

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Рукопис

ВАВРИК ПЕТРО РУСЛАНОВИЧ

УДК 519.722:519.816

ДИСЕРТАЦІЯ
МОДЕЛІ ТА МЕТОДИ АНАЛІЗУ ЕВОЛЮЦІЙНИХ ПРОЦЕСІВ
РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ У СОЦІАЛЬНИХ МЕРЕЖАХ

124 – Системний аналіз

Подається на здобуття наукового ступеню доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

(підпис, ініціали та прізвище здобувача)

Науковий керівник: **Івохін Євген Вікторович**, доктор фізико-математичних наук, професор.

Київ – 2023

АНОТАЦІЯ

Ваврик П.Р. Моделі та методи аналізу еволюційних процесів розповсюдження інформації у соціальних мережах. – Рукопис.

Дисертація на здобуття наукового ступеня доктора філософії із спеціальності 124 – Системний аналіз. Київський національний університет імені Тараса Шевченка МОН України, Київ, 2023.

Підготовка дисертаційної роботи здійснювалась на кафедрі системного аналізу і теорії прийняття рішень факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка.

Дисертаційна робота присвячена дослідженню та формалізації еволюційних процесів розповсюдження інформації в соціальних мережах, на основі побудови ефективних моделей відображення складних мереж та моделюванні взаємодій між сутностями в складних мережах.

З швидким розвитком інформаційних технологій та загальним збільшенням розміру технологічних процесів, все частіше виникає необхідність в проведенні систематичних досліджень таких складних систем. Під час проведення досліджень, науковці найчастіше стикаються з двома принциповими проблемами: збором актуальних даних про системи та побудовою математичних моделей для опису процесів, що відбуваються в них. У випадку простих систем, математичний апарат вже сформований, проте як тільки система починає проявляти динамічні характеристики, існуючий математичний апарат не здатний забезпечувати дослідників актуальною інформацією про стан системи.

Найчастіше для формалізації складних систем використовуються міждисциплінарні підходи, що базуються на окремих методах фізики, біології та соціології: теорія складних мереж, теорія графів, теорія хаосу, клітинні автомати, теорія інформації, теорія ігор, моделі дифузії, фрактальний аналіз, аналіз часових рядів, диференціальні рівняння, диференціальні рівняння в частинних похідних, різниці рівняння. Проте, зі зростанням складності систем та потребі в їх

комплексному аналізі, все частіше застосовуються методи агентного моделювання. Дослідженню складних систем присвячено роботи *Albert-László Barabási*, *Stephen Wolfram*'а, *Stuart Kauffman*'а, *Robert Axtell*'а, *Joshua M. Epstein*'а та інших. Варто зазначити, складні системи зазвичай сильно відрізняються та залежать від області моделювання.

Соціальні мережі є яскравим прикладом складних систем, в яких постійно взаємодіють користувачі, породжуючи складні процеси поведінки, пов'язані з створенням публікацій, розповсюдженням інформації та утворенням розповсюдження пливу.

Зв'язки між сутностями в соціальних мережах утворюють складні мережі, дослідженнями яких активно займався *Albert-László Barabási*, вдосконалюючи теорію випадкових графів та алгоритмів переважного приєднання.

Саме тому дисертаційна робота зосереджена на всебічному дослідженні соціальних мереж та мереж економічних зв'язків як складних систем. В основі дослідження складних систем лежить побудова динамічних моделей, що формалізують процеси, характеристики, та поведінку окремих складових складних систем та системи в цілому, що дозволяє моделювати її з врахування зовнішніх чинників. Важливим аспектом побудови моделей є збереження кількісних та якісних характеристик вхідних складних систем.

Найбільш ефективними підходами до математичного моделювання процесів розповсюдження інформації в соціальних мережах, а також економічних зв'язків між людьми, є агентне моделювання складних мереж з використанням синтетичних популяцій. Впродовж останніх років, в умовах пандемії *COVID-19* та активного використання соціальних мереж для політичної реклами та розповсюдження пропаганди, з'явилося багато досліджень, які використовують схожу методологію, проте більшість з них зіткнулися з різноманітними особливостями їх застосування до вирішення реальних проблем.

Незалежно від області моделювання складних систем, дослідники часто нехтують важливістю методології. Серед частих випадків застосування хибних підходів: перенавчання моделей, надмірне спрощення моделей, ігнорування

динамічної природи складних систем, використання лише історичних даних, ігнорування механізмів зворотного зв'язку, відсутність процесу калібрування та верифікації результатів, ігнорування неоднорідності складних систем, застосування статистичних моделей при описі динамічних систем, ігнорування зовнішніх чинників, ігнорування невизначеності процесів в складних системах, звуження масштабу складних мереж. Кожен з цих підходів гарантовано унеможливить використання таких моделей для систематичного аналізу складних систем та застосування їх під час підтримки прийняття рішень.

Саме ефективність застосування агентного моделювання для дослідження динамічних процесів взаємодії в соціальних мережах та мережах економічних зв'язків зробили його основною методологією роботи.

Було проведено огляд актуальних досліджень розповсюдження впливу та інформації в соціальних мережах. Детально проаналізовано основні особливості моделювання процесу розповсюдження інформації в контексті наборів даних, на яких проводилося дослідження обраної методології, та розв'язання проблеми вибору оптимальних цільових аудиторій для оптимізації впливу. Виходячи з результатів розглянутих досліджень, було встановлено принципові недоліки підходів до дослідження розповсюдження інформації, що характеризуються системною відсутністю реалістичних моделей соціальних мереж на основі публічних даних та застосуванню методології, що не здатна описувати динамічну структуру процесів та взаємодій в соціальних мережах та соціальних мереж загалом.

Було розглянуто актуальні дослідження соціальних зв'язків в економічних мережах, як важливого аспекту соціальної взаємодії. Детально проаналізовано основні особливості моделювання процесів в економічних мережах. Виходячи з результатів розглянутих досліджень, було встановлено схожі недоліки моделювання, в основі яких – відсутність моделей реальних економічних мереж, а також методології, що не враховує поведінковий аспект.

Для моделювання розповсюдження інформації в соціальних мережах та моделювання процесів в економічних мережах, було запропоновано використати методологію аналізу складних мереж з використанням агентного моделювання та синтетичних популяцій.

Було формалізовано процеси розповсюдження інформації в соціумі, які згодом масштабовані як основна складова сучасних соціальних мереж. Було розглянуто основні механізми та моделі розповсюдження інформації в соціальних мережах.

Було досліджено динамічні фактори впливу на розповсюдження інформації в соціальних мережах та запропоновано застосування моделювання соціальних мереж, яке базується на дослідженні поведінки користувачів. Було проведено дослідження особливостей різних соціальних мереж, формалізовано їх основні сутності та способи взаємодії.

Для ефективного моделювання процесу розповсюдження інформації, було запропоновано модель мережі перетину аудиторій на основі публічних даних. В умовах обмеження соціальними мережами доступу до даних, було показано як модель мережі подібності публікацій та мережі взаємодій дозволяють оцінювати перетин аудