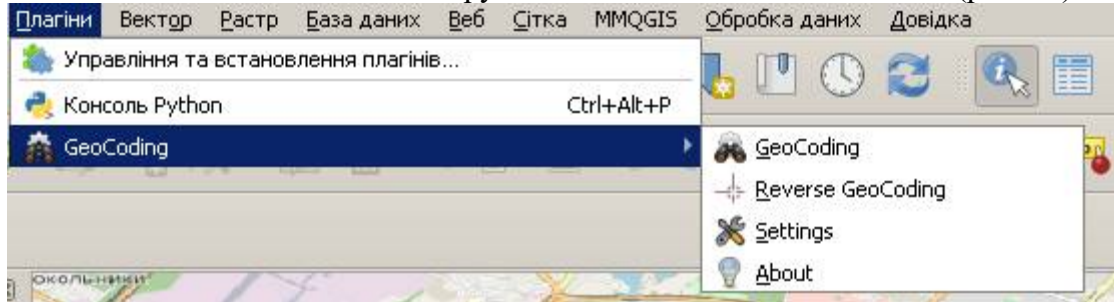


Рис. 15. Меню розміщення точок на карті за адресами

Більш докладно про роботу з різними модулями QGIS можна дізнатися, наприклад, з guides.library.ucsc.edu/DS/Resources/QGIS.

Крім виконання базових завдань з картографування певних подій, у програмі QGIS можна працювати з даними аерофотозйомки, тобто зі знімками високої роздільної здатності. Це є дуже важливим з точки зору роботи сучасних систем безпечного міста.

Для того, щоб навчитися працювати з зовнішніми наборами даних (dataset), можна використати спеціалізовану базу просторових даних GeoSeer. Ця база містить багато посилань на різні електронні мапи, у тому числі результати аерофотозйомки. У якості ключових слід для пошуку наборів даних можуть бути використані назви країн або міст.

Що стосується України, то на даний момент набори даних з результатами аерофотозйомки знайти вкрай складно, натомість присутні кадастрові відомості, розташування будівель, транспортних мереж тощо (geoseer.net/rl.php?ql=5a44cdaf8765c964&p=2&q=ukraine#, geoseer.net/rl.php?ql=87d88cdab250695d&p=1&q=ukraine#).

Для підключення до серверу з наборами геоданих слід скопіювати відповідне посилання на сервер, відкрити диспетчер джерел даних, обрати вкладку WMS/WMTS створити новий шар із посиланням на ресурс та натиснути кнопку підключити (рис. 16). Після появи наборів даних, слід обрати ті, які потрібні, після чого натисканням кнопки Додати завантажити їх у новий шар.

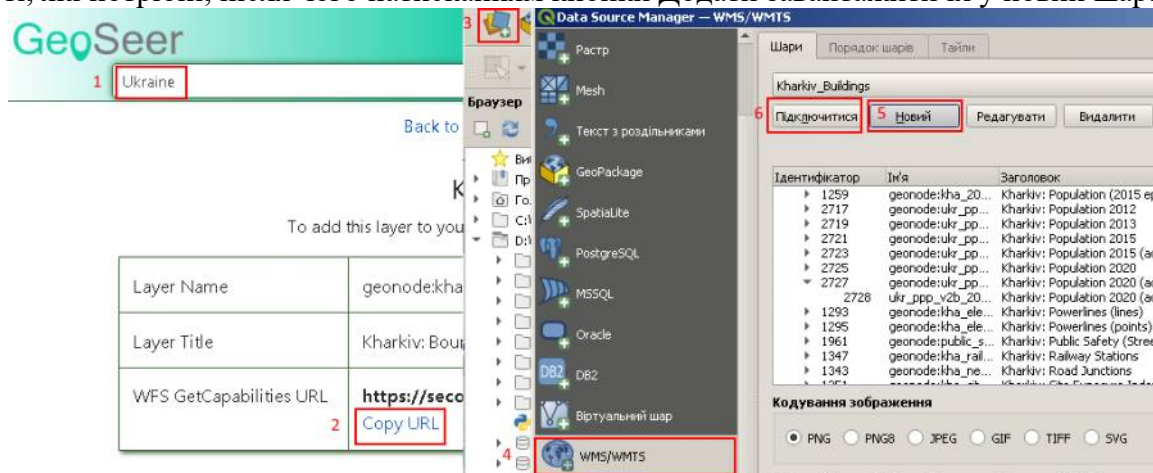


Рис. 16. Підключення зовнішніх наборів геоданих

На рис. 17-18 показано результати завантаження різних наборів даних до QGIS.

