УДК 004.415:620.9

DOI: https://doi.org/10.17721/3041-2323.2025.224-242

Тетяна ПРОСЯНКІНА-ЖАРОВА, д.т.н., доц.
ORCID ID: 0000-0002-9623-8771
e-mail: t.pruman@gmail.com

Олексій ШОЛОХОВ, к.ф.-м.н.
ORCID ID: 0000-0002-8676-3724
e-mail: gyroalex@knu.ua

Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ, Україна
Київський національний університет імені Тараса Шевченка, Київ, Україна

ІНТЕЛЕКТУАЛЬНА СИСТЕМ ВИЯВЛЕННЯ ІНФОРМАЦІЙНОГО ВПЛИВУ НА НАСЕЛЕННЯ КРАЇНИ-ЖЕРТВИ З БОКУ КРАЇНИ-ВІЙСЬКОВОГО АГРЕСОРА

У статті розглянуто створення системи, здатної до автоматичного виявлення та аналізу фейкових новин, зокрема визначення емоційної складової таких новин; виявлення першоджерел надходження повідомлень з такими новинами та шляхів їх розповсюдження. Спочатку аналізується цифровий простір взагалі – для виявлення його сегментів, що дотикаються українського цифрового простору чи проникають в нього, шляхів такого проникнення, і, далі, способи протидії. Отримані результати можуть бути використані для побудови комплексних систем пасивної та активної протидії у гібридній війни.

Ключові слова: ***цифровий простір, соціальні мережі, фейкові новини, аналіз тональності повідомлень, інформаційна боротьба, кібератака, соціальні графи, вузли впливу.***

Вступ

Дедалі більше людей усвідомлює, що ми живемо в епоху війни. Помітною її складовою є не фізичне знищення нападником населення країни, яку він хоче підкорити, а трансформація людей ментально на власну користь. Це інформаційна війна: у ЗМІ, соціальних мережах, кіно, телебаченні, комп'ютерних іграх. Це

тонка, маніпулятивна і непряма форма війни за проектування своєї версії колишніх і поточних подій, що закладаються поволі як у дитячі, так і у дорослі розуми – заміна смислових образів.

Інформаційна війна – це одна з форм інформаційного протиборства, комплекс заходів щодо інформаційного впливу на масову свідомість для зміни поведінки людей і нав'язування їм цілей, які не відповідають їхнім інтересам, а також, природно, захист від подібних впливів.

Як відомо, інформаційна війна – це дії, розпочаті для досягнення інформаційної переваги через завдання шкоди інформації та процесам, що базуються на інформації та інформаційних системах ворога при одночасному захисті власної інформації та процесів, що базуються на інформації та інформаційних системах. Основні методи інформаційної війни – блокування або перекручування інформаційних потоків і процесів прийняття рішень супротивником» (Горбулін, 2009, с.7). Передусім, що слід робити у такій війні, це викривати і спростовування фейки перед аудиторією і на випередження (Гороховський, 2018). Ефективна протидія в умовах інформаційно-гібридного протистояння ґрунтується, зокрема, на застосуванні аналогічних технологічних інструментів, які використовує противник, проте з протилежною функціональною спрямованістю та цільовим призначенням. Конкретним завданням дослідження є ідентифікація каналів і механізмів поширення дезінформації, а також розроблення підходів до її нейтралізації шляхом системної інтеграції сучасних інформаційних технологій.

Наукова новизна дослідження полягає у розробленні та теоретичному обґрунтуванні комплексної системи протидії інформаційно-психологічним операціям, спрямованим на інтенсифікацію суспільної напруги та формування деструктивних наративів громадського невдоволення. Запропонована система – це методологічно цілісна інтеграція відомих інструментальних засобів автоматизованого виявлення дезінформації, оцінювання рівня її впливу, ідентифікації джерел

поширення, а також формування адекватних механізмів реагування і нейтралізації.

