



DOI 10.28925/2663-4023.2025.28.862

УДК 004.738.5:351

Васильєва Ольга Олександрівна

аспірантка

Національний університет «Ченівська політехніка», Чернівці, Україна

ORCID ID: 0000-0002-8565-0525

olga.vasiljeva37@gmail.com

ТИПОЛОГІЯ АГЕНТІВ ТА ЇХ ВЛИВ НА ПОШИРЕННЯ ДЕЗІНФОРМАЦІЇ В СОЦІАЛЬНІЙ МЕРЕЖІ X

Анотація. Останнім часом з'являється чимало досліджень, що аналізують роль соціальних мереж у різних типах комунікації, зокрема у поширенні інформаційних впливів, таких як дезінформація, фейки та спрямовані інформаційні кампанії. Такі явища становлять загрозу інформаційній безпеці будь-якої держави. Наразі фіксується активне використання мережі X (Twitter) для впливу як на європейський, так і на український сегменти користувачів соціальних мереж. Враховуючи важливість цього питання, дане дослідження було спрямоване на аналіз алгоритму поширення контенту з урахуванням різних типів користувачів (так званих «агентів»). У межах дослідження проведено аналіз чотирьох типів агентів і визначено їхні ключові характеристики, що ймовірно визначають рівень впливу під час інформаційної операції, який, своєю чергою, є основним критерієм при побудові імітаційної моделі поширення контенту в соціальній мережі.

Ключові слова: інформаційна безпека; імітаційне агентне моделювання; соціальні мережі; дезінформація; інформаційні операції.

ВСТУП

Зростаюче значення соціальних мереж у формуванні суспільного дискурсу пояснює підвищений інтерес до аналізу того, як відбуваються інформаційні впливи та поширення дезінформації в онлайн-середовищі. Актуальність дослідження зумовлюється двома основними чинниками. По-перше, використання ботів, тролів та інших інструментів для маніпулювання громадською думкою стало справжнім викликом для інформаційної безпеки, оскільки такі технології впливають на політичні процеси, громадську довіру та якість суспільних дискусій. По-друге, зростає кількість звичайних користувачів, які мимоволі поширюють неперевірену або маніпулятивну інформацію, посилюючи ефект дезінформаційних кампаній. За таких умов постає потреба у детальному аналізі динаміки поширення контенту з урахуванням різноманіття типів учасників соціальної мережі – від лідерів думок і звичайних користувачів до автоматизованих та деструктивних агентів.

Постановка проблеми. Більшість існуючих робіт зосереджується або на формальних ознаках поширення (кількість постів, хештегів, ретвітів), або на детальному аналізі одного чи двох типів агентів (приміром, ботів і тролів), без урахування комплексної взаємодії з іншими типами користувачів. Натомість у запропонованому дослідженні робиться акцент на чотирьох відмінних типах агентів (лідери думок, боти, тролі та звичайні користувачі), які суттєво впливають на динаміку поширення контенту й інформаційні потоки в мережі X (Twitter). Зазначені вище підходи закладають важливе методологічне підґрунтя, але не дають комплексного уявлення про багатокomпонентну систему соціальної мережі, де кожен тип агента має свою специфічну роль. Таким чином, **унікальність та новизна** цього дослідження полягає в тому, що воно пропонує:



1. Сформувані основні характеристики чотирьох типів агентів за визначеними показниками для їх ідентифікації в мережі.
2. Розрахувати коефіцієнт впливу для кожного типу агента.

Завдяки такому підходу можна отримати більш точне уявлення про механізми поширення дезінформації та впливу, що, своєю чергою, відкриває перспективи для розроблення ефективніших стратегій протидії інформаційним загрозам.

Аналіз останніх досліджень і публікацій. Останніми роками з'явилася низка робіт, присвячених вивченню поширення контенту та інформаційних операцій у соціальних мережах. Так Ferrara, E. (2017) [1] у своєму дослідженні дослідив роботу ботів у X (Twitter) під час президентської виборчої кампанії у Франції 2017 року, виявивши, що автоматизовані акаунти впливали на ключові етапи суспільного обговорення та потенційно сприяли поширенню дезінформації. Bessi, A., & Ferrara, E. (2016) [2] показують, як боти в X (Twitter) брали участь у дискусіях щодо виборів у США 2016 року, та демонструє масштаби їхнього впливу. Дослідженням впливу фейкових новин у соціальних мережах приділяли увагу автори Bovet, A., & Makse, H. A. (2019) [3], які також проаналізували механізми поширення фейкових новин під час виборчої кампанії у США 2016 року, звертаючи увагу на мережеву структуру поширення контенту та роль окремих впливових користувачів. Також звертає на себе увагу одне з найвідоміших досліджень теми фейків і дезінформації: автори Vosoughi, S., Roy, D., & Aral, S. (2018) [4] аналізують величезний обсяг даних із X (Twitter) і виявляють, що фейкові новини поширюються швидше та охоплюють ширшу аудиторію, ніж правдиві. Роль лідерів думок та інфлюенсерів в алгоритмі розповсюдження контенту досліджували Aral, S., & Walker, D. (2012) [5]. Хоча робота не концентрується виключно на дезінформації, автори пропонують методологію виявлення ключових інфлюенсерів (лідерів думок) у соціальних мережах, що важливо й для аналізу інформаційних операцій. У роботі González-Bailón, S., Borge-Holthoefer, J., & Moreno, Y. (2013) [6] проаналізовано, як відбувається поширення протестних настроїв у соціальних мережах, зокрема виділено користувачів, які мають прихований вплив на великий масштаб мережі. У контексті референдуму щодо Brexit автори Howard, P. N., & Kollanyi, B. (2016) [7] демонструють, як боти у X (Twitter) впливали на громадську думку, поширюючи повідомлення із визначеними хештегами та гаслами. Starbird, K., Arif, A., & Wilson, T. (2019) [8] висвітлюють роль тролів та координованих груп користувачів, які поширюють дезінформацію. Автори підкреслюють колективний характер таких інформаційних операцій.

Метою статті є визначити ключові характеристики лідерів думок, ботів, тролів і звичайних користувачів, які являються маркерами їх ідентифікації, розрахувати коефіцієнт впливу для кожного типу агента та провести аналіз визначених параметрів, що дозволить виокремити найбільш вагомі характеристики для кожного типу агентів та визначити їх вплив на ефективність інформаційної операції в соціальних мережах.