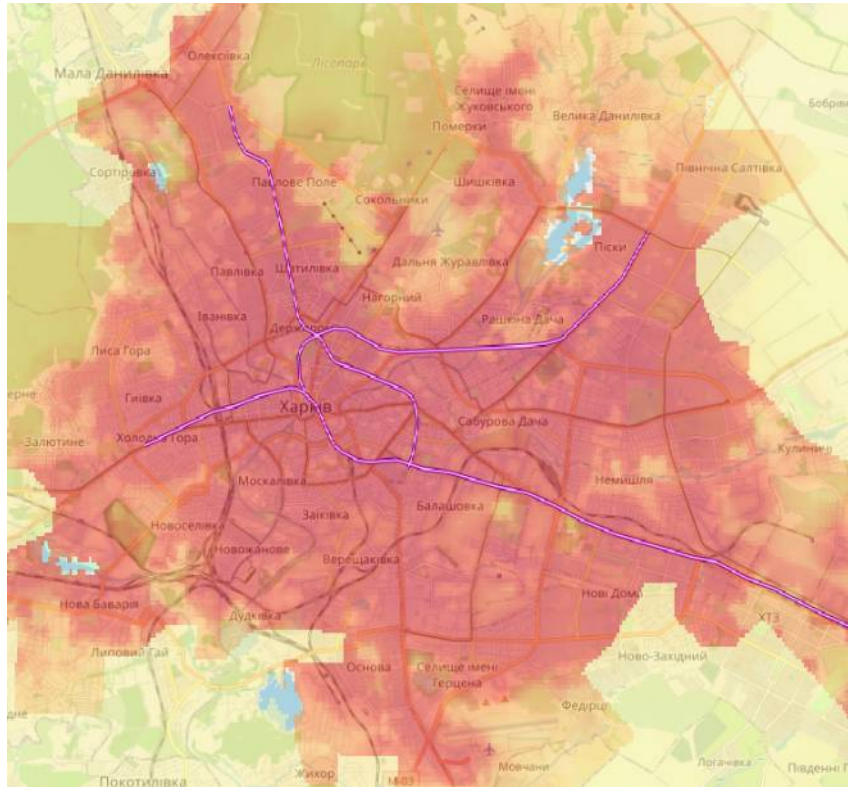


Рис. 17. Теплова мапа, яка відображає густину населення

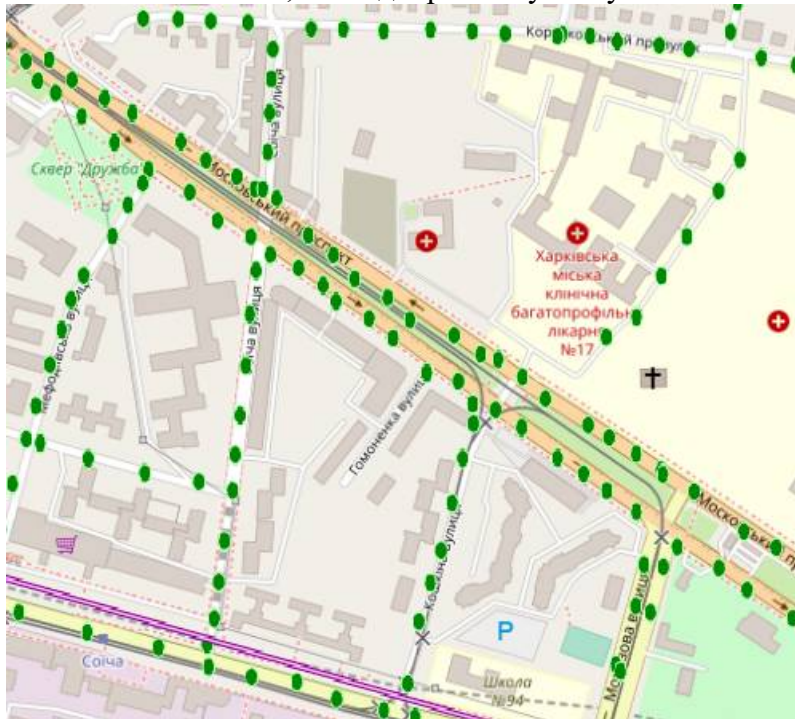


Рис. 18. Відомості про розташування світлофорів



Рис. 19. Наближена ділянка з набору даних

аерофотозйомки окремого міста (geoseer.net/rl.php?ql=62014683a336b412&p=1&q=airphoto#)

Якщо після переміщення по карті буде складно знайти один із шарів, можна натиснути на ньому правою кнопкою миші та обрати Збільшити до шару.

