НАУКОВА РОБОТА

ТЕМА: «Виявлення джерел деструктивного інформаційного впливу в мережі Інтернет»

ШИФР «Інформвплив»

ЗМІСТ

АНОТАЦІЯ

наукової роботи на тему: Виявлення джерел деструктивного інформаційного впливу в мережі Інтернет

Актуальність роботи. На цей час ресурси мережі Інтернет стають домінуючим джерелом інформації для людей. В умовах жорстокої конкурентної боротьби, в процес інформування втручаються інформаційні джерела, що створюються з метою маніпулювання свідомістю людей.

У цій роботі пропонується оригінальний підхід до виявлення маніпуляцій них джерел, що базується на математичній лінгвістиці, кластерному аналізі і теорії графів. Враховується, що для максимального залучення уваги користувачів мережі маніпуляційні джерела найчастіше генерують повідомлення із заголовками, що містять ненормативну лексику, так званні меми, сенсаційні епітети тощо. Такі слова і словосполучення можуть виступати маркерами маніпуляцій.

Створення методології і інформаційних технологій виявлення джерел деструктивного інформаційного впливу в мережі Інтернет сьогодні актуально для задач змістовного аналізу мережевої інформації, дослідження суспільної думки, виявлення інформаційних атак і операцій, фільтрації впливу на людей. Проблема сьогодні остаточно не розв’язана, їй займаються дослідники в усьому світі – вона складна і потребує великих витрат.

Мета роботи – створення методології, теоретичних і технологічних засад виявлення джерел маніпуляційного інформаційного впливу шляхом автоматизованого аналізу інформації із соціальних мереж.

Завдання полягає у розв’язанні часткових поставлених задач:

1. Аналіз існуючих підходів до визначення джерел інформаційного впливу через соціальні мережі.
2. Запропонувати та обгрунтувати алгоритми автоматизованого виявлення джерел маніпуляційного інформаційного впливу в мережі Інтернеті, визначення взаємозв’язків між джерелами, основних кластерів джерел інформаційного впливу.
3. Створити інструмент для виявлення і кластеризації джерел маніпуляційного інформаційного впливу через соціальні мережі.

Результати роботи можна використовувати для побудови системи вибору достовірних джерел інформації для задач підтримки прийняття рішень на основі моніторингу мережі Інтернет, в якості готового до застосування засобу виявлення і фільтрації маніпулятивних джерел інформації в умовах гібридних війн.

Наукова робота: 40 с., 9 рис., 20 джерел.

Ключові слова: ІНФОРМАЦІЙНИЙ МАНІПУЛЯТИВНИЙ ВПЛИВ, МЕРЕЖА ДЖЕРЕЛ ІНФОРМАЦІЇ, ВЗАЄМОЗВ'ЯЗКИ ІНФОРМАЦІЙНИХ ДЖЕРЕЛ, СОЦІАЛЬНІ МЕРЕЖІ, КЛАСТЕРИЗАЦІЯ ДЖЕРЕЛ ІНФОРМАЦІЇ

ВСТУП

Актуальність роботи. На цей час ресурси мережі Інтернет стають домінуючим джерелом інформації для багатьох людей. Спочатку передбачалося, що ці ресурси, зокрема, соціальні мережі, самі будуть наповнюватися людьми інформацією, яка стане доступною для їх друзів, а інколи й для більш широкого співтовариства. Але в умовах інформаційних війн, жорстокої конкурентної боротьби, у цей процес втручаються інформаційні джерела, що створюються з метою маніпулювання свідомістю людей [16, 17]. Зазвичай ці джерела наповнюються ботами, спеціальними програмами.

У цій роботі пропонується оригінальний підхід, що базується на математичній лінгвістиці, кластерному аналізі і теорії графів. Враховується, що для максимального залучення уваги користувачів мережі маніпуляційні джерела найчастіше генерують повідомлення із заголовками, що містять ненормативну лексику, так званні меми, багатими сенсаційними епітетами тощо. Справді, такі слова і словосполучення можуть виступати маркерами маніпуляцій, але одне, два повідомлення із такими словами (із тисяч) може містити й доволі об’єктивне джерело. Однак, як показано у цій роботі, велика кількість повідомлень із подібними словами у заголовках достатньо точно маркують маніпуляційні джерела.

Створення методології і інформаційних технологій виявлення джерел деструктивного інформаційного впливу в мережі Інтернет сьогодні актуально для задач змістовного аналізу мережевої інформації, дослідження суспільної думки, виявлення інформаційних атак і операцій, фільтрації маніпулятивного впливу на людей, маркетингу. Ця проблема сьогодні остаточно не розв’язана, їй займаються дослідники в усьому світі [4, 7, 10, 11,14]. Побудова великих промислових систем виявлення маніпуляцій в електронних ЗМІ – це складна проблема, яка потребує великих ресурсних витрат.

Мета роботи – створення методології, теоретичних і технологічних засад виявлення джерел маніпуляційного інформаційного впливу в мережі Інтернет

шляхом автоматизованого аналізу інформації із соціальних мереж (мессенджера Telegram).

Завдання полягає у розв'язанні часткових поставлених задач:

1. Аналіз існуючих підходів до визначення джерел інформаційного впливу через соціальні мережі.
2. Запропонувати та обгрунтувати алгоритми автоматизованого виявлення джерел маніпуляційного інформаційного впливу в мережі Інтернеті, визначення взаємозв'язків між джерелами, основних кластерів джерел інформаційного впливу.
3. Створити інструмент для виявлення і кластеризації джерел маніпуляційного інформаційного впливу через соціальні мережі.

Об'єкт роботи – методи виявлення джерел маніпуляційного інформаційного впливу шляхом автоматизованого аналізу інформації із соціальних мереж (мессенджера Telegram).

Предмет роботи – алгоритми і засоби виявлення джерел маніпуляційного інформаційного впливу на основі інтелектуального аналізу тексту, математичної статистики та теорії графів.

Практичне значення отриманих результатів полягає в створенні програмного забезпечення для автоматизованого вибору достовірних джерел інформації, що надасть можливість подальшого застосування у задачах підтримки прийняття рішень на основі моніторингу соціальних мереж. Крім того, розроблене програмно-алгоритмічне забезпечення можна використовувати на практиці в якості готового засобу виявлення і фільтрації маніпулятивних джерел інформації в умовах гібридних війн.

РОЗДІЛ 1 АНАЛІЗ СУЧАСНОГО СТАНУ ТЕХНОЛОГІЙ ВИЗНАЧЕННЯ ДЖЕРЕЛ МАНІПУЛЯТИВНОГО ІНФОРМАЦІЙНОГО ВПЛИВУ ЧЕРЕЗ СОЦІАЛЬНІ МЕРЕЖІ

Ця робота спрямована на часткове розв'язання проблеми фальшивих новин (Fake News), яка посилюється значним поширенням, поряд із вже звичними веб-сайтами, також соціальними мережами, мессенджерами. Наприклад, багато хто вважає, що фальшиві новини в соціальних мережах в значній мірі сприяють отриманню результатів виборів в США в 2016 році. Ми хотіли створити просту у використанні систему для визначення достовірності тверджень користувача або статті.

Відомо, що серед головних ознак маніпуляційних повідомлень можна назвати такі [13, 20]:

1. Наявність ненормативної лексики, засилля мемів у заголовку.
2. Заголовок має надмірну емоційність, є категоричним, закликає до дій, (на кшталт "Ви ніколи не повірите!", «Такого ви не могли уявити»).
3. Відсутність посилання на перше джерело, специфічна назва сайту, забагато реклами тощо.
4. Виражена мотивація чи зацікавленість у джерела поширити певну інформацію.
5. Відсутність конкретики, яку можна перевірити – місця, часу, імен, назв установ тощо.

1.1 Системи виявлення маніпуляційного інформаційного впливу

На цей час розвиток засобів автоматизації обробки текстів різко знизили вартість поширення фейкових новин у глобальному інформаційному просторі. Більшість користувачів Інтернету стикаються з фальшивими новинами, принаймні, раз на тиждень. Так, судячи з результатів масштабного

європейського дослідження - Євробарометра «Фейк-ньюз і дезінформація онлайн», опублікованим у 2018 році, більше третини респондентів (37%) стикаються з підробленими новинами кожен день [8]. 85% респондентів з різних країн хто вважає фейкові новини проблемою для їхньої країни і 83 відсотки розглядають їх як проблему для демократії в цілому [8]. Тому зараз як ніколи важливо навчитися розпізнавати маніпулятивну інформацію, щоб вміти надалі протистояти їй.