Основна частина та результати

Єдиний інформаційний простір мережі Інтернет загалом не підлягає юрисдикції конкретної держави, тому пряма заборона ворожих ІТ-джерел або конфіскація серверів, на яких розташовані ворожі медійні ресурси, здебільшого юридично неможлива. Крім того, віртуальна сутність цього простору дозволяє легко уникати адресних заборон. Відповідно, засоби та методи протидії мають характеризуватися переважно віртуалізованим і проактивним характером, передбачаючи превентивне моделювання та прогнозування ворожих інформаційних впливів, зокрема сценаріїв поширення маніпулятивного контенту через різні комунікаційні канали. Такий підхід передбачає здійснення випереджувальних заходів реагування, що охоплюють системне викриття фальсифікованої інформації, оприлюднення верифікованих даних щодо реального перебігу подій, а також представлення повних повідомлень із належним контекстуальним супроводом.

Сучасний стан щодо цифрової гігієни та шкільною освітою щодо цього не є задовільним, а його зміна – повільний процес за визначенням та обумовлений факторами, якими не завжди можливо керувати. У демократичній країні неможливо обмежити доступ до цифрового простору і ментально шкідливого контенту, тому єдиний шлях зниження впливу такого контенту – створення у більшості населення імунітету до всіх видів ворожого інформаційного впливу. Цей імунітет створюється і забезпечується, як вже згадувалось, системою освіти та виховання. Звичайно, для неповнолітніх громадян доступ до певних інтернет-ресурсів можна обмежити юридично, але ж не для дорослих, оскільки це суперечить демократичним принципам. Крім того, введення та впровадження юридичних обмежень та освітніх процесів у систему освіти також потребує часу і тому основною зброєю залишається постійний пошук і виявлення ворожих інформаційних атак і у перспективі дія на випередження – проактивні контрзаходи.

Якщо стійке ментальне здоров'я населення – імунітет до маніпуляції свідомістю через інформаційний простір, – є *зброя стратегічна* в інформаційній війні, втілення відповідних юридичних та освітніх процесів – є *оперативний рівень*, то розглянута далі система – *тактична зброя*.

Досвід провідних держав світу (табл.1) свідчить про різні підходи до протидії інформаційним загрозам: США інвестують у ШІ для виявлення дезінформаційних кампаній, Китай застосовує стратегію «інтелектуалізації» з контролем інформаційних потоків, Ізраїль зосереджується на аналітиці даних і виявленні ботмереж.

Таблиця 1

Зарубіжний досвід протидії інформаційним загрозам

Держава	Ключові програми/ приклади	Інвестиції	Основний фокус	Результати
США	DARPA SMISC (Social Media in Strategic Communication); Detect Fakes (AI-алгоритми для фото/відео)	>$1 млрд/рік	Виявлення дезінформаційних кампаній, deepfake, OSINT	Система моніторингу соцмереж, технології детекції маніпуляцій
Китай	Великий державний цензорський апарат; використання LLM для аналізу інформаційних потоків	~$2 млрд/рік	Контроль соцмереж, виявлення «шкідливого контенту»	Побудова централізованих систем контролю інформації

Ізраїль	Cybereason, NSO Group (технології аналітики даних); алгоритми для моніторингу Telegram та X/Twitter	~$0.8млрд/рік	Кібербезпека, пошук бот-мереж, інформаційна розвідка	Системи для швидкої ідентифікації ботів та пропагандистських мереж
ЄС	EUvsDisinfo (стратегічна комунікація ЄС); EDMO (European Digital Media Observatory)	~$0.5млрд/рік	Протидія кремлівській пропаганді, аналіз медіа	База з 15k кейсів дезінформації, створення незалежних хабів з перевірки фактів
Україна	Центр протидії дезінформації (РНБО); Stop Fake; «Bellingcat-стиль» OSINT-розслідування SOC Prime	<$0.2млрд	Виявлення фейків, розслідування наративів, кіберзахист	96% кібератак заблоковано (2025); тисячі фейків спростовано та внесено в базу

Напрями, які вразливі до інформаційних загроз:

- **Розвідка та аналітика**. Фейки в супутникових знімках чи відео з БПЛА можуть дезорієнтувати командування. Неправильна класифікація об'єктів тягне за собою ризик хибних рішень (Zhang, 2024).

Вплив: зниження точності прогнозів, маніпуляції розвідданими.

- **Логістика**. Дезінформація про маршрути чи ресурси здатна паралізувати постачання. Фейкові повідомлення про поломки/брак ресурсів можуть зупинити операції (Tilley, 2024).

Вплив: зрив ланцюгів постачання, зайві витрати часу та коштів.