Описаний інструментарій роботи з QGIS дає можливість у загальних рисах уявити процес картографування та відповідні можливості, які воно надає. Як вбачається з наведеного це дуже перспективний напрям з точки зору роботи правоохоронних органів, зокрема в частині впровадження різних моделей безпечного міста. У цьому контексті хотілося б навести рішення HawkEye II від компанії Persistent Surveillance Systems (pss-1.com/hawkeye-ii), яке полягає у постійному виконанні аерофотозйомки певної місцевості з легкомоторних літаків, на яких встановлено фотообладнання високої якості. Зроблені фотознімки, прив'язані до часу, можуть бути використані для виявлення надзвичайних ситуацій, контролю за кордонами та розслідування злочинів. Наприклад, після вчинення злочину шляхом перегляду фотознімків можна побачити маршрут пересування правопорушника, що у подальшому дозволить притягнути його до відповідальності.

Завдання

1. Створіть таблицю, яка буде складатися зі 100 рядків і більше та міститиме інформацію про графіті у вашому районі за такою формою:

Широта	Довгота	Назва об'єкту	Адреса	Кількість	Опис
49,9993	36,24431	Будинок актора	вулиця Манізера, 3, Харків, Харківська область, 61000	2	Реклама наркотиків

2. Імпортуйте підготовлені відомості до програми QGIS.

3. Додайте до карти додаткові шари, які відображатимуть будинки, вулиці тощо та густину населення в місцевості, яка охоплює об'єкти з таблиці.

4. Відпрацюйте зміну параметрів шарів та створіть теплову мапу на основі імпортованих об'єктів.

5. Визначте ділянки з найбільшою густиною графіті.

6. Відпрацюйте навички пошуку адреси за точкою на карті та зворотного пошуку координат за адресою.

7. Порядок виконання завдань відобразіть у звіті, який у якості підсумку має містити короткий аналітичний висновок.

3. Рекомендована література (основна, допоміжна),
інформаційні ресурси в Інтернеті
Основна

1. Манжай О. В. Курс лекцій з дисципліни.
2. Ratcliffe J. H. Intelligence-led Policing. 2nd edn. New York, NY: Routledge, 2016. 234 p.
3. Wang Liang & Zhao Jihong Solomon Contemporary police strategies of crime control in U.S. and China: a comparative study. *Crime, Law and Social Change*. 2016. № 5(66). pp. 525-537.
4. Манжай О. В. Аналіз методології кримінальної розвідки в зарубіжних країнах. *Вісник Луганського державного університету внутрішніх справ імені Е.О. Дідоренка*. 2016. № 3(75). С. 256-265.
5. Потильчак А. О. Щодо співвідношення термінів «кримінальна розвідка» та «кримінальний аналіз» // *Прикарпатський юридичний вісник*. 2017. Вип. 1. Т. 5. С. 174-177.
6. Потильчак А. О. Що таке розвідувальні відомості в контексті моделі Intelligence-Led Policing? // *Visegrad journal on human rights*. 2019. № 6/3. с. 162-166.
7. Манжай О. В., Потильчак А. О. Особливості географічного профілювання у правоохоронних органах // *Право і безпека*. 2020. № 3(78). С. 13-21 (DOI: <https://doi.org/10.32631/pb.2020.3.01>).
8. Манжай О. В., Потильчак А. О. Особливості картографування злочинних проявів // *Право і безпека*. 2020. № 4(79). С. 66-72 (DOI: <https://doi.org/10.32631/pb.2020.4.10>).

Допоміжна

9. Манжай О. В., Жицький Є. О. Кримінальна розвідка та її співвідношення з оперативним обслуговуванням. *Jurnalul Juridic National: Teorie si Practică*. 2015. № 3(13). С. 100-105.

Інформаційні ресурси

10. inteltechniques.com

- 11.