Завданнями дослідження є:

1. сформувані основні характеристики чотирьох типів агентів за визначеними показниками для їх ідентифікації в мережі;
2. розрахувати коефіцієнт впливу для кожного типу агента;

Завдяки такому підходу можна отримати більш точне уявлення про механізми поширення дезінформації та впливу, що, своєю чергою, відкриває перспективи для розроблення ефективніших стратегій протидії інформаційним загрозам.



РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Сучасні соціальні мережі набули статусу потужного каналу комунікації, перетворившись на середовище, де мільйони користувачів щоденно обмінюються інформацією та формують громадську думку. Водночас швидкість і масштаб поширення контенту в цих мережах роблять їх сприятливим майданчиком для здійснення інформаційних операцій, спрямованих на маніпуляцію, дезінформацію та вплив на широкі верстви населення. У цьому контексті мережа X (Twitter) привертає дедалі більшу увагу дослідників, оскільки там активно розгортаються як локальні, так і міжнародні інформаційні кампанії.

Різноманітні за природою учасники комунікації у соціальних мережах – від звичайних користувачів і лідерів думок до ботів та тролів – мають різні показники активності та здійснюють неоднаковий вплив на інформаційне середовище. Саме тому агентне моделювання, що дозволяє деталізовано аналізувати поведінку та взаємодію різних типів агентів, стає актуальним підходом до дослідження феномену поширення контенту.

Слід зазначити, що для більш структурованого підходу моделювання інформаційних операцій в соціальних мережах допускається враховувати ряд інших факторів, таких як зовнішні чинники та властивості контенту. Хоча в даному дослідженні при агентному моделюванні основними параметрами для моделі будуть визначені показники активності різних типів агентів та їхня взаємодія між собою.

Аналізу робіт авторів (Vosoughi, S., Roy, D., & Aral, S. (2018).) [4], (Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016).) [9], (Berger, J. (2013) [10] дає змогу виокремити такі властивості контенту як ключові фактори успішного чи масового поширення контенту в соціальних мережах:

- значимість
- актуальність
- неодноразовість
- повнота
- новизна
- швидкість його поширення
- віральність (вірусність)

Можемо припускати, що для використання в інформаційній операції, контент має бути обов'язково значимим та актуальним, і це вплине на показник швидкості поширення. Новизна може бути, але це не обов'язковий показник. За часту, як видно із сьогодення може бути пролонгована інформаційна операція з використанням сталих наративів, і новизну може становити лише 10% контенту, якій використано як привід для пролонгації зацікавленості користувачів темою, що наводить нас на інший показник – неодноразовість. Беручи до уваги техніки та тактики підготовки наративів, що розповсюджуються, показник повнота не є необхідним, а скоріше навпаки, чим коротше повідомлення, тим швидше воно розповсюджується та охоплює більший масив цільової аудиторії, тобто більшу кількість користувачів.

Фактор зовнішніх чинників під час проведення інформаційної операції можемо сприймати як константу, оскільки такі або постійно з'являються в силу об'єктивних причин, або стимулюються організаторами інформаційної операції. Вони включають: існування зовнішніх факторів впливу (політичні події, заяви, зустрічі), існування зовнішніх агентів впливу (публікації в засобах масової інформації) та наявність стадій –



характерних етапів інформаційної операції (наприклад, процесу дифузії інновацій, каскади).

В попередніх публікаціях [11] було запропоновано виділити чотири типи агентів для подальшої побудови агентної моделі: лідери думок, боти, тролі та звичайні користувачі. З метою ідентифікації цих типів агентів в мережі наводяться характеристики для кожного типу за визначеними показниками [12]: непрямі зв'язки (Indirect links), частота взаємодій (Frequency), стабільність зв'язків (Stability), множинність (Multiplexity), сила зв'язку (Strength), напрямок (Direction), симетрія (Symmetry). При цьому теоретично можемо співвіднести кожен з наданих характеристик з чотирма визначеними типами агентів.

– **Перший тип агентів – лідери думок**, які публікують контент для подальшої популяризації. Цьому типу притаманні такі характеристики як авторитетність, центральність, значимість думок, високий ступінь до впливу на аудиторію та високий ступінь прийняття рішень в умовах невизначеності. Лідери думок публікують (генерують) власний контент, або використовують контент із зовнішнього джерела.

Indirect links (непрямі зв'язки): часто мають розгалужену «мережу впливу» через посередників, їхні публікації ретвітять/поширюють інші користувачі, які, у свою чергу, далі транслюють ці думки, що створює високий рівень непрямих зв'язків.

Frequency (частота взаємодій): зазвичай середня або висока, це спричинено кількістю регулярних публікацій та реакцією на запити аудиторії.

Stability (стабільність зв'язків): мають доволі стабільне коло підписників.

Multiplexity (множинність): висока ймовірність кількох рівнів контакту, тобто цей тип агентів може взаємодіяти з аудиторією одночасно в різних соцмережах (Twitter/X, Instagram, Facebook, YouTube). Мультиплексність часто висока.

Strength (сила зв'язку): має сильні зв'язки із «близьким оточенням» (команда, колеги), але з підписниками – «слабкі» або «помірні» зв'язки. Можна стверджувати про полярний розподіл сили зв'язків: мало дуже сильних, багато слабких.

Direction (напрямок): здебільшого односторонній у плані «фолловер → лідер думок» чи «лідер думок → аудиторія». Кількість людей, які читають лідера думок, значно перевищує кількість, за якими стежить він сам. Водночас, коли лідер думок згадує когось (відповідає), це створює спрямований канал впливу з високим резонансом.

Symmetry (симетрія): у зв'язках «лідера думок» і масової аудиторії асиметрія переважає: більшість зв'язків не є взаємними (follower→followed).

– **Другий тип агентів – акаунти-боти** (спеціальна програма, що виконує автоматично і за заданим розкладом певні дії через ті ж інтерфейси, що й звичайний користувач). Використовуються для штучного нарощування «лайків» чи коментарів з метою підняття рейтингу публікації серед користувачів мережі, дають змогу посилити вплив у інформаційному середовищі шляхом залучення нових цільових аудиторій (бот поширює новину, але майже нічого не змінює, він на відміну від троля є механічним, троль – креативним).

Indirect links (непрямі зв'язки): можуть породжувати непрямі зв'язки, особливо якщо бот «читує» чи поширює контент інших і таким чином запускає «ланцюжковий» ефект. Але часто ці зв'язки залишаються слабо вираженими або мають низьку довіру.