Відомо, що «статистична обробка множини підроблених статей дозволяє виділити набори ключових слів, які з певною часткою ймовірності сигналізують про можливість, що стаття є підробленою» [20].

До фейкових новин відносяться новини, які не містять строго фактичної інформації, не підходять стандартам журналістської етики, а також ті, що відповідають багатьом іншим характеристикам, які будуть надані нижче у цій роботі.

Щоб дійсно виявити підроблені новини, у загальному випадку потрібне застосування алгоритмів, що дозволяють інтерпретувати людську мову, у загальному випадку мають здійснюватися [20]:

1. Перевірка оригінальності URL-адреси повідомлення на відповідність домену сайту. Сайти-підробки можуть цілком імітувати великі новинні сайти.
2. Аналіз фото публікації за принципом збігів такого на сайтах зі списку довірених ресурсів.
3. Виконується перевірка дати публікації. Фейкові новини, як правило, не датовані, тому що говорять про події, яких не було.
4. Визначення наявності граматичних і пунктуаційних помилок - часто підроблені статті формуються з використанням автоматичних генераторів текстів без якісної вичитки. Якщо статтю становить людина, то зазвичай робиться це недбало і з різних фрагментів текстів за шаблоном, що позначається на грамотності тексту.

5. Перевірка на наявність внутрішньої узгодженості аналізованої статті. Підроблені або вводять в оману статті часто мають велику неузгодженість між різними частинами тексту самої статті. Наприклад, алгоритм сканує і оцінює заголовок, основний текст, анотацію новини і т.д. на предмет того, чи узгоджені факти, представлені в статті між собою, чи немає між ними конфліктів і логічних невідповідностей.
6. Пошук збігів фрагментів різних статей. Часто основна стаття, що вийшла на авторитетному ресурсі, копіюється повністю або шляхом рерайта з внесенням в неї неправдивих перекручень фактів і подій.
7. Аналіз посилань вихідних зі статті, ймовірність того, що стаття буде помилковою вище, якщо вона посилається також на помилкові статті, і навпаки. Схожий принцип діє і при визначенні репутації сайтів.
8. Пошук сигнальних слів (сенсаційних). Статті, що містять в своїх заголовках і ключових словах привертають увагу (сенсаційні) затвердження часто мають тенденцію бути підробленими. Статистична обробка безлічі підроблених статей дозволяє виділити набори ключових слів, які з певною часткою ймовірності сигналізують про можливість, що стаття є підробленою.

Слід відмітити, що саме реалізації пункту 8 присвячене програмно-методичне забезпечення, що реалізовано у цій роботі.

До числа компаній, які борються з фальшімвимі новинами, ботами та тролями, приєдналися Facebook, Youtube, Google. Зокрема, Facebook запустив кампанію, в якій використовуються як повідомлення Facebook, так і рекламні оголошення в газетах, щоб надати споживачам поради про те, як ідентифікувати фейк-ньюз [2]. У липні 2018 року про заходи щодо припинення поширення відео-роликів з фальшивими новинами заявив Youtube [3].

Ниже наведено перелік ресурсів для виявлення фейкових повідомлень [15]:

Botometer

Веб-сайт Botometer (первинна назва BotOrNot) був створений в Університеті Індіани в відповідь на поширення в Twitter ботів, що публікують фейковий новини. Цей сайт оцінює аккаунти за шкалою від одного до п'яти, де один означає, що обліковий запис належить реальним користувачам, а п'ятіркою позначаються фейковий аккаунти. Оцінка проводиться на основі твіттів, історії публікацій та згадок іншими користувачами. Більш детальну інформацію можна знайти тут.

Визначник фейковий новин (Detecting Fake News)

Ця програма, яку можна знайти на GitHub, використовує технології машинного навчання і байєсовські моделі для пошуку фейковий новин.

FactCheck.org

На цьому сайті користувачі можуть задавати питання про достовірність інформації, що звучить в заявах політиків, а команда сайту проводить розслідування і пропонує докладне пояснення. Пояснення включає інформацію про те, ким була зроблена заява, коли воно прозвучало і як команда його перевіряла. У сайту також є спеціальна функція для перевірки наукової інформації – SciCheck.

FakeBananas

Розробники з коледжу Суортмор створили Fake Bananas - модель машинного навчання, визначальну фейковий новини з точністю 82 відсотки за допомогою технологій машинного навчання. Програма шукає в авторитетних онлайн-виданнях статті, пов'язані з темою висловлювання, яке потрібно перевірити, і аналізує, чи згодні автори статей до укладеного в висловлюванні твердженням. Якщо достовірні джерела згодні з ним, програма оцінює затвердження як правдиве. Хоча сервіс не розміщений в публічному доступі, програму можна використовувати в інших розробках.

Hoaxy

Hoaxy – це онлайновий інструмент, що візуалізує поширення статей в Інтернеті. Орієнтований на перевірку фейковий новин сайт створює кольорові

інтерактивні графіки, даючи користувачам можливість побачити, як різні заяви поширюються в Twitter. Сайт створено в 2016 році, це спільний проект Центру досліджень комплексних мереж і систем (Center for Complex Networks and Systems Research) та Інституту мережевих наук Університету Індіани (Indiana University Network Science Institute).

Politifact

Лауреат Пулітцерівської премії сайт Politifact перевіряє заяви політиків і блогерів і оцінює ці твердження за шкалою від "правда" до "зовсім забрехався" (pants on fire). Сайт був створений у 2007 році редакцією The Tampa Bay Times, а зараз його роботою керує Інститут Пойнтера. Міжнародна мережа з перевірки фактів включила Politifact в свій список кращих ресурсів.

Snopes

З 1994 року Snopes перевіряє достовірність заяв, статей, постів в соціальних медіа та фотографій. Не обмежуючись простими заявами - "правда" або "брехня", Snopes використовує більш детальні категорії: "правда", "брехня", "суміш того й іншого", "в основному правда", "в основному брехня", "застаріла інформація" , "неправильно зрозуміла інформація" та ін. На сайті також можна знайти список сайтів, що поширюють фейковий новини.

1.2 Ознаки маніпуляційного інформаційного впливу і методи його виявлення

Наявні класифікаційні ознаки фейкових новин враховують такі ознаки [15]:

I. Класифікація, в основу якої покладено критерії, які можуть застосовуватися при виявленні фейкових новин:

за формою їх надання;

– За змістом;

– За тематикою;

– За призначенням для певної вікової категорії;

– За джерелом інформації.

II. Класифікація, в якій критерієм відбору постає мета створення «фейкових» новин:

– «Випадкові фейки» – наслідки редакційного поспіху, некомпетентності журналістів;
– Фейки, створені в рамках інформаційної війни. З метою посилення міжнародних конфліктів;
– Фейки, створені з комерційною метою.
– Фейки, створені для залучення трафіку.
– Фейки, створені з незрозумілою метою.

III. Класифікація за критерієм відбору «неправдивої» інформації за мотивами створення:

– розважальні;
– хибний зв'язок, коли заголовки не віддзеркалюють контент;
– оманливе вживання інформації з метою дискредитації;
– неправдива ситуація;
– вигадані джерела;
– маніпуляційний контент ;
– сфабрикований контент, який на сто відсотків є хибним.

Першим кроком, очевидно, для вирішення завдання виявлення джерела фейковий новин можна вважати отримання даних великого обсягу (корпус). Це завдання вирішується за допомогою роботів, реалізованих, зокрема, в системі Cyber Aggregator [18] для десятка соціальних мереж, в тому числі і мессенджера Telegram. Потім можна аналізувати використовувану в тексті лексику. В цьому випадку тексти корпусу необхідно спочатку піддати автоматичній обробці: розмітити частини мови, виявити емоційно забарвлені слова, імена (в тому числі медіаперсон) і все це порахувати. Робити цю підготовчу роботу лінгвістам допомагають готові програми (наприклад, LIWC, MyStem). Подальше завдання дослідника полягає в тому, щоб, проаналізувавши отримані дані, визначити, які лексичні ознаки є значущими для класифікації. Серед таких маркерів можуть

застосовуватися довжина слів, частотність прикметників, спілок, числівників, цитат, знаків оклику та емоційної лексики.