- **Кібербезпека**. Інформаційні кампанії маскують реальні кібератаки. Фейкові повідомлення про «витік даних» підривають довіру до систем (CGC: Cyber Grand Challenge, 2016).

Вплив: підвищення вразливості критичної інфраструктури.

- **Інформаційна боротьба**. Масове поширення фейків, bot-мереж і deepfake-матеріалів. Емоційні маніпуляції (страх, агресія, паніка) (Лозинська, 2025).

Вплив: розхитування суспільної стабільності та зниження морального духу.

Основні завдання виявлення та протидії рейковим новинам:

- виявлення фейкових новин – текстів, зображень, відео, аудіо, діпфейків;
- виявлення ботмереж – кластерізації акаунтів, автоматичних патернів;
- аналіз емоційності текстів – страх, агресія, пропаганда;
- антифекові дії у соцмережах – маркування, блокування, трекінг-поширення;
- розпізнавання діпфейків – зображень, відео, аудіо.

Для полегшення створення системи ненобхідно ввести визначення. Назвемо цифровою платформою окремий застосунок цифрового простору: сайт з форумом, месенджер з чатами, соціальну мережу з групами тощо. Цифрові платформи збирають і аналізують інформацію про всі види активності користувача і сприяють обміну отриманими даними, як віртуальними, так і реальними. Маються на увазі час і місце використання (відвідування), можливі записи та коментарі в соціальних мережах, різного роду «лайки», маршрути переміщення споживачів у реальному світі (геолокація), пошукові запити, покупки, перегляди того чи іншого контенту тощо. (SpiderFoot automates OSINT for threat intelligence and mapping your attack surface, 2025). Далі зібрані дані обробляються та аналізуються за допомогою сучасних інформаційних технологій штучного інтелекту таким чином, щоб отримані аналітичні висновки могли застосовуватися з метою,

поставленою бенефіціаром цього процесу, наприклад, інфлюенсєром ворожої країни. У табл. 2 наведені завдання системи.

Таблиця 2

Прототип системи аналізу фейкових новин

Збір даних	Автоматичне збирання інформації з відкритих джерел: соцмереж (Facebook, X/Twitter, Telegram, TikTok), новинних сайтів, форумів. Використовуються API та web-scraping. Дані проходять очистку від дублікатів, визначається мова та час публікації (Wang, Y., 2023).
NLP-аналіз текстів (фейк/наратив/топіки)	Система аналізує тексти на предмет маніпулятивних наративів. Виконується семантичне порівняння з базою відомих фейків та тематичне моделювання (LDA/BERT) для виявлення нових трендів. На цьому етапі отримується перша оцінка: чи текст схожий на фейковий.
Емоційний/афективний аналіз	Додатково виконується аналіз тональності повідомлень (позитивна, негативна, нейтральна). Система визначає рівень агресії, страху та закликів до дії, що дозволяє зрозуміти психологічний вплив повідомлення на аудиторію.
Виявлення каналів поширення/ботмереж	Створюється соціальний граф: користувачі, групи та канали. Аналізуються зв'язки між ними, виявляються підозрілі акаунти та бот-мережі. Визначаються ключові «вузли впливу», через які інформація поширюється найшвидше.
Трасування «ланцюжка поширення» (каскади)	Система відслідковує шлях повідомлення від початкового джерела до кінцевих репостів. Будується дерево поширення, визначається швидкість росту каскаду, час піку, ключові точки входу на різні платформи.
Візуалізація результатів/звіти	На виході формується інтерактивна карта поширення фейку з підсвіткою основних акторів, хвиль та каналів. Генерується список ключових поширювачів, а також автоматичні звіти для аналітиків і служб кіберзахисту.
Зворотний зв'язок та навчання з експертом (HITL/active learning)	HITL: аналітики підтверджують або спростовують фейки, додають ярлики наративів і працюють з низькою впевненістю. Це дозволяє донавчати модель, підвищуючи її точність і адаптивність.

Для розв'язання задачі – побудови системи аналізу та протидії, спочатку дослідимо цифровий простір (соцмережі, канали у соціальних месенджерах, форуми, чати) (Shodan Search Engine for the Internet of Everything, 2025) і процеси поширення інформаційних загроз у ньому (Wang, 2023). Для цього знадобляться такі методи та інструменти:

1. NLP: семантичне зіставлення з базами фейків, тематичне моделювання (LDA, BERT).
2. Sentiment Analysis: визначення емоційного тону та інтенсивності.
3. Graph Neural Networks: аналіз соціальних графів, виявлення вузлів впливу (Maltego Technical Documentation and User Guide, 2025).
4. Computer Vision: виявлення deepfake-матеріалів (pyLDAvis, 2025).
5. Active Learning (HITL): залучення експертів для донавчання системи.

Спочатку докладно розглянемо перші два пункти у їх застосуванні.

NLP – *Natural Language Processing* – обробка природної мови; LDA – *Latent Dirichlet Allocation* – приховане розподілення Діріхле; BERT – *Bidirectional Encoder Representations from Transformers* – представлення двонаправлених кодерів з трансформаторів.

На рис. 1 показано процес оброблення даних технологіями за першими двома пунктами (Topic Modeling BERT+LDA, 2025).

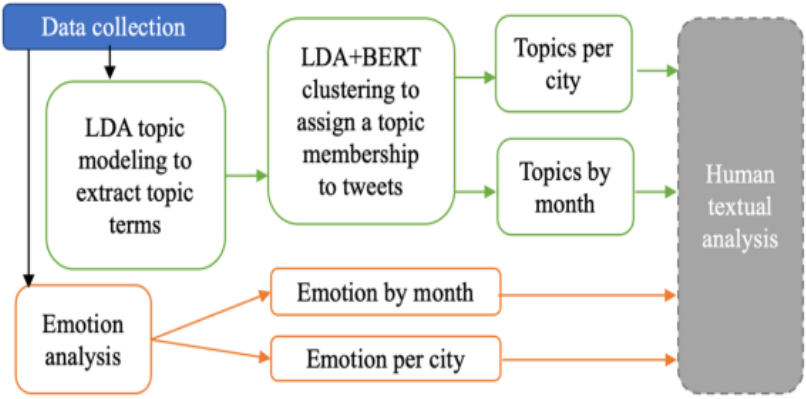

Рис.1. *Робочий процес аналізу даних*

Тематичне LDA-моделювання використовується для виявлення латентних семантичних просторів, або, іншими словами, основних тенденцій у текстових даних. LDA-модель – це генеративна ймовірнісна модель, що базується на словах, темах та документах. Вона обчислює розподіл слів за темами як представлення тем (латентні семантичні тенденції), а розподіл тем за документами – як представлення семантики документів стосовно виявлених «тем». Розроблювана система має вибирати таку кількість тем, яка генерує високий показник когерентності та створює теми, які можна інтерпретувати. Кількість тем – це гіперпараметр, який обирають експерти для ініціалізації моделі. Остаточне число тем обирається на основі аналізу когерентності та оцінки інтерпретованості людиною. Аналіз когерентності порівнює когерентність тематичних моделей з урахуванням різної кількості тем. Розмір теми занадто малий, щоб відобразити повторювану тему, її можна виключити з подальшої інтерпретації. Аналіз використовує NPMI – *Normalized Point Mutual Information* – нормалізовану поточкову взаємну інформацію та косинусну подібність для вимірювання тематичної когерентності. Інакше, NPMI – це нормалізована міра ймовірності спільної появи слів. Оцінка, виконана людиною, якісно визначить, чи є витягнуті ключові слова теми змістовними