Frequency (частота): як правило, висока (або дуже висока) за рахунок автоматичного постингу великої кількості повідомлень, часто за шаблоном.

Stability (стабільність): діяльність бота може бути стабільною (довготривала й послідовна) або епізодичною (кампанійного типу). Однак стабільність особистих



«зв'язків» у ботів низька, адже вони рідко формують тривалі чи «емоційні» зв'язки з реальними користувачами.

Multiplexity (множинність): як правило, низька, оскільки боти зазвичай працюють на одній або кількох пов'язаних платформах, але не схильні до багаторівневої взаємодії.

Strength (сила зв'язку): здебільшого слабкі зв'язки. Контакт із ботом не ґрунтується на довірі чи емоційній близькості, тож «сила» таких зв'язків незначна.

Direction (напрямок): Боти часто поширюють контент (*outbound direction*) і мають мало вхідних зв'язків. Водночас інші користувачі можуть підписуватися чи не підписуватися на бота, тож є асиметричний спрямований канал.

Symmetry (симетрія): здебільшого асиметрія – рідкісні випадки, коли бот «свідомо» читає (авторизує, відповідає) інших. Навіть якщо бот фолловить когось, це зазвичай автоматизоване й не означає «справжнього» взаємного зв'язку.

– **Третій тип агентів** – акаунти-тролі – звичайна дійова особа, яка надає «кольорового забарвлення» публікаціям. Вони оперують великими групами і подають новину з різних ракурсів. За цим типом агентів можуть знаходитись блогери, лідери думок, які репостять «вірусний» контент на власні сторінки, тим самим одразу (в один клік) збільшуючи потенційні цільові аудиторії (перегляди, лайки, репости) в рази. Тролі за часту коментують контент, тим самим залучаються до активної багатосторонньої комунікації. Наявність специфічних соціальних норм відсутня, але наявна висока ігрова взаємодія агентів.

Indirect links (непрямі зв'язки): можуть виникати значні непрямі ефекти, оскільки тролі часто провокують конфлікти, які потім «відлунюються» через репости інших. Втім, троль зазвичай свідомо не розвиває непрямі ланцюжки, та це може статись «автоматично» через обговорення.

Frequency (частота): середня або висока – тролі зазвичай дуже активні, але діяльність може бути нерівномірною (спалахи активності в певних темах).

Stability (стабільність): низька. Тролі часто змінюють теми, можуть переходити з одного каналу дискусії в інший, створюють нові акаунти. Стійких зв'язків із «жертвами» чи з аудиторією не підтримують.

Multiplexity (множинність): залежить від «стратегії» тролінгу – іноді це лише один майданчик, іноді – кілька (Twitter, Facebook, форуми). Проте емоційно глибоких чи комплексних зв'язків зазвичай немає.

Strength (сила зв'язку): переважно слабка, але з високим рівнем «емоційного збурення»; тобто троль може викликати сильну негативну реакцію, проте це не «сила зв'язку» у класичному розумінні довіри чи дружби.

Direction (напрямок): тролі найчастіше «нападають» або провокують (*outbound direction*), а інші реагують на них. Тому в них наявна значна асиметричність у комунікації.

Symmetry (симетрія): дуже рідко зустрічається симетрична взаємодія; зв'язки зазвичай односторонні (тролеві відповідають, але сам троль не прагне «порозуміння»). «Взаємне читання» може бути, якщо тролі об'єднуються в групу, але це швидше виняток.

– **Четвертий тип агентів** – реальні люди, які є потенційною цільовою аудиторією інформаційної операції, вони «споживають» контент та сприймають його достовірним, тобто підпадають під інформаційний та психологічний вплив і стають «інфікованими» даним контентом, вірять, що інформація правдива та популяризують її далі в мережі, базуючись на своїй вірі.

Indirect links (непрямі зв'язки): у типового користувача є зв'язки зі знайомими й друзями, які далі можуть бути пов'язані з іншими, тож непрямі ланцюжки можуть бути



досить довгими (ефект «шести рукоприкладств»). Здебільшого помірна кількість непрямих зв'язків.

Frequency (частота): сильна варіативність від дуже низької (рідко пишуть) до досить високої (активні блогери). У середньому – помірна.

Stability (стабільність): зазвичай помірна. Основні дружні або родинні зв'язки доволі стабільні, але онлайнова активність може змінюватися з часом (хтось перестає користуватися соцмережею або заводить нові акаунти).

Multiplexity (множинність): у реальному житті люди взаємодіють і онлайн, і офлайн. Тож для близьких друзів чи колег множинність каналу висока. В середньому рівень мультиплексності цього типу агентів вищий, ніж у ботів, але нижчий, ніж у лідерів думок із багатьма медіаплатформами.

Strength (сила зв'язку): у більшості зв'язків із друзями, близькими знайомими – помірна або сильна. Із «випадковими» фолловерами чи онлайн-знайомими – слабка. Тобто спостерігається полярне розподілення за силою зв'язку, але масштаби менші, ніж у лідерів думок.

Direction (напрямок): багато зв'язків є взаємними (undirected) або двосторонніми (друзі читають одне одного). Утім, бувають і односторонні стосунки (наприклад, коли хтось читає знайомого, а той не фолловить у відповідь).

Symmetry (симетрія): зазвичай відносно висока (друзі, родичі найчастіше «фоловлять» чи «дружать» одне з одним). Але симетричність не абсолютна – особливо якщо йдеться про підписки на відомих блогерів чи інфлюенсерів.

На основі емпіричних досліджень проаналізовано три варіанти розповсюдження контенту в соціальній мережі X (Twitter). Під час дослідження на основі запропонованих характеристик агентів визначено, які саме агенти і в який проміжок часу долучаються до розповсюдження визначеного контенту.

У результаті досліджень поширення інформації в соціальній мережі X (Twitter) визначено, що процес відбувається у три фази, залежно від часу, що минув із моменту публікації повідомлення.

Перша фаза охоплює перші 60 хвилин після публікації. У цей період поширення відбувається переважно серед підписників лідера думок, які складають його безпосередню аудиторію.

Друга фаза починається після першої години і триває до 10 годин після публікації. У цей період відбувається поділ між всіма агентами розповсюдження інформації. Поширення інформації в цей час визначається коефіцієнтами впливу кожного агента, що забезпечує експоненційне чи лінійне зростання охоплення.