1.3 Актуальність і мета розробки програмного та інформаційного забезпечення

В умовах інформаційних війн в процес інформування втручаються інформаційні джерела, що створюються з метою маніпулювання свідомістю людей.

Якщо на цей час у світі створюються технології і сервіси, що призначені для виявлення і аналізу окремих фейкових новин, то задачам виявлення джерел таких новин приділено суттєво менша увага, тому що ця інформація менш цікава для пересічного користувача. Проте саме виявлення джерел маніпуляційного інформаційного впливу важливо для підрозділів великих корпорацій, державних установ, засобів масової інформації. Наявні системи автоматичного визначення маніпулятивності окремих новин надають точність виявлення таких новин до 75%. Разом з цим, інтеграційний підхід, що запропоновано у цій роботі забезпечує точність понад 90% для визначення саме джерел фейкових новин. І цьому факту надано математичне обґрунтування. Крім того, вирішується задача автоматичного групування таких інформаційних джерел. У цій роботі пропонується підхід до виявлення маніпуляцій них джерел, що базується на математичній лінгвістиці, кластерному аналізі і теорії графів.

Створення методології і інформаційних технологій виявлення джерел деструктивного інформаційного впливу в мережі Інтернет сьогодні актуально для задач змістовного аналізу мережевої інформації, дослідження суспільної думки, виявлення інформаційних атак і операцій, фільтрації впливу на людей. Проблема сьогодні остаточно не розв’язана, їй займаються дослідники в усьому світі – вона складна і потребує великих витрат.

Висновок по 1 розділу

За результатами аналізу сучасного стану методологій і інформаційних технологій виявлення джерел деструктивного інформаційного впливу в мережі Інтернет було встановлено, що побудова відповідних систем і сервісів – складна і витратна проблема, що на цей час спрямована, насамперед, на виявлення ознак рейкових новин, а не на виявлення маніпуляцій них інформаційних джерел і на їх групування.

Разом з цим, встановлено, що існує декілька підходів до виявлення фейкових новин з текстових корпусів, що приводить, відповідно, до різних видів реалізації технологій. Повідомлення можуть бути визначені як фейкові за допомогою технологій розпізнавання образів, глибинного навчання (Deep Learning), байєсівських алгоритмів машинного навчання, але їх точність не перевищує на цей час 75%.

Разом з цим, інтеграційний підхід, що запропоновано у цій роботі забезпечує точність понад 90% для визначення саме джерел фейкових новин. І цьому факту надано математичне обґрунтування. Крім того, вирішується задача автоматичного групування таких інформаційних джерел. У цій роботі пропонується підхід до виявлення маніпуляцій них джерел, що базується на математичній лінгвістиці, кластерному аналізі і теорії графів.

Тому пропонується розробити систему виявлення джерел деструктивного інформаційного впливу на основі аналізу текстових корпусів з мережевих ЗМІ, соціальних мереж, мессенджерів.

РОЗДІЛ 2 МЕТОДИКА ВИЯВЛЕННЯ І АНАЛІЗУ ДЖЕРЕЛ МАНІПУЛЯЦІЙНОГО ІНФОРМАЦІЙНОГО ВПЛИВУ В МЕРЕЖІ ІНТЕРНЕТІ

Методика виявлення джерел маніпуляційного інформаційного впливу в мережі Інтернеті, визначення взаємозв'язків між джерелами, основних кластерів джерел інформаційного впливу базується на лінгвістичному ймовірнісному підході, кластерному аналізі і теорії графів.

Для реалізації методики застосовуються власні програмні компненти, а також система аналізу і візуалізації графів Gephi (http://gephi.org) [5, 6].

2.1 Етапи методики виявлення і аналізу джерел маніпуляційного інформаційного впливу

Основні етапи, що охоплює представлена методика – це етап навчання системи і етап сталого функціонування. Методика, таким чином реалізована у вигляді автоматизованої системи, що охоплює модулі навчання, модулі сталого функціонування і зовнішній пакет програм – систему Gephi. Обидва етапи системи пов'язані із ланцюжками взаємонезалежних кроків, зокрема, етап навчання системи полягає у створенні детального словника слів і словосполучень, що маркують маніпуляційні повідомлення, складається із (Додаток Г):

1) Формування масиву вхідних повідомлень, що скануються із соціальних мереж або меседжерів (у цій як полігон для представлення методики використовувалися повідомлення мессенджера Telegram);
2) Формування тимчасового словника слів, що можуть маркувати повідомлення маніпуляційного характеру (термінів);
3) Фільтрація масиву вхідних повідомлень за допомогою цього словника;

4) Змістовний аналіз результатів фільтації із застосуванням програмних компонентів виявлення найбільш значущих слів [12, 19] і коригування тимчасового словника.
5) Оформлення детального словника слів і словосполучень, що маркують маніпуляційні повідомлення.

Другий етап (етап сталого функціонування), містить такі кроки (Додаток Г)):

1) Фільтрація масиву вхідних повідомлень за допомогою детального словника слів і словосполучень, що маркують маніпуляційні повідомлення (враховуються тільки заголовки повідомлень);
2) Результати фільтрації, представлені у вигляді записів, що склдаються із двох полів – назва джерела і заголовок повідомлення сортуються за назвами повідомлень. Джерела, що зустрічаються найчастіше всього виводяться для подальшого аналізу, як найбільш вірогідні маніпуляційні джерела інформації. Співвідношення між рівнем маніпуляційності окремого повідомлення і джерела цього повідомлення наведено у п. 2.2. цієї роботи.
3) Далі визначається взаємозв'язок між джерелами маніпуляційної інформації. Для цього у відповідність кожному джерелу інформації ставиться вектор, елементи якого відповідають словнику слів і словосполучень, що маркують маніпуляційні повідомлення. Значення цих елементів відповідає кількості появ слів із цього словника в повідомленнях даного джерела. Матриця взаємозв'язку джерел складається із елементів, що відображають близькість відповідних джерел.
4) Отримана матриця взаємозв'язку джерел завантажується у систему Gephi, де вона штатними засобами цієї системи кластерізується і візуалізується (за класами модулярності).

Людина-аналітик визначає інформаційну спрямованість кожного кластера, що було визначено системою автоматично.

2.2 Формальне співвідношення між маніпулятивністю окремих повідомлень і джерел інформації

Нехай G – множина слів, що маркують маніпуляції: $G = \{g_i\}_{i=1}^{|G|}$. Позначимо множину джерел як S: $S = \{s_k\}_{k=1}^{|S|}$.

Ймовірність того, що повідомлення d є маніпулятивним, якщо воно містить слово g_i, позначимо як p_i. Відповідно, ймовірність того, що повідомлення d не є маніпулятивним, якщо воно містить слово g_i, позначимо як q_i $(p_i + q_i = 1)$.

Нехай повідомлення d містить декілька слів з G: $\exists i : g_i \in d,\ g_i \in G$.

Ймовірність того, що повідомлення d не є маніпулятивним, у цьому випадку дорівнює:

$$q(d) = \prod_{i:g_i \in d,\ g_i \in G} (1 - p_i).$$

Тоді ймовірність того, що повідомлення d є маніпулятивним, дорівнює:

$$p(d) = 1 - \prod_{i:g_i \in d,\ g_i \in G} (1 - p_i).$$

Ймовірність того, що всі повідомлення джерела S_k не є маніпулятивним, дорівнює:

$$q(S_k) = \prod_{d \in S_k} \prod_{i:g_i \in d,\ g_i \in G} (1 - p_i).$$

З того, що $0 \le 1 - p_i \le 1$, випливає, що для будь якого документа d^j:

$$q(S_k) \le \prod_{g_i \in d^j,\ g_i \in G,} (1 - p_i).$$

Тобто для ймовірності того, джерело S_k містить маніпулятивні повідомлення (є маніпулятивним), має місце нерівність:

$$p(S_k) = 1 - q(S_k) = 1 - q(S_k) = 1 - \prod_{d \in S_k} \prod_{i: g_i \in d,\ g_i \in G} (1 - p_i) \geq 1 - \prod_{g_i \in d^j,\ g_i \in G,} (1 - p_i).$$

Тому ймовірність маніпулятивності джерела завжди вище маніпулятивності будь якого окремого повідомлення, що відноситься до нього.