та зв'язними. Тематична LDA-модель забезпечує розподіл ймовірностей повідомлення у соціальній мережі за всіма темами. Однак вона не призначає повідомлення до тематичної категорії. Кодування всіх повідомлень у тематичні категорії є надзвичайно трудомістким для людини. Тому для цього застосовано кластеризацію LDA+BERT для автоматичного кодування документів у тематичну категорію. Нагадаємо, що кластеризація *k*-Mean – це метод зменшення розмірності, який призначає документ одному з k кластерів. Мотивацією змішування вектора документів LDA та вектора речень BERT є усунення невід'ємної слабкості підходу BoW – *Bag of Words* – «мішок слів» у статистичному моделюванні (наприклад, тематичне моделювання LDA). BoW ігнорує залежність слів, яку неможливо виявити за допомогою таких методів, як TF-IDF або додаткової інженерії ознак. Сильна сторона тематичних моделей LDA полягає у використанні глобальної статистики корпусу для ідентифікації латентних семантичних просторів (тобто, тем). Однак підхід BoW ігнорує послідовну залежність, яка може впливати на значення слова в локальному контексті (тобто реченні). Тематична модель LDA може фіксувати найімовірніші слова з латентного семантичного простору, наприклад, «захід». Однак вона не надає жодної інформації про те, чи означає «захід» певну подію (свято, семінар, збори), чи географічну орієнтацію у «Я був на заході» чи «Новини надійшли з заходу». Ситуація значно ускладнюється внаслідок стійкої міжмовної інтерференції мови країни-агресора і української мови, та ще обтяжується спільними кириличними літерами у запису слів. Наприклад: *рос. уродливый (некрасивий) – укр. уродливий (красивий), рос. просторічне запамятовать (забути) – укр. запам'ятати (не забути); рос. рожа (морда, пика) – укр. рожа (мальва, троянда), рос. пыльный (прикметник до пыль) – укр. пильний (уважний, негайний, настійний, ретельний) та ін*. Оскільки кластеризація виконується для кожного документа, важливе локальне розуміння. Включення слів, вивчені моделями глибокого навчання, такими як Word2Vec

та BERT, використовують локальні контексти. Вони можуть вводити «локальні семантики» для заданої послідовності/речення. Їх можна використовувати для збалансування слабкості векторів документів, обчислених тематичними моделями LDA. Якщо попередньо навчити на великому корпусі включення BERT, тоді можна усунути неоднозначність полісемії на рівні речень. Конкретні кроки кластеризації LDA+BERT показано на рисунку 2.

Для створення контекстного представлення кожного повідомлення використовується не реєстроване попередньо навчене речення BERT з бібліотеки Hugging Face. Потім включення речень BERT, об'єднане з векторами речень LDA, вводиться в простий автокодер зі щільним шаром.

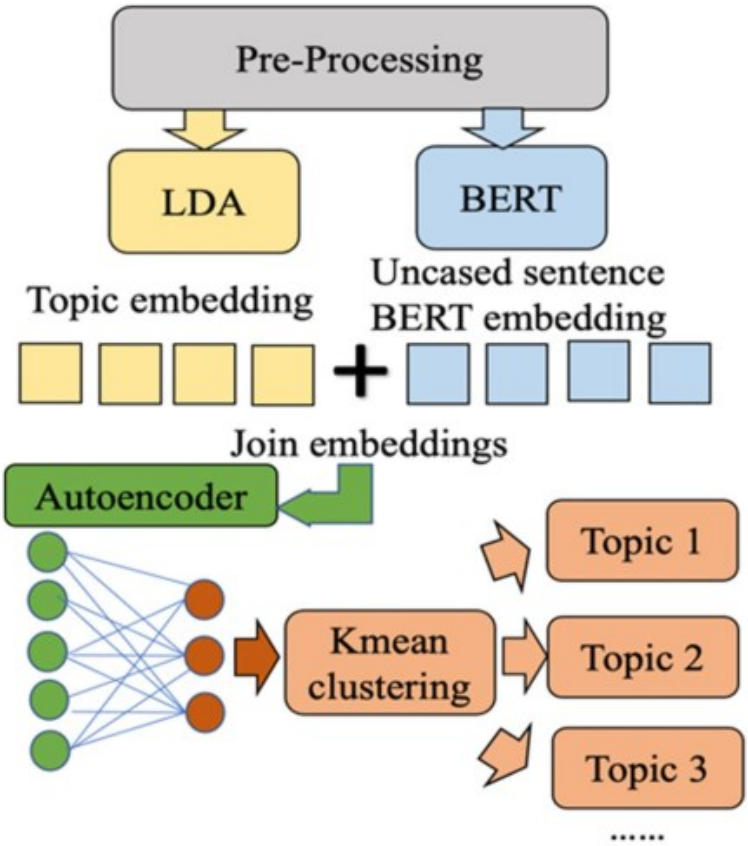

Рис.2. *Модель кластеризації LDA+BERT*

Далі кожне повідомлення кодується у представлення латентного векторного простору. Наприкінці до латентного представлення застосовується кластеризація методом *k*-Means і повідомлення включається у категорію теми. Використовувалась така ж кількість тем *k* в LDA для виконання кластеризації *k*-Mean. На рис. 2 показано архітектуру моделі «Аналіз емоцій». Мета аналізу емоцій полягає в тому, щоб дати аналітику-людині уявлення про настрій обговорень. Аналіз емоцій є розширенням аналізу настроїв. Розпізнавання емоцій допомагає глибше

зануритися в думки учасників дискусій (цільової аудиторії повідомлення), а також точніше розуміти розмови за допомогою моделей глибокого навчання.