Третя фаза настає після 10 годин до визначеного в дослідженні часу, у нашому випадку 24 години. Основну роль відіграють звичайні користувачі, оскільки вплив ботів і тролів поступово зменшується і вже не має статистично значущого впливу на подальше поширення інформації.

Основні початкові параметри:

lom_people – змінна, яка відповідає значенню кількості підписників лідера думки;

lom_count – змінна, яка відповідає значенню кількості розповсюдження серед підписників лідера думки;

bot_count – змінна, яка відповідає значенню кількості розповсюдження завдяки ботам;



$troll_count$ – змінна, яка відповідає значенню кількості розповсюдження завдяки троллям;

$person_count$ – змінна, яка відповідає значенню кількості розповсюдження завдяки звичайним користувачам;

k_person – відсоток кількості звичайних користувачів з кількості підписників лідера громадської думки;

k_bot – відсоток кількості ботів з кількості підписників лідера громадської думки;

k_troll – відсоток кількості тролів з кількості підписників лідера громадської думки;

a – коефіцієнт, який відображає розповсюдження інформації від ботів на користувачів соціальної мережі у 2 фазі;

q – коефіцієнт, який відображає розповсюдження інформації від тролів на користувачів соціальної мережі 2 фазі;

b – коефіцієнт, який відображає розповсюдження інформації від звичайних користувачів на користувачів соціальної мережі;

v – коефіцієнт, який відображає розповсюдження інформації від ботів на користувачів соціальної мережі 3 фазі;

m – коефіцієнт, який відображає розповсюдження інформації від тролів на користувачів соціальної мережі 3 фазі.

Ця структура дозволяє чітко відмежувати етапи інформаційної динаміки та використовувати окремі математичні моделі для кожного з них.

1. Лідери думок (Opinion Leaders).

Роль у соціальній мережі:

- Впливають на думку великої кількості підписників.
- Часто є експертами або авторитетами в певній тематиці.
- Генерують унікальний контент, що активно поширюється іншими користувачами.

Ключові характеристики:

1) Кількість підписників lom_people – чим більше аудиторія, тим більший вплив.

2) Рівень взаємодії з аудиторією En – середня кількість лайків, коментарів, репостів на одиницю контенту.

3) Частота публікацій Ps – середня кількість постів на добу.

4) Довжина життєвого циклу поста Pl – період активного обговорення (в годинах).

5) Коефіцієнт впливу α – оцінка рівня довіри до контенту на основі взаємодій.

Для визначення кількості підписників, які фактично взаємодіють із контентом лідера думки, використовуємо наступну формулу, врахувавши коефіцієнт масштабування K :

$$lom_count = lom_people * (1 - e^{-(Ps * Pl * En * \alpha) / K})$$

Коефіцієнт впливу l лідера думок на користувачів в соціальній мережі визначаємо як:



$$l = lom_count / lom_people$$

2. Боти (Bots).

Роль у соціальній мережі:

- Автоматизовані акаунти, що генерують або поширюють контент.
- Використовуються для масового розповсюдження інформації або дезінформації.
- Часто мають низьку взаємодію з аудиторією, але високу частоту публікацій.

Ключові характеристики:

- 1) Частота публікацій Pf – кількість постів у день (зазвичай висока).
- 2) Кількість репостів Nr – середня кількість репостів на день.
- 3) Коефіцієнт автоматизації β – частка контенту, що є автоматично згенерованим.
- 4) Рівень взаємодії Li – середня кількість відповідей на коментарі (зазвичай низький).
- 5) Стійкість до блокування Br – оцінка ймовірності, що бот не буде видалений.

За ключовими характеристиками визначимо коефіцієнт впливу ботів a в окремо взятому посту. Кількість ботів на пряму залежить від рівня взаємодії Li , та визначається з кількістю однотипних коментарів Liz :

$$k_bot = Li / Liz$$

Коефіцієнт впливу ботів a визначаємо як (розповсюджується на 2 фазу):

$$a = Pf * Nr * \beta$$

Емпіричним шляхом встановлено, що кількість репостів Nr , які несуть дійсний інформаційний вплив Nrd на залучення аудиторії, становить 20%:

$$Nr = Nrd * 20 / 100$$

На 3 фазі починається виявлення та блокування ботів, тому коефіцієнт впливу a змінюється на коефіцієнт впливу v :

$$v = a * Br$$

Коефіцієнт блокування залежить від часу та визначається за формулою:

$$Br(t) = \exp(-Br(t-60)), t > 60$$

3. Тролі (Trolls).

Роль у соціальній мережі:



- Провокують конфлікти, дискредитують інформацію, поширюють дезінформацію.

- Орієнтовані на емоційні реакції інших користувачів.
- Мають середню або високу взаємодію через провокаційний контент.

Ключові характеристики:

1) Кількість конфліктних коментарів N_{cc} – кількість негативних або провокаційних коментарів.

2) Рівень залученості Loi – середня кількість реакцій на контент.

3) Частота атак Af – кількість провокаційних дій у день.

4) Тривалість активності Ad – середній час, протягом якого троль активний у мережі.

5) Коефіцієнт агресії δ – рівень токсичності коментарів (оцінка за NLP-аналізом).

За ключовими характеристиками визначимо коефіцієнт впливу тролів q в окремо взятому посту. Кількість тролів на пряму залежить від кількості конфліктних коментарів N_{cc} , визначаємо аналізуючи пост по його коментарям N_{ccz} . Визначення агента троль залежить від кількості конфліктних або провокаційних коментарів:

$$k_{troll} = N_{cc} / N_{ccz}$$

Коефіцієнт впливу q тролів на користувачів в соціальній мережі визначаємо як (розповсюджується на 2 фазу):

$$q = Af * Ad * \delta$$

Ad залежить від часу та коефіцієнта впливу k :

$$Ad = t * k$$

На 3 фазі суттєву роль відіграє рівень залученості Loi тролів, тому коефіцієнт впливу q змінюється на коефіцієнт впливу m :

$$m = q * Loi$$

4. Звичайні користувачі (Regular Users).

Роль у соціальній мережі:

- Основна маса учасників, що пасивно або активно споживають контент.
- Можуть поширювати контент, але не впливають на широкий загал.

Ключові характеристики:

1. Кількість лайків та репостів N_{lr} – середня кількість взаємодій з контентом на день.