2.3 Формування мережі взаємозв'язку джерел

Мережа взаємозв'язку джерел формується за кореляційним принципом [9].

Кожному джерелу s_k із $S = \{s_k\}_{k=1}^{|S|}$ ставиться у відповідність вектор $\overline{w^k} = (w_1^k, w_2^k, ..., w_n^k)$, де $n = |G|$ – кількість елементів в множині слів G, а кожне значення w_i^k відповідає кількості повідомлень в джерелі s_k, заголовки яких містять слово g_i.

Кореляція між джерелами s_i і s_j (a_{ij}) визначається як косинус кута між відповідними векторами $\overline{w^i}$ та $\overline{w^j}$:

$$A_{ij} = \frac{\left(\overline{w^i}, \overline{w^j}\right)}{\left|\overline{w^i}\right|\left|\overline{w^j}\right|} = \frac{\sqrt{\sum_{k=1}^{n} w_k^i w_k^j}}{\sqrt{\sum_{k=1}^{n} \left(w_k^i\right)^2} \sqrt{\sum_{k=1}^{n} \left(w_k^j\right)^2}}.$$

Квадратну матрицю A з елементами A_{ij}, $A = \|A_{ij}\|$, будемо називати матрицею суміжності, або матрицею взаємозв'язку джерел.

2.4 Кластеризація і візуалізація мережі джерел

Для кластеризації мережі джерел застосовується стандартний механізм, вбудований в систему Gephi, а для виділення кластерів застосовується властивість модулярності.

Модулярність — один з мережевих параметрів, який було введено для вимірювання ступеня розбиття мережі на модулі (кластери) [1].

Модулярність мережі обчислюється як різниця між долею ребер всередині кластера в мережі, що розглядається, і очікуваній долі ребер всередині кластера в мережі, в якій вершини мають той самий ступень, що і в первинній, але ребра розподілені випадково.

Для розрахунку модулярності використовується матриця суміжності A складається з елементів A_{vw} , значення яких дорівнюють 0, якщо вузол v не з'єднано з вузлом w, і вазі зв'язку між v і w, якщо ці вузли з'єднані між собою (Див. п. 2.3).

Модулярність мережі можна виразити формулою:

$$Q = \frac{1}{(2m)} \sum_{vw} \left[A_{vw} - \frac{k_v k_w}{(2m)} \right] \delta(c_v, c_w) \quad ,$$

де A_{vw} – елемент матриці суміжності A, m – кількість ребер у графі, k_v, k_w – ступені вузлів v і w відповідно, δ – дельта функція Кронекера (показує, чи знаходяться вузли v і w в одному модулі).

Отже, модулярність – це міра якості кластеризації, на базі якої будується широкий клас алгоритмів виявлення груп в мережах.

Gephi (https://gephi.org/) – на цей час найпопулярніша програма візуалізації і аналізу графів. Gephi забезпечує швидку компоновку, ефективну фільтрацію і інтерактивне дослідження даних, є одним з кращих варіантів для візуалізації великомасштабних мереж. Gephi - це мультиплатформне програмне забезпечення яке розповсюджується з відкритим кодом згідно з ліцензіями CDDL 1.0 і GNU General Public License v3.

Існує три основні режими створення нового графа в Gephi:

– через інтерфейс «Граф» в режимі «Обробка»;

– через «Лабораторію даних»;

– через експортування даних графа із зовнішнього файла (найпростіше із файла у форматі CSV).

В технології, що пропонується, застосовується саме третій режим. Тобто основний варіант експорту даних графа із зовнішнього файлу - це завантаження

вихідних мережевих даних у форматі CSV, в якому елементи розділені знаком «крапка з комою». В цьому випадку в CSV-файлі фактично повинна знаходитися розширена тегами (назвами вузлів) матриця суміжності мережі. Нижче наводиться приклад для мережі з п'яти вузлів:

```
1)  ;Node1;Node2;Node3;Node4;Node5
2)  Node1;0;1;0;1;0
3)  Node2;1;0;0;1;0
4)  Node3;0;1;0;0;1
5)  Node4;1;1;1;0;0
6)  Node5;0;1;0;1;0
```

Після завантаження в систему Gephi і обробки вже описаним способом, отримуємо відображення (Рис. 2.1):

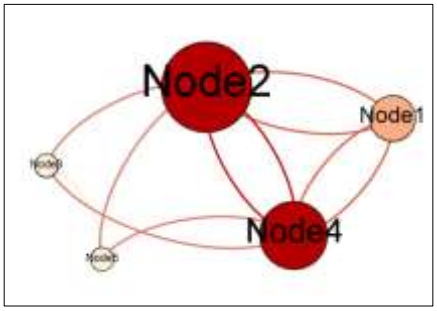


Рис. 2.1 – Відображення графа,
завантаженого із файлу в форматі CSV

При аналізі великих і щільних мереж швидке компонування (впорядкування вузлів графів) є вузьким місцем, оскільки більшість складних алгоритмів компонування вимогливі до параметрів процесора, пам'яті і часу виконання. У той же час Gephi поставляється з ефективними алгоритмами компонування, такі як Yifan-Hu, Force-directed, які можуть використовуватися для таких мереж. Зокрема, алгоритм Yifan-Hu є ідеальним варіантом для застосування після інших, більш швидких і грубих алгоритмів.

Ранжирування дозволяє встановлювати розміри вузлів і фарбувати графи на основі зазначеного користувачем атрибута, такого як ваги, міри (за замовчуванням) і багато іншого, в залежності від доступних полів в нашому наборі даних. Багато з таких полів, наприклад, значення PageRank, модулярності,

можна отримати в режимі «Лабораторія даних» у вкладці «Статистика», яка знаходиться в правій частині інтерфейсу. Цей розділ дозволяє отримувати різні спеціальні розрізи статистики для подальшого аналізу різних параметрів графів.

При послідовному натисканні всіх кнопок «Запуск» поспіль, можна отримати значення характеристик графа.

Розраховані характеристики доступні у вкладці «Таблиця даних».

Використовуємо результати роботи в режимі «Лабораторія даних» для візуалізації. Після розрахунку відповідних параметрів з'являється можливість змінювати розміри і кольору вузлів не тільки за ступенем, але і за іншими характеристиками, таким як модулярність, кластерність, тощо.

Якщо повернутися до ранжирування в режимі «обробка», то бачимо додаткові можливості (Рис. 2.2).

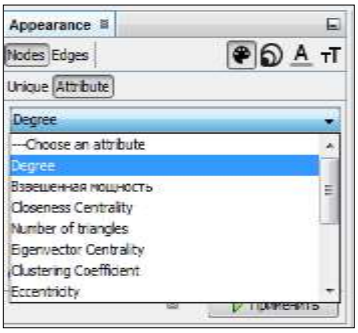


Рис. 2.2 – Критерії ранжирування для кольорів вузлів

Застосувавши ранжирування за класами модулярності можна виділити кластери мережі.

2.5 Результати, що передбачається отримати

В результаті виконання цієї роботи передбачається отримати:

1. Отсортований за критерієм маніпулятивності перелік інформаційних джерел
2. Мережу джерел із визначеними взаємозв’язками
3. Виділені класи (кластери) джерел маніпулятивного інформаційного впливу

4. Методику і інструментальні засоби для подальшого аналізу інформаційного простору щодо джерел маніпуляційного інформаційного впливу.

Висновок по 2 розділу

В результаті розробки розділу запропоновано та обґрунтовано методику автоматизованого виявлення джерел маніпуляційного інформаційного впливу в мережі Інтернеті, визначення взаємозв’язків між джерелами, основних кластерів джерел інформаційного впливу базується на лінгвістичному ймовірнісному підході, кластерному аналізі і теорії графів.

За допомогою даної методики забезпечується вибір найважливіших термінів, створення словника маркерів маніпуляційного інформаційного впливу, формується ранжируваний список інформаційних джерел, створюється мережа взаємозв’язків між джерелами. Побудовану мережу можна використовувати в якості основи класифікації , виявлення спрямованості маніпуляцій них джерел. Побудовану методику можна реалізувати в якості інструментального засобу виявлення і фільтрації маніпулятивних джерел інформації в умовах гібридних війн.

РОЗДІЛ 3 РЕАЛІЗАЦІЯ СИСТЕМИ ВИЯВЛЕННЯ І КЛАСТЕРИЗАЦІЇ ДЖЕРЕЛ ІНФОРМАЦІЙНОГО ВПЛИВУ ЧЕРЕЗ СОЦІАЛЬНІ МЕРЕЖІ

3.1 Формування масиву повідомлень мессенджера Telegram

Для формування масиву повідомлень мессенджера Telegram програмно здійснюється сканування відповідних (заздалегідь визначених) Telegram-каналів за адресами вигляду:

https://tlgrm.ru/channels/[Канал]

де [Канал] – це назва каналу, наприклад, https://tlgrm.ru/channels/@The_Idiot.

В межах сканованих сторінок програмно знаходиться адреса типу:

https://api.tlgrm.ru/v3/channels/[Код каналу]/feed

де [Код каналу] – це цифровий код каналу, наприклад, https://api.tlgrm.ru/v3/channels/1108035346/feed, а за вказаною адресою знаходиться JSON-фід, у якому в Юнікод-кодуванні знаходяться повідомлення, розмічені наступним чином:

"id":(\d+?)

"title":"(.+?)"

"text":"(.+?)"

Нижче наведено фрагмент повідомлення у форматі JSON:


```
{"_id":11080353468516,
"channel_id":1108035346,
"id":8516,
"date":1579879618000,
"text":"\u041d\u0430\u0432\u0430\u043b\u044c\u043d\u043e\u043c\u0443
\u0441\u0440\u043e\u0447\u043d\u043e \u043d\u0443\u0436\u043d\u043e
\u043f\u0440\u043e\u0442\u0435\u0441\u0442\u0438\u0440\u043e\u0432\u0430\u0442\u
044c \u043d\u043e\u0432\u0443\u044e \
. . .
"title":"\u0412\u044b\u043f\u0443\u0441\u043a\u0430\u0439\u0442\u0435
\u041a\u0440\u0430\u043a\u0435\u043d\u0430!"}
```

Далі повідомлення програмно збираються і перетворюються до записів бази даних з відповідними формату JSON полями.

Для проведення випробовувань системи було зібрано понад 300 тисяч повідомлень з 6 тисяч джерел Telegram-каналів за 2019 рік.

3.2 Фільтрація вхідних повідомлень і сортування джерел інформації

Для фільтрації вхідних повідомлень за словником слів і словосполучень розроблений окремий програмний модуль, текст якого наведено у Додатку А. На вхід цієї програми подається масив вхідних повідомлень (Див. п. 3.1) і текстовий файл із словником, кожний рядок якого відповідає мему, що ймовірно маркує маніпуляцій не інформаційне повідомлення. На виході цієї програми створюється текстовий файл (fn.txt), кожний рядок якого наведений у такому форматі (Рис. 3.1):

Джерело: Заголовок,

де «Джерело» – це назва каналу джерела, а «Заголовок» – це заголовок повідомлення.

Після того, як отримано цей текстовий файл він сортується за алфавітом стандартною утілитою sort:

>sort –o fn.txt fn.txt

Далі на основі цього файлу за допомогою програмного модуля, вихідний код якого наведено у Додатку В, створюється файл (an.txt), в якому відображено назву джерела-каналу і кількість повідомлень, що відносяться до відповідних каналів у форматі:

Кількість: Джерело,

де «Джерело» – це назва каналу джерела, а «Кількість» – це кількість повідомлень, що відповідають цьому джерелу. Сортування цього файлу за кількістю джерел дозволяє отримати рейтинг джерел за кількістю ймовірно маніпуляцій них повідомлень. Найбільш рейтингові джерела надалі застосовуються при побудові мережі взаємозв’язку джерел. Слід зазначити, що

детальний аналіз визначених джерел інформації показав точність віднесення їх класу маніпулятивних джерел перевищила 90% (366 дерел з 407 відібраних).

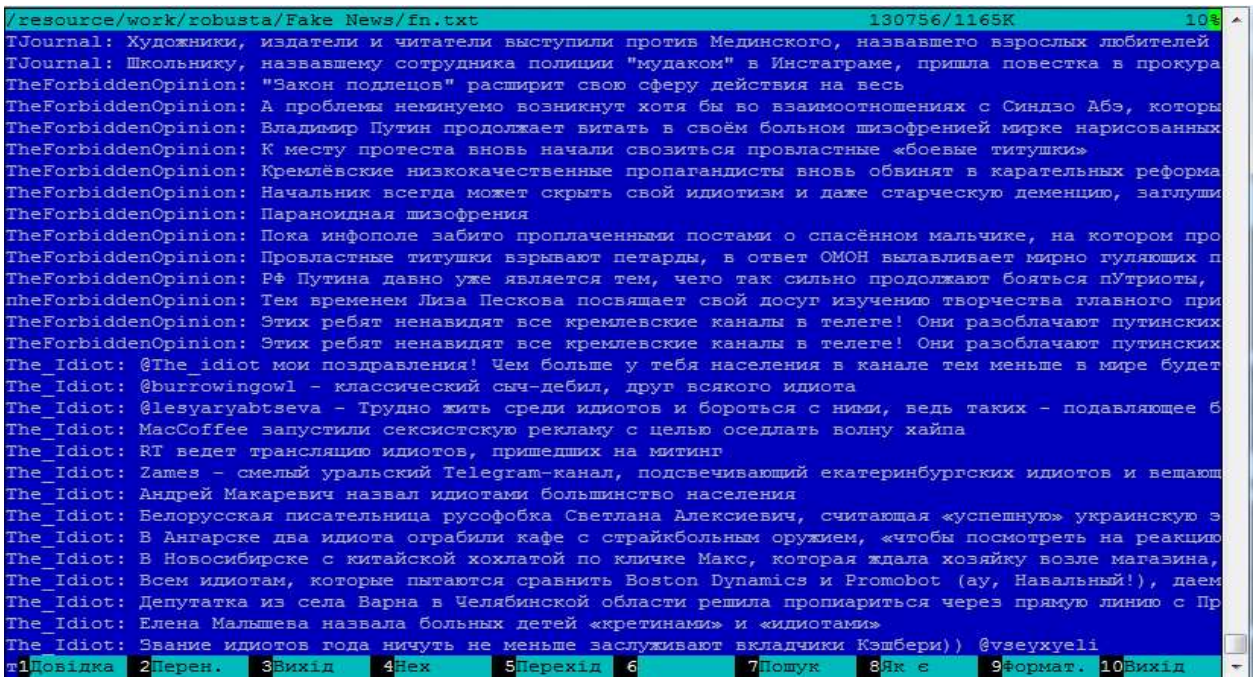


Рис 3.1– Фрагмент вихідного файлу програми фільтрації повідомлень

3.3 Формування матриці взаємозв’язку вузлів мережі джерел

Формування матриці взаємозв’язку вузлів мережі джерел здійснюється у відповідності із п. 2.3 за допомогою програмного модуля, вихідний код якого наведено у Додатку Б. Цією програмою досліджується заздалегідь визначена множина джерел інформації (верхня частина рейтину, визначеному у п.2). На вхід програми подається цей перелік джерел, словник слів і словосполучень маніпулятивної спрямованості і вихідний файл програми фільтрації повідомлень, фрагмент якого наведено на Рис. 3.1. На виході програми – файл у форматі CSV, структуру якого описано у розділі 2.4.

3.4 Візуалізація мережі засобами Gephi

На останньому етапі формування і аналізу мережі взаємозв’язку джерел здійснюється її відображення за допомогою програмного пакету аналізу і візуалізації графів Gephi (https://gephi.org/). Для завантаження мережі джерел до

баз даних цієї системи приведено відповідну матрицю суміжності до загальноприйнятого формату CSV.

Після завантаження файлу в середовище Gephi, засобами цієї системи (Рис.3.2) здійснюється кластиризація вихідної мережі, обрахунок значень модулярності для кожного із кластерів (Додаток Д, відображені статистичні дані мережі, де вузли відсортовані за класом модулярності, другий ключ сортування – ступінь вузла) і візуалізація мережі джерел інформації, що мають ознаки маніпуляційності (Рис. 3.3).

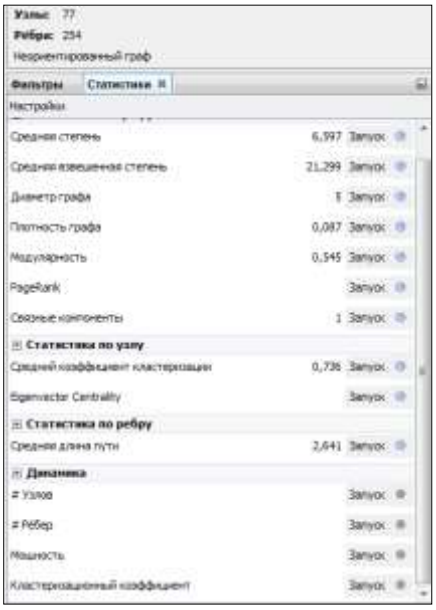


Рис. 3.2 – Аналіз мережі – режим «Лабораторія даних». Тека «Статистика»

На Рис. 3.4 наведено верхню частину рейтингу джерел інформації за їх ступенем (як одним із показників важливості в мережі).

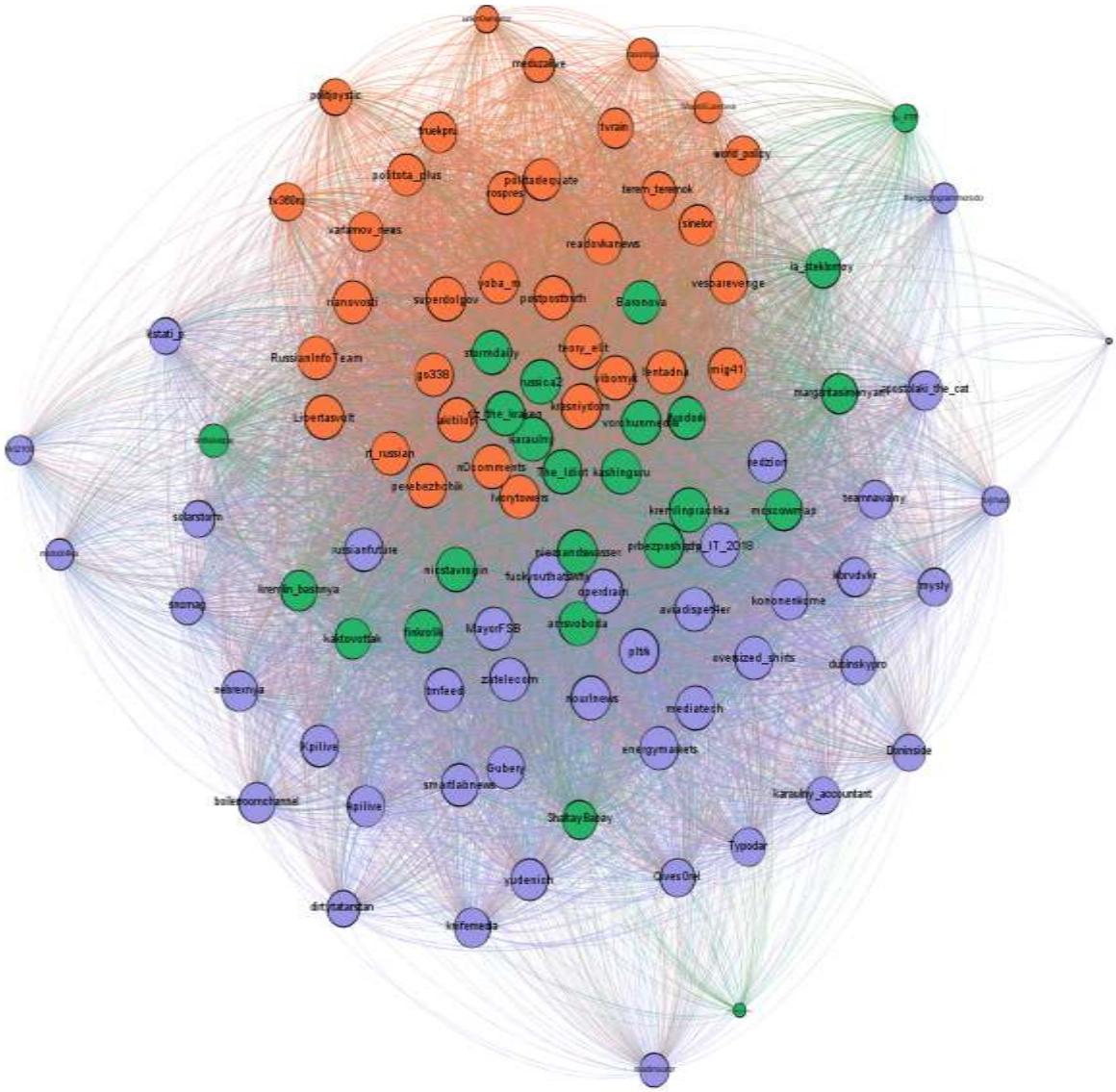


Рисунок 3.3 – Мережа джерел інформації у середовищі Gephi

Висновки по 3 розділу

В результаті розробки розділу було сформовано масив публікацій із мессенджера Telegram. Розроблений комлпекс програмних модулів для автоматизованої побудови і візуалізації мережі понять предметної області, за допомогою якого оброблено новини з мережевих ЗМІ та створено мережу понять предметної області. За допомогою програмного пакету аналізу і візуалізації графів Gephi візуалізовано мережу джерел маніпулятивної інформації, що допомагає краще її сприймати та обробляти. Дослідження, проведені на реальному масиві даних дозволили визначити основні три класи модулярності, а разом із тим, аналітики визначили зміст, притаманний цим класам – 1-й клас –

російські провладні джерела «правого» спрямування, 2-й клас – російські протестні джерела «лівого», але шовіністичного спрямування (Навальний, тощо), 3-й клас – джерела, в яких критикується сучасна політика РФ.

Id	Суммарная мощность
kashinguru	244
rlz_the_kraken	240
fuckyouthatswhy	238
karaulny	236
krasniydom	216
niemandswasser	214
go338	196
yoba_m	194
nicstavrogin	188
fyodork	178
arhsvoboda	174
pltrk	174
perebezhchik	172
operdrain	168
russica2	166
ia_steklomoy	164
vibornyk	152
Gubery	146
politadequate	146
karaulny_accountant	144
akitilop	138
vorchunmedia	134
vesparevenge	134
The_Idiot	130
redzion	128
stormdaily	128
prbezposhady	126
pro_IT_2018	126
kbrvdvkr	124
kremlin_bashnya	120
apostolaki_the_cat	120
Ivorytowers	120
terem_teremok	118
mig41	118
RussianInfoTeam	118
margaritasimonyan	116
mediasrachi	114
solarstorm	114
tmfeed	112
readovkanews	112
finkrolik	110

Рис. 3.4 – Найбільш маніпулятивні джерела Telegram, отримані за лінгвістичним алгоритмом

ВИСНОВКИ

По-перше, на сьогоднішній день визначення і ранжирування джерел маніпулятивного інформаційного впливу через соціальні мережі та автоматичного встановлення зв'язків між ними на рівні всього інтернет-простору – складна і витратна проблема. Встановлено, що існує декілька підходів до рішення цієї задачі та різні способи ранжирування за критеріями інформаційного впливу та кластеризації джерел інформації, що приводить, відповідно, до різних результатів, різних представлень мереж взаємозв’язків джерел інформації. Визначено, що найбільш результативний на цей час підхід – лінгвомережевий, на базі якого і проводилися ці дослідження.

По-друге, у роботі наведена методика автоматизованого виявлення джерел маніпуляційного інформаційного впливу в мережі Інтернеті. Крім того запропоновано новий метод визначення взаємозв’язків між джерелами та основних кластерів джерел інформаційного впливу.

По третє, було розроблено комлпекс програмних модулів для автоматизованого виявлення і кластеризації джерел маніпуляційного інформаційного впливу через соціальні мережі, побудови і візуалізації мережі джерел маніпуляційного інформаційного впливу. За допомогою програмного пакету аналізу і візуалізації складних мереж Gephi візуалізовано мережу взаємозв’язків джерел, що допомагає краще її сприймати та обробляти.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Albert-László Barabási. Network Science. Cambridge University Press, 2016.