Щодо реалізації BERT, тут використовується попередньо навчена модель BERT з бібліотеки Hugging Face. Ця модель має різні завдання класифікації повідомлень, включаючи настрої, токсичність та емоції. Завдання емоцій має чотири мітки: «радість», «оптимізм», «сум», «гнів» Щоб позначити повідомлення, розміщені у відповідної соціальної мережі, варто проаналізувати кожне з них відповідно до моделі. Виходом моделі є єдина мітка зі списку емоцій. Мета аналізу тексту, проведеного людиною, що базується на темах та емоціях, заданих NLP-моделями, полягає в тому, щоб запропонувати уточнену інтерпретацію для кожного територіального утворення або групи користувачів. Текстовий аналіз дозволяє виявити локальну перспективу з повідомлень за загальною кількісною оцінкою. На рис. 3 у вигляді блок-схеми представлено один з можливих варіантів методики виявлення позитивної чи негативної реакції на певні повідомлення певних груп у соціальних мережах (Dehghani, 2024).

Було побудовано систему з застуванням технологій: LLM (GPT, LLaMA, Mistral) – для класифікації текстів і наративів; Sentiment Analysis – для аналізу емоційності; Graph Neural Networks (GNN) – для моделювання поширення інформації; Computer Vision (CNN, GAN-detectors) – для виявлення діпфейків. Було отримано такі KPI (табл.3) та результати, зокрема, карта поширення фейку (мережевий граф з основними вузлами); список ключових акторів (бот-мережі, інфлюенсери, первинні джерела); автоматичні звіти для аналітиків та кіберзахисників.