2. Рівень активності Al – кількість постів та коментарів на день.

3. Коефіцієнт впливу η – частка оригінального контенту у профілі.

4. Частота заходів в соцмережу Fs – середня кількість входів у день.

5. Індивідуальний коефіцієнт мотивації λ – рівень зацікавленості користувача в участі в інформаційних кампаніях.

За ключовими характеристиками визначимо коефіцієнт впливу звичайних користувачів в окремо взятому посту. Кількість звичайних користувачів буде

залишковою за виключенням кількості ботів та тролів, визначаємо аналізуючи пост по його коментарям:

$$k_{person} = lom_count - (bot_count + trol_count)$$

Коефіцієнт впливу b звичайних користувачів на користувачів в соціальній мережі враховуємо однаковим для всіх фаз та визначаємо як:

$$b = Nlr * Al * \eta * Fs * \lambda$$

Використовуючи наступні данні (рис.1) побудуємо графік коефіцієнтів впливу розповсюдження інформації в соціальній мережі “X (Twitter)”, для кожного типу агента. Програмний код написаний на мові програмування python наведено в додатку 1.

Для розрахунку коефіцієнтів впливу було використано наступні значення:

- 1) Коефіцієнта впливу лідера думок: $F = 7000$, $Ps = 3$, $Pl = 5$, $En = 120$, $\alpha = 0.1$, $K = 10$;
- 2) Коефіцієнта впливу ботів: $Pf = 5$, $Nrd = 20$, $\beta = 0.00055$, $Br = 0.1$;
- 3) Коефіцієнта впливу тролів: $Af = 250$, $t = 600$, $k = 0.001$, $\delta = 0.000114$, $Loi = 0.059$;
- 4) Коефіцієнта впливу звичайних користувачів: $Nlr = 17$, $Al = 6$, $\eta = 0.0013$, $Fs = 5$, $\lambda = 0.0009$.

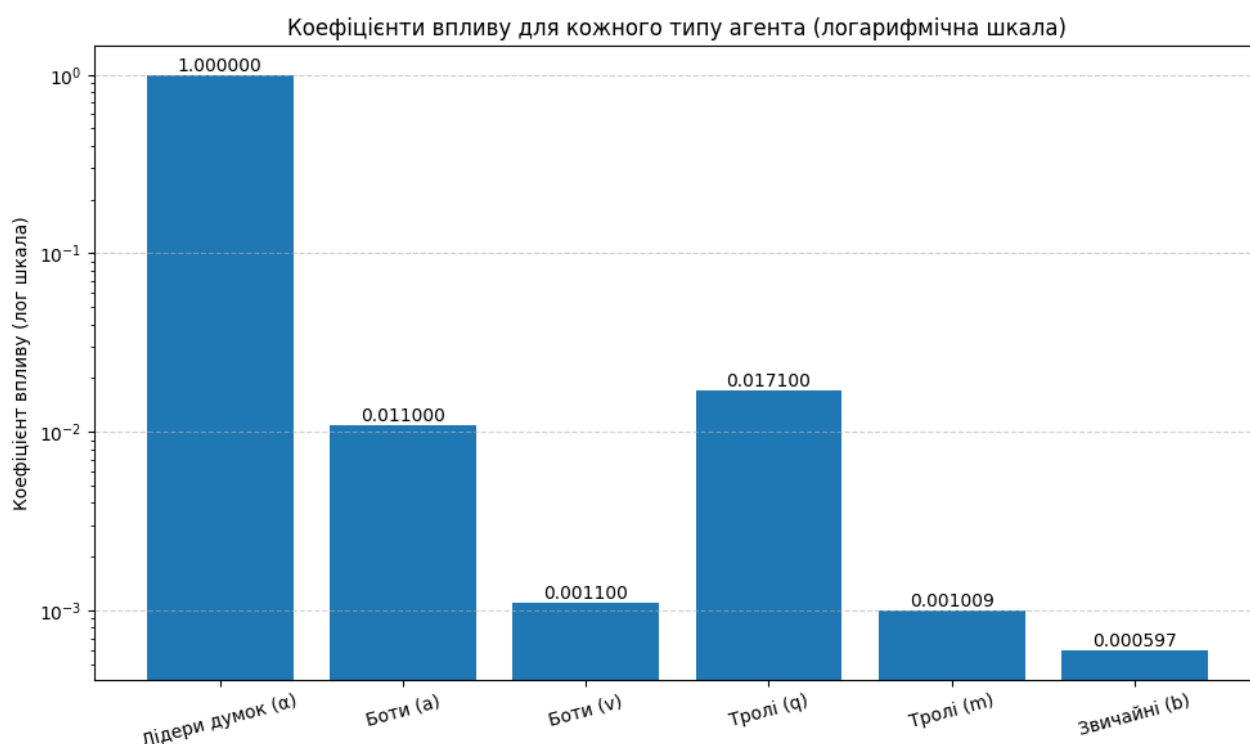


Рис.1 – Графік коефіцієнтів впливу кожного типу агента на розповсюдження інформації в соціальній мережі X (Twitter)

Вплив на користувачів здійснюється поетапно та залежить від першоджерела, а саме:

Лідери думок (негайне максимальне охоплення);

Тролі (фаза 2) – високий емоційний вплив;
 Боти (фаза 2) – автоматизована присутність;
 Боти/тролі (фаза 3) – вплив падає через блокування / зниження активності;
 Звичайні користувачі – мінімальний, довгостроковий вплив.

Проведемо моделювання змінюючи основні характеристики кожного типу агента, для знаходження більш значних та вагомих характеристик (рис. 2-5). Програмний код написаний на мові програмування python наведено в додатку 2.

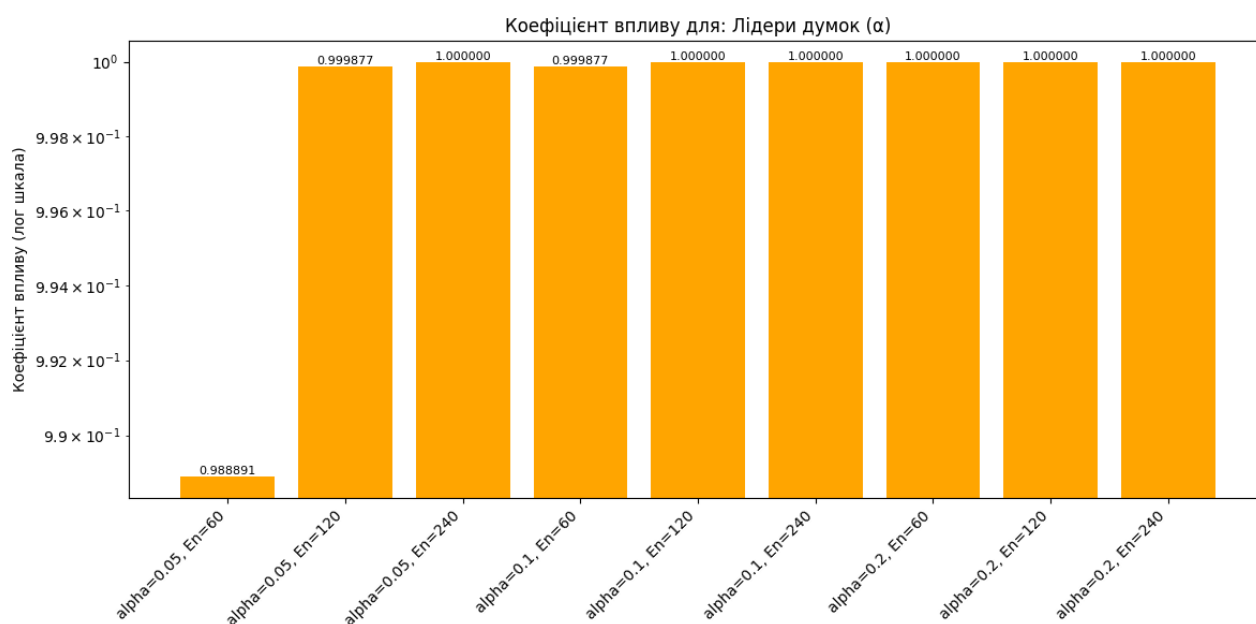


Рис.2 – Моделювання розрахунку коефіцієнту впливу лідера думок при різних вхідних параметрах

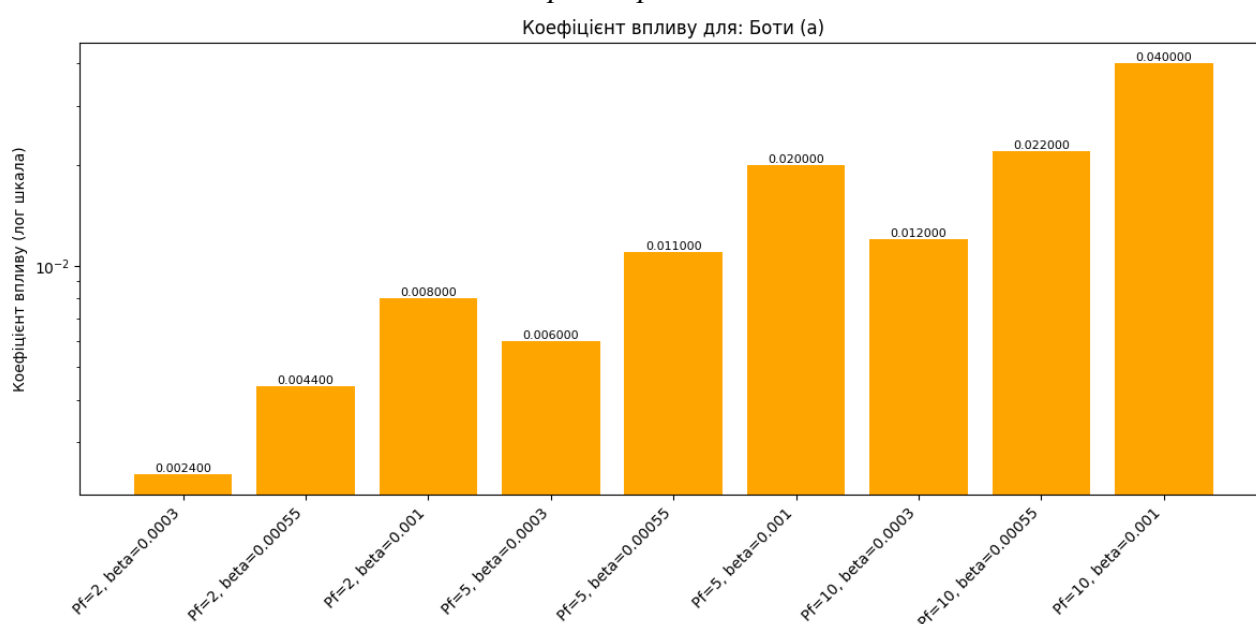


Рис.3 – Моделювання розрахунку коефіцієнту впливу ботів при різних вхідних параметрах

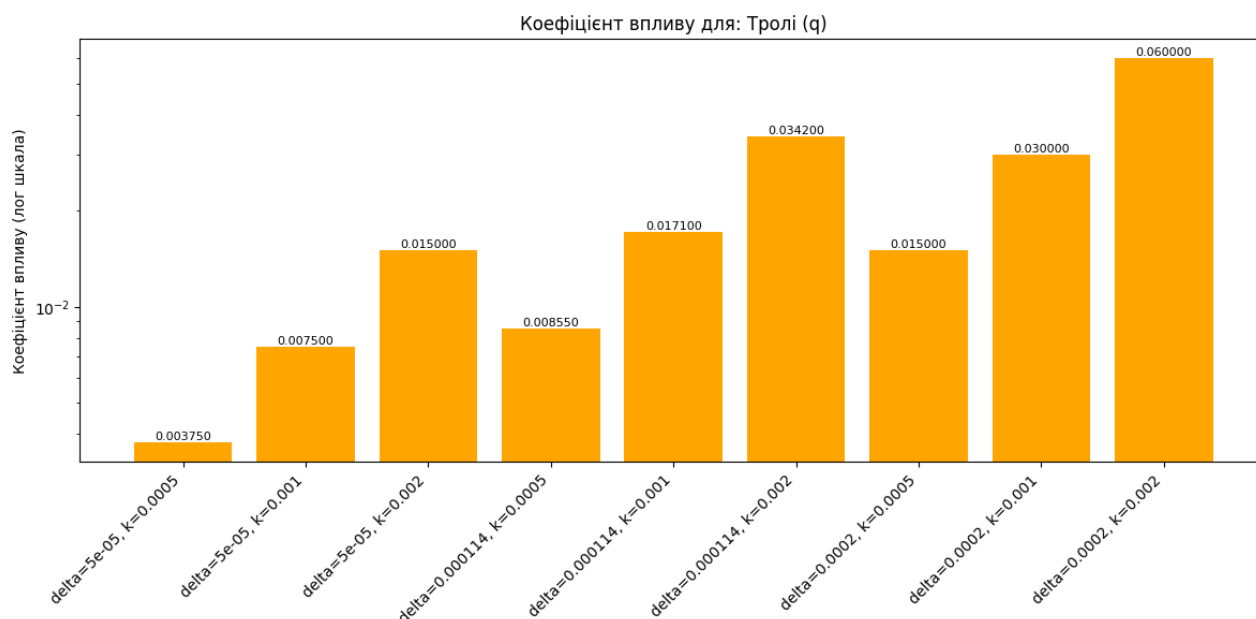


Рис.4 – Моделювання розрахунку коефіцієнту впливу тролів при різних вхідних параметрах