2. Al-Heeti A. Facebook Will Fight Fake News with Real Newspaper Ads (and More). *CNET*. May 23, 2018.

3. Building a better news experience on YouTube, together. URL: https://youtube.googleblog.com/2018/07/building-better-news-experience-on.html

4. Busari S. How fake news does real harm. *TED talks conference,* 2017.

5. Ken Cherven. Mastering Gephi Network Visualization. Packt Publishing, 2015.

6. Ken Cherven. Network Graph Analysis and Visualization with Gephi. 2013 Packt Publishing, 2013.

7. Niall J. Conroy, Victoria L. Rubin, Yimin Chen. Automatic deception detection: Methods for finding fake news // asis&t, 2015. – Vol. 52, Iss. 1. – pp. 1-4. DOI: 10.1002/pra2.2015.145052010082.

8. Fake news and disinformation online. Report. European Union, 2018. URL: https://ec.europa.eu/digital-single-market/en/news/final-results-eurobarometer-fake-news-and-online-disinformation.

9. John W. Foreman. Data Smart. Using Data Science to Transform Information into Insight. Wiley, 2013.

10. How Content Discovery Platforms Can Fight Fake News via Web Scraping and AI. URL: https://www.promptcloud.com/blog/fight-fake-news-web-scraping-artificial-intelligence.

11. David M. J. Lazer, Matthew A. Baum, Yochai Benkler etc. The science of fake news. *Science*, 09 March 2018. – Vol 359, Issue 6380. – pp. 1094-1096. DOI: 10.1126/science.aao2998.

12. Ortuño M., Carpena P., Bernaola P., Muñoz E., Somoza A.M. Keyword detection in natural languages and DNA. *Europhys*. Lett., 2002. – 57. – P. 759-764.

13. David Sumpter. Outnumbered: From Facebook and Google to Fake News and Filter-bubbles. Bloomsbury Sigma, 2018.

14. Clint Watts. Messing with the Enemy: Surviving in a Social Media World of Hackers, Terrorists, Russians, and Fake News. Harper Paperbacks, 2019.

15. Главацька Ю.Л. Класифікація «фейкових» новин у сучасному медіапросторі: синергетичний аспект. *Науковий вісник Херсонського державного університету*. 2019. Випуск 1. С. 275-280

16. Горбулін В.П., Додонов О.Г., Ланде Д.В. Інформаційні операции та безпека суспільства: Загрози, протидія, моделювання: монографія. К.: Інтертехнологія, 2009. 164 с.

17. Додонов А.Г., Ландэ Д.В. Моделирование и анализ тематических информационных потоков. *Информационное противодействие угрозам терроризма*, 2013. № 20. С. 52–59.

18. Кальян Н.А., Матіїшин О.Т. Система контент-моніторингу соціальних мереж з питань кібербезпеки. "*Інтелектуальний потенціал - 2019*" - збірник наукових праць молодих науковців і студентів з нагоди 30-річчя кафедри кібербезпеки та комп'ютерних систем і мереж ХНУ. Ч.1: Комп'ютерні системи та кібербезпека. Хмельницький: ПВНЗ УЕП, 2019. С. 28-30.

19. Ланде Д.В. Методи оцінки рівня дискримінантної сили слів у текстах з правової тематики. *Правова інформатика*. 2012. № 3 (35). С. 5-9.

20. Некрасов Г.А., Романова И.И. Разработка поискового робота для обнаружения веб-контента с фейковыми новостями. *Инновационные, информационные и коммуникационные технологии*. 2017. № 1. С. 128-130.

ДОДАТОК А

Лістинг модуля фільтрації аккаунтів маніпулятивних джерел і заголовків відповідних повідомлень

```
#!/usr/bin/perl
use Encode;
use utf8;

$x="man.src";   # Словник
open F,"$x";
$i=0;
$shablon="";  # Шаблон пошуку
while ($_=<F>) {
    chomp();
    $buf=$_;
    $buf=~s/\s+$//g;
    $buf=~s/\]/ /g;
    if (length($buf)>1) {
        $syt[$i]=$buf;
        $shablon.=$buf."|";  # Шаблон пошуку
        $i++;
        $iu++;
    }
}

$shablon=~s/\|$//;
utf8::decode($shablon);
close F;

$N=$i;
# Масив із повідомлень Telegram із Cyber Aggregator
open F,"tlg.t";

while ($_=<F>) {
    $buf=$_;
    if ($buf=~m/^\*\*\*/) {
        $_=<F>;
        chomp();
        $buf=$_;
        utf8::decode($buf);
        if ($buf=~m/($shablon)/) {
            $_=<F>;
            $buf1=$_;
            utf8::decode($buf1);
            if ($buf1=~m/^User: (.+?)\.(.+)$/) {
               $aut=$1;
               $aut=~s/@//g;
```

```
                # Виведення результатів фільтрації:
                print $aut,": ",$buf,"\n";
              }
          }
    }
}
close F;
```

ДОДАТОК Б

Лістинг модуля обрахунку продуктивності аккаунтів

```
#!/usr/local/bin/perl -w

$x="fn.txt";
open F,"$x";
while ($_=<F>) {
    chomp();
    $buf=$_;
    if ($buf=~m/^(.+?):/) {
        $a=$1;
        if (exists $name{$a}) {
            $name{$a}++;
        }
        else {
            $name{$a}=1;
        }
    }
}
close F;

foreach my $x (sort {lc $a cmp lc $b} keys %name) {
    print "\n$name{$x}\t$x";
}
```

ДОДАТОК В

Лістинг модуля побудови таблиці інформаційних взаємозв'язків аккаунтів у форматі *.csv

```
#!/usr/bin/perl –w

use Encode;
use utf8;

$x="man.src";  # Словник

open F,"$x";
$i=0;
$shablon="";
while ($_=<F>) {
    chomp();
    $buf=$_;
    utf8::decode($buf);
    $buf=~s/\s+$//g;
    $buf=~s/\]/ /g;
    if (length($buf)>1) {
        $syt[$i]=$buf; # Слово-індикатор
        $i++;
    }
}
close F;
$N=$i;

$i=0;
open F,"an.txt";  # Масив аккаунтів
while ($_=<F>) {
    chomp();
    $buf=$_;
    if ($buf=~m/^.*?\t(.+)$/) {
        $name[$i]=$1;
    }
    $i++;
    if ($i>199) { #99
        last;
    }
}
close F;
$M=$i;
for ($i=0; $i<$M; $i++) {
    for ($j=0; $j<$N; $j++) {
        $mtr[$i][$j]=0;  # Матриця «аккаунт-слово»
    }
}
# Масив повідомлень Telegram із Cyber Aggregator
```

```
open F,"fn.txt";
while ($_=<F>) {
    $buf=$_;
    utf8::decode($buf);
    if ($buf=~m/^(.+?):(.+)$/) {
        $nam=$1;
        $doc=$2;
        for ($i=0; $i<$M; $i++) {
            if ($name[$i] eq $nam) {
                # Заповнення матриці «аканут-слово»
                                for ($j=0; $j<$N; $j++) {
                   if ($doc=~m/$syt[$j]/) {
                       $mtr[$i][$j]=1; #++;
                   }
                }
                next;
            }
        }
    }
}
close F;
for ($i=0; $i<$M; $i++) {
    for ($j=0; $j<$M; $j++) {
        $MTR[$i][$j]=0; # Матриця взаємозв’язку джерел
    }
}
for ($i=0; $i<$M; $i++) {
    for ($j=0; $j<$M; $j++) {
        if ($i!=$j) {
        for ($k=0; $k<$N; $k++) {
            $a=$mtr[$i][$k];
            $b=$mtr[$j][$k];
            if ($a>$b) {$a=$b; }
               $MTR[$i][$j]+=$a; # Заповнення матриці $MTR
            }
        }
    }
}
for ($i=0; $i<$M; $i++) {
    print ";$name[$i]"
}
for ($i=0; $i<$M; $i++) {
    print "\n$name[$i]";
    for ($j=0; $j<$M; $j++) {
        $x=$MTR[$i][$j];
        if ($x<5) { $x=0 } else { $x=1 };
        print ";$x";
    }
}
```