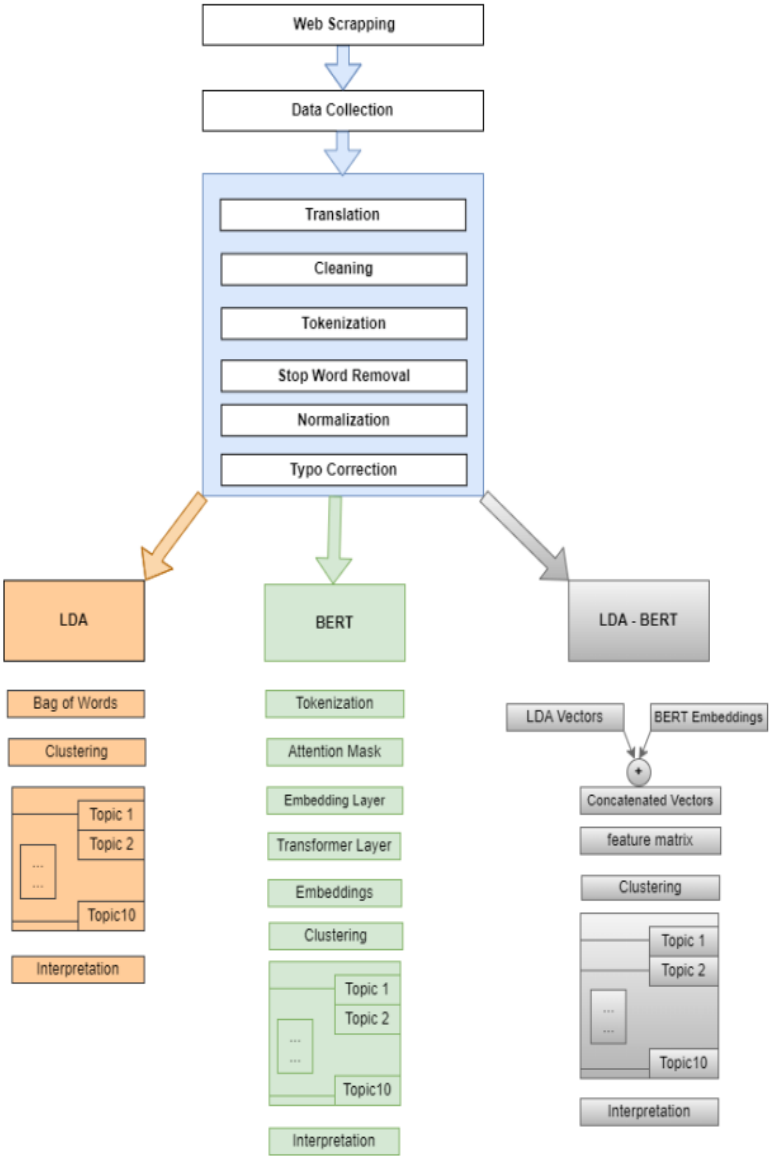

Рис.3. *Графічне представлення методології, закладеної у систему*

Таблиця 3

Антифейкова система: технології, результати, KPI

Модуль	Ключові метрики
Збір даних	Coverage ≥ 90%; Lag до індексації ≤ 60 сек; Deduplication ratio 15–40%
NLP-аналіз текстів	Macro-F1 ≥ 0.85; Recall@10 ≥ 0.9; Topic coherence ≥ 0.5
Емоційний аналіз	F1 для емоцій ≥ 0.80; ROC-AUC (токсичність) ≥ 0.95; Drift alert > 2σ
Виявлення каналів поширення	Точність детекції ботів ≥ 90%; Modularity Q ≥ 0.4; % підозрілих акаунтів у топ-10%
Трасування, поширення	R_0 (репродукція); Cascade depth/breadth; Time-to-detection ≤ 5 хв
Візуалізація результатів	Alert latency ≤ 2 хв; Uptime ≥ 99.5%; MTBFA ≥ 7 днів
HITL/Active Learning	ΔF1 +1–3 п.п.; Throughput ≥ 60 зразків/год; узгодженість κ ≥ 0.75

Рис.4. *Прототип дизайну системи*

Дискусія і висновки

У роботі показано створення автоматичної системи виявлення та розповсюдження шейків. Для полегшення сприйняття наведено візуалізацію. Зазначимо, що у разі великої кількості отримувачів фейків візуалізація втрачає інформативність – головним стає табличне виведення джерел фейків з певними атрибутами. Джерел, зазвичай, декілька, відповідно атрибути – назви груп, імена блогерів, адреси сайтів, уможливлюють блокування таких ресурсів або ведення контрпропаганди.

Основне досягнення – це методологія аналізу та виявлення, яка дозволяє застосовувати різні інструменти, що зручно для аналітиків, які користуються іншими програмними застосунками, зокрема власної розробки.

Отримані результати підтверджують застосовність методології та ефективність прототипу системи.

Список використаних джерел

Горбулін, В.П., Додонов, О.Г., Ланде, Д.В. (2009). *Інформаційні операції та безпека суспільства: загрози, протидія, моделювання*. Київ: Інтертехнологія.

Гороховський, О. (2018). *Фактчек як тренд розслідувань: можливості та перспективи*. Прага: KUFR

Лозинська, О.В., Марків, О.О., Висоцька, В.А. (2025). Метод виявлення розповсюджувачів дезінформації на основі графового представлення структури соціальної мережі. *Центрально-український науковий вісник. Технічні науки*, *11*(42), 70-78. https://doi.org/10.32515/2664-262X.2025.11(42).2.70-78

CGC: Cyber Grand Challenge (2016). https://www.darpa.mil/research/programs/cyber-grand-challenge

Dehghani, F., & Zaman, L. (2024). Exploring Players' Perspectives: A Comprehensive Topic Modeling Case Study on Elden Ring. *Information*, *15*(9), 573. https://doi.org/10.3390/info15090573

Maltego Technical Documentation and User Guide (2025). https://docs.maltego.com/en/support/home

pyLDAvis (2025). https://pyldavis.readthedocs.io/en/latest/readme.html

Shodan Search Engine for the Internet of Everything (2025). https://www.shodan.io/

SpiderFoot automates OSINT for threat intelligence and mapping your attack surface (2025). https://github.com/smicallef/spiderfoot

Tilley, S. (2024). Smart Logistics: Navigating the AI Frontier in Sustainment Operations. *Army Sustainment Magazine, October 17, 2024.* https://www.army.mil/article/280377/