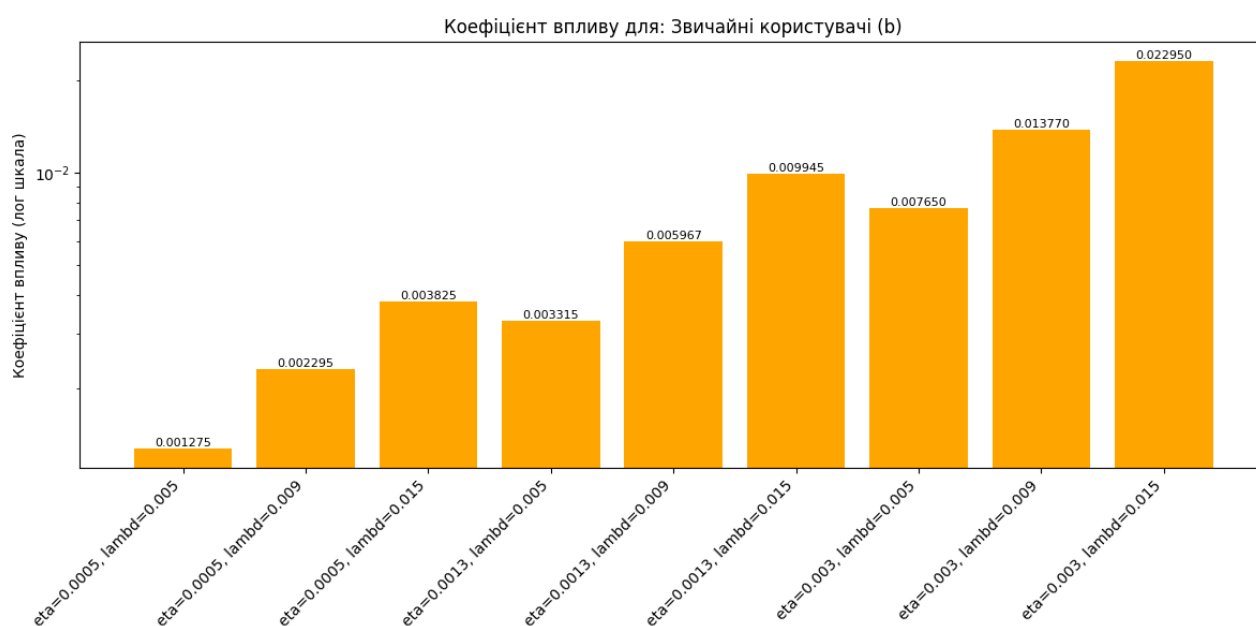


Рис.5 – Моделювання розрахунку коефіцієнту впливу звичайних користувачів при різних вхідних параметрах

На основі моделювання і обчислених коефіцієнтів, можна зробити такі висновки:

1) Лідери думок l , ключові параметри: α (довіра), En (взаємодія). Висновок: навіть незначне збільшення α або En швидко доводить вплив до 1, тобто лідери думок досягають максимальної ефективності при середньому рівні активності та довіри.



Чутливість висока лише при дуже низьких значеннях α/En , після чого графік "насичується". Найбільш вагомий параметр – En (взаємодія), але ефект значно підсилюється при довірі ($\alpha \geq 0.1$).

2) Боти a , ключові параметри: Pf (частота публікацій), β (автоматизація). Висновок: вплив ботів зростає лінійно при збільшенні Pf , а також помітно змінюється при навіть невеликих змінах автоматизації. Наприклад, перехід β з 0.0003 до 0.001 при фіксованому Pf зростання впливу ботів майже втричі. Найбільш вагомий параметр – β (автоматизація), оскільки він масштабує весь ефект.

3) Тролі q , ключові параметри: δ (агресія), k (коефіцієнт активності). Висновок: вплив тролів експоненційно зростає при збільшенні обох параметрів. Найбільше значення q спостерігається при максимальному k і δ . Значення зростають від 0.0037 до 0.06 – майже в 16 разів у межах сценарію. Обидва параметри мають однакову чутливість, особливо у комбінації.

4) Звичайні користувачі b , ключові параметри: η (оригінальний контент), λ (мотивація). Висновок: лінійне зростання впливу, але він залишається низьким у порівнянні з іншими агентами. Найвищі значення досягаються лише при максимальному η та λ . Від 0.001 до 0.023 – ефект є, але м'який. Найбільш вагомий параметр – λ (мотивація), що визначає готовність поширювати контент.

Таблиця 1

Поведінкові характеристики та їх вплив на ефективність агентів

№	Агент	Найважливіший параметр	Тип зростання	Коментар
1	Лідери думок	En, α	Швидке насичення	Достатньо середніх значень для максимуму
2	Боти	β	Лінійне/експоненційне	Залежить від ступеня автоматизації
3	Тролі	δ, k	Експоненційне	Вибуховий ріст при провокаціях
4	Звичайні користувачі	λ	Лінійне	Найменший вплив, але стабільний

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Проведене дослідження дозволило здійснити формалізовану оцінку впливу різних типів агентів на поширення інформації в соціальній мережі X (Twitter). Було розроблено та застосовано математичні моделі для обчислення коефіцієнтів впливу з урахуванням поведінкових характеристик чотирьох категорій агентів: лідерів думок, ботів, тролів і звичайних користувачів.

На основі проведених симуляцій та графіків впливу (рис.1–5) встановлено наступне.