ДОДАТОК Г

Рис. Г1 – Блок-схема етапу навчання системи

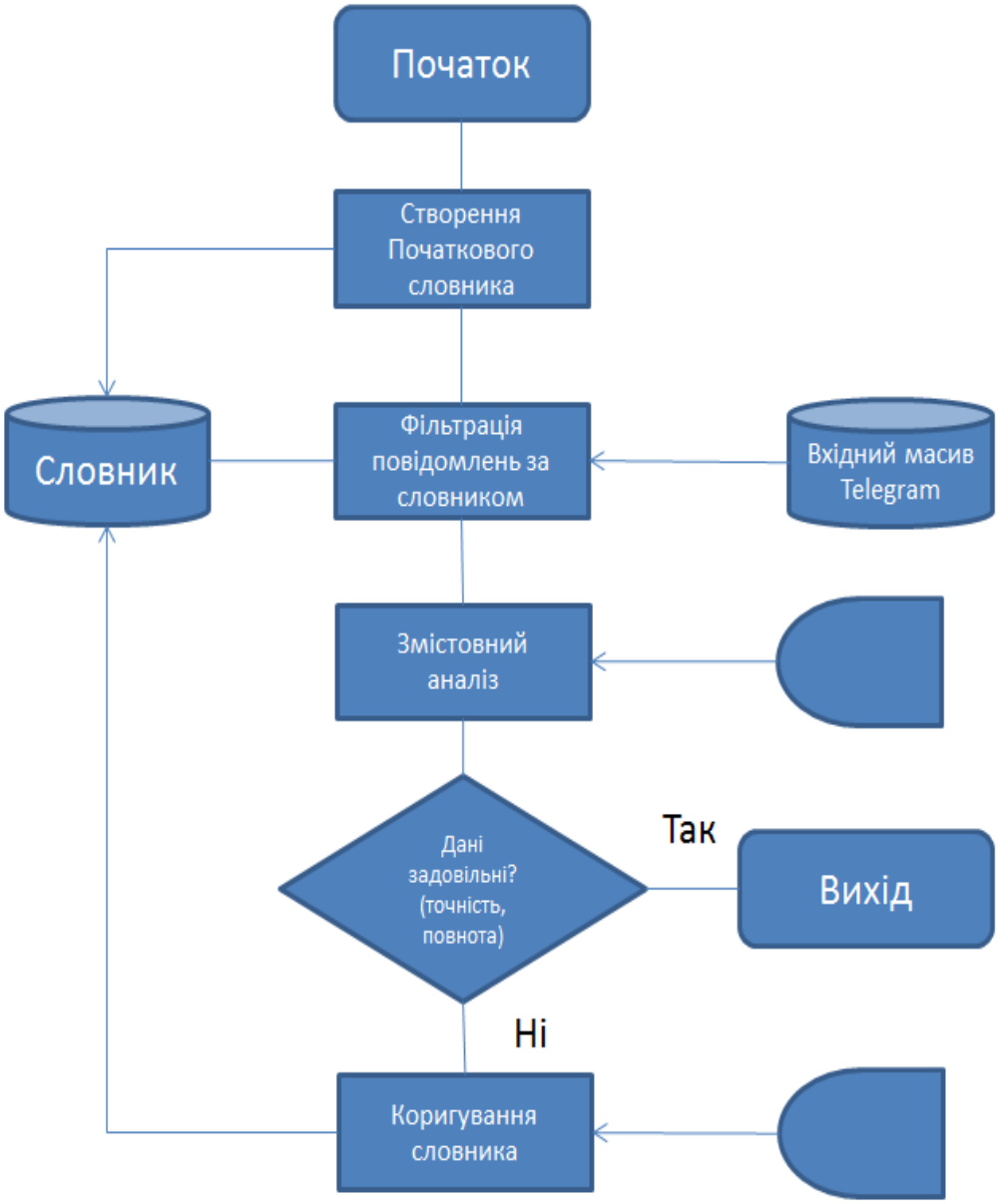
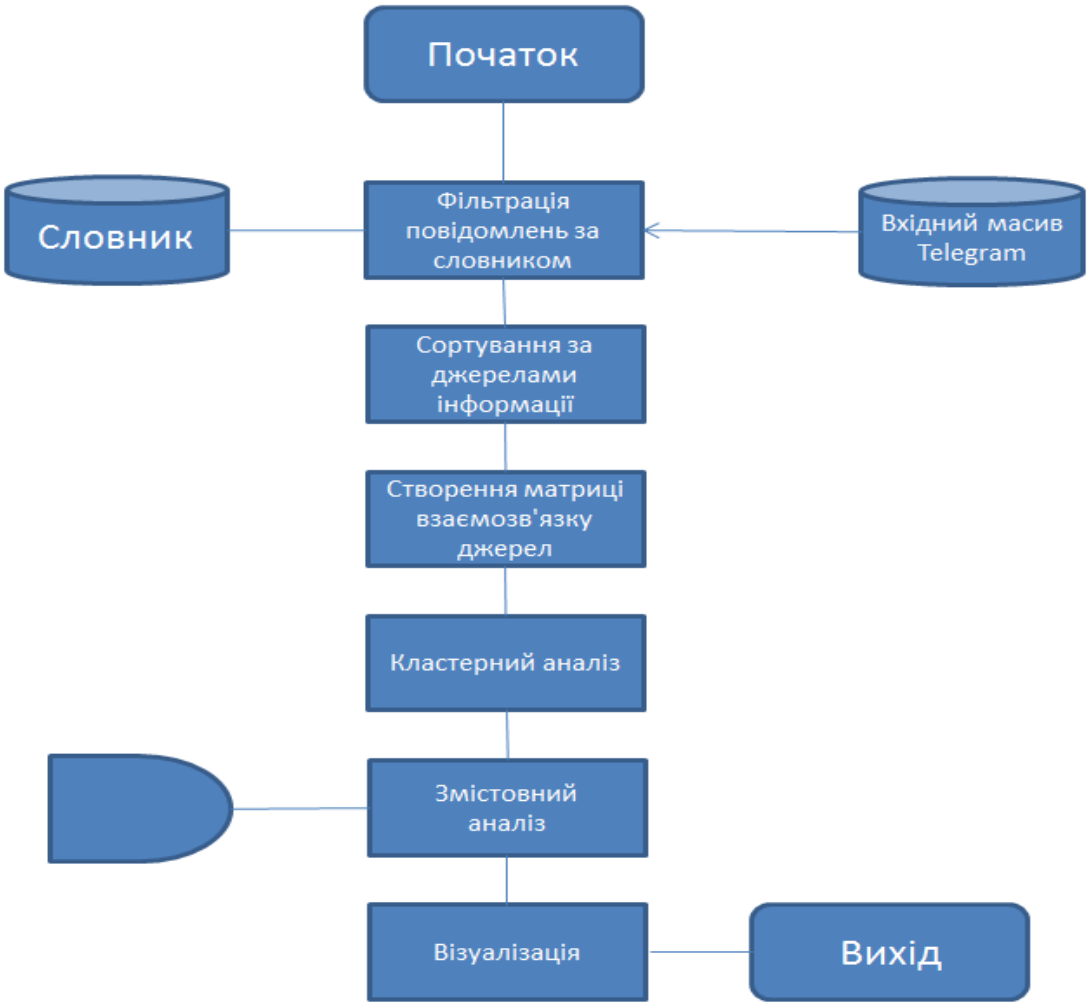


Рис. Г2 – Блок-схема етапу сталого функціонування системи



ДОДАТОК Д

Рисунок Д1 – Вузли відсортовані за класом модулярності

Id	Modularity Class	Входящая мощность
kbrvdvkr	0	86
energymarkets	0	96
dirtytatarstan	0	80
apostolaki_the_cat	0	87
Typodar	0	86
Kpilive	0	92
Doninside	0	82
CivesOrel	0	85
closET174	0	11
karaulny	1	98
kashinguru	1	99
niemandswasser	1	98
arhsvoboda	1	98
rlz_the_kraken	1	97
The_Idiot	1	97
prbezposhady	1	98
vorchunmedia	1	97
russica2	1	98
kremlinprachka	1	98
nicstavrogin	1	98
fyodork	1	93
Baronova	1	95
stormdaily	1	96
kremlin_bashnya	1	90
ia_steklomoy	1	87
finkrolik	1	94
ru_FTP	1	62
antiskrepa	1	74
margaritasimonyan	1	88
kaktovottak	1	91
ShaltayBabay	1	85
moscowmap	1	92
tinkoffjournal	1	30
krasniydom	2	99
mig41	2	93
akitilop	2	97
yoba_m	2	94
vibornyk	2	97
vesparevenge	2	93
Ivorytowers	2	97
rt_russian	2	96