Topic Modeling BERT+LDA (2025) https://www.kaggle.com/code/dskswu/topic-modeling-bert-lda

Wang, Y., Willis, E., Yeruva, V.K. et al. (2023). *A case study of using natural language processing to extract consumer insights from tweets in American cities for public health crises*. BMC Public Health 23, 935. https://doi.org/10.1186/s12889-023-15882-7

Zhang, Li Ang, Yusuf Ashpari, and Anthony Jacques (2024). *Understanding the Limits of Artificial Intelligence for Warfighters. Volume 3, Predictive Maintenance*. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/research_reports/RRA1722-3.html

References

Horbulin, V.P., Dodonov, O.H., Lande, D.V. (2009). *Informatsiini operatsii ta bezpeka suspilstva: zahrozy, protydiia, modeliuvannia*. Kyiv: Intertekhnolohiia.

Horokhovskyi, O. (2018). *Faktchek yak trend rozsliduvan: mozhlyvosti ta perspektyvy*. Praha: KUFR

Lozynska, O.V., Markiv, O.O., Vysotska, V.A. (2025). Metod vyiavlennia rozpovsiudzhuvachiv dezinformatsii na osnovi hrafovoho predstavlennia struktury sotsialnoi merezhi. *Tsentralno-ukrainskyi naukovyi visnyk. Tekhnichni nauky*, 11(42), 70-78. https://doi.org/10.32515/2664-262X.2025.11(42).2.70-78

CGC: Cyber Grand Challenge. https://www.darpa.mil/research/programs/cyber-grand-challenge

Dehghani, F., & Zaman, L. (2024). Exploring Players' Perspectives: A Comprehensive Topic Modeling Case Study on Elden Ring. *Information*, *15*(9), 573. https://doi.org/10.3390/info15090573
Maltego Technical Documentation and User Guide (2025). https://docs.maltego.com/en/support/home
pyLDAvis (2025). https://pyldavis.readthedocs.io/en/latest/readme.html
Shodan Search Engine for the Internet of Everything. https://www.shodan.io/
SpiderFoot automates OSINT for threat intelligence and mapping your attack surface (2025). https://github.com/smicallef/spiderfoot
Tilley, S. (2024). Smart Logistics: Navigating the AI Frontier in Sustainment Operations. *Army Sustainment Magazine, October 17, 2024.* https://www.army.mil/article/280377/
Topic Modeling BERT+LDA (2025) https://www.kaggle.com/code/dskswu/topic-modeling-bert-lda
Wang, Y., Willis, E., Yeruva, V.K. et al. (2023). *A case study of using natural language processing to extract consumer insights from tweets in American cities for public health crises*. BMC Public Health 23, 935. https://doi.org/10.1186/s12889-023-15882-7
Zhang, Li Ang, Yusuf Ashpari, and Anthony Jacques (2024). *Understanding the Limits of Artificial Intelligence for Warfighters. Volume 3, Predictive Maintenance*. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/research_reports/RRA1722-3.html

Отримано редакцією журналу / Received: 26.09.25
Прорецензовано / Revised: 27.09.25
Схвалено до друку / Accepted: 01.10.25

Tetiana PROSIANKINA-ZHAROVA, DSc (Techn.), associate professor
ORCID ID: 0000-0002-9623-8771
e-mail: t.pruman@gmail.com

Oleksii SHOLOKHOV, candidate of phys.-math. science
ORCID ID: 0000-0002-8676-3724
e-mail: gyroalex@knu.ua

Institute of Telecommunications and Global Information Space of The National Academy of Sciences of Ukraine
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

DEVELOPMENT OF AN INTELLECTUAL SYSTEM FOR DETECTION OF INFORMATION IMPACT ON THE POPULATION OF THE VICTIM COUNTRY BY THE MILITARY AGGRESSOR COUNTRY

The article considers the creation of a system capable of automatically detecting and analyzing fake news, in particular, determining the emotional component of such news; identifying the primary sources of messages with

such news and ways of their distribution. First, the digital space in general is analyzed - to identify its segments that touch or penetrate the Ukrainian digital space, the ways of such penetration, and then methods of counteraction. The results obtained can be used to build complex systems of passive and active counteraction in hybrid warfare.

Keywords: digital space, social networks, fake news, message tone analysis, information warfare, cyberattack, social graphs, nodes of influence.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.