- Лідери думок демонструють найвищий вплив уже на початку, що обумовлено довірою до них α та рівнем взаємодії En . Вплив насичується при середніх значеннях і швидко досягає максимуму.
- Боти є ефективними у 2 фазі поширення. Їх вплив зростає лінійно або експоненційно при підвищенні автоматизації β та частоти публікацій Pf . У 3 фазі ефективність падає через блокування.



- Тролі мають вибуховий емоційний вплив, особливо при високих значеннях агресії δ та активності k . Вони здатні швидко ескалювати ситуацію в інформаційному середовищі.
- Звичайні користувачі мають мінімальний, але стабільний вплив, який нарощується в третій фазі. Важливими факторами є мотивація λ та присутність оригінального контенту η .

Проведений аналіз параметрів чутливості дозволив виокремити найбільш вагомі характеристики для кожного типу агентів та визначити їх вплив на ефективність інформаційної операції. Це дає змогу:

- здійснювати більш точне прогнозування динаміки поширення повідомлень;
- конструювати стратегії впливу з акцентом на потрібний тип агентів;
- розробляти сценарії інформаційних операцій з урахуванням фазової структури залучення аудиторії.

Найбільшу ефективність в інформаційній кампанії демонструють агенти, що активні в перших двох фазах – зокрема, лідери думок, тролі та боти. Їх вплив у комбінації є критичним для успішної реалізації інформаційної операції. За відсутності ботів і тролів ефективність значно знижується. Звичайні користувачі формують фонове, пасивне поширення, і можуть лише підсилювати ефект лідерів у пізні фази.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday*, 22(8). <https://doi.org/10.5210/fm.v22i8.8005>
2. Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. presidential election online discussion. *First Monday*, 21(11). <https://doi.org/10.5210/fm.v21i11.7090>
3. Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications*, 10, Article 7. <https://doi.org/10.1038/s41467-018-07761-2>
4. Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
5. Aral, S., & Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, 337(6092), 337–341. <https://doi.org/10.1126/science.1215842>
6. González-Bailón, S., Borge-Holthoefer, J., & Moreno, Y. (2013). Broadcasters and hidden influentials in online protest diffusion. *American Behavioral Scientist*, 57(7), 943–965. <https://doi.org/10.1177/0002764213479371>
7. Howard, P. N., & Kollanyi, B. (2016). Bots, #StrongerIn, and #Brexit: Computational propaganda during the UK-EU referendum. *COMPROM Research Note 2016.1*. <https://doi.org/10.2139/ssrn.2798311>
8. Starbird, K., Arif, A., & Wilson, T. (2019). Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), Article 127. <https://doi.org/10.1145/3359229>
9. Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016). The structural virality of online diffusion. *Management Science*, 62(1), 180–196. <https://doi.org/10.1287/mnsc.2015.2158> (передбачимо DOI, але пошук не виявив прямий DOI; вказано журнал і дату)
10. Berger, J. (2013). *Contagious: Why things catch on*. New York, NY: Simon & Schuster. ISBN: 978-1451686579
11. О.Васильєва, Технічні науки та технології: науковий журнал / Національний університет «Чернігівська політехніка». – Чернігів: НУ «Чернігівська політехніка», 2023. – № 2(32). – 418 с., 193-200
12. Daniel J. Brass A SOCIAL NETWORK PERSPECTIVE ON HUMAN RESOURCES MANAGEMENT, *Research in Personnel and Human Resources Management*, 39-79. DOI:10.1093/oso/9780195159509.003.0019

**Vasylieva Olha**

graduate student

Chernihiv Polytechnic National University, Chernihiv, Ukraine

ORCID ID: 0000-0002-8565-0525

olga.vasiljeva37@gmail.com

TYOLOGY OF AGENTS AND THEIR INFLUENCE ON THE SPREAD OF DISINFORMATION IN SOCIAL NETWORKS

Abstract. Recently, there have been numerous studies analyzing the role of social networks in various types of communication, particularly in the spread of informational influences such as disinformation, fake news, and targeted information campaigns. Such phenomena poses a threat to the information security of any state. Currently, there is active use of the X network (Twitter) to influence both European and Ukrainian social media users. Given the importance of this issue, this study aimed to analyze the algorithm for content dissemination, taking into account different types of users (so-called “agents”). The study analyzed four types of agents and identified their key characteristics, which are likely to determine the level of influence during an information operation, which, in turn, is the main criterion for building a simulation model of content distribution on social networks.

Keywords information security; simulation agent modeling; social networks; disinformation; information operations.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday*, 22(8). <https://doi.org/10.5210/fm.v22i8.8005>
2. Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. presidential election online discussion. *First Monday*, 21(11). <https://doi.org/10.5210/fm.v21i11.7090>
3. Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications*, 10, Article 7. <https://doi.org/10.1038/s41467-018-07761-2>
4. Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
5. Aral, S., & Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, 337(6092), 337–341. <https://doi.org/10.1126/science.1215842>
6. González-Bailón, S., Borge-Holthoefer, J., & Moreno, Y. (2013). Broadcasters and hidden influentials in online protest diffusion. *American Behavioral Scientist*, 57(7), 943–965. <https://doi.org/10.1177/0002764213479371>
7. Howard, P. N., & Kollanyi, B. (2016). Bots, #StrongerIn, and #Brexit: Computational propaganda during the UK-EU referendum. *COMPROM Research Note 2016.1*. <https://doi.org/10.2139/ssrn.2798311>
8. Starbird, K., Arif, A., & Wilson, T. (2019). Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), Article 127. <https://doi.org/10.1145/3359229>
9. Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016). The structural virality of online diffusion. *Management Science*, 62(1), 180–196. <https://doi.org/10.1287/mnsc.2015.2158> (DOI expected, but search did not find a direct DOI; journal and date are provided)
10. Berger, J. (2013). *Contagious: Why things catch on*. New York, NY: Simon & Schuster. ISBN: 978-1451686579
11. Vasylieva O, *Technical Sciences and Technologies: Scientific Journal / National University “Chernihiv Polytechnic.”* – Chernihiv: NU “Chernihiv Polytechnic,” 2023. – No. 2(32). – 418 p., 193-200
12. Daniel J. Brass A SOCIAL NETWORK PERSPECTIVE ON HUMAN RESOURCES MANAGEMENT, *Research in Personnel and Human Resources Management*, 39-79. DOI:10.1093/oso/9780195159509.003.0019



This work is licensed under Creative Commons Attribution-noncommercial-sharealike 4.0 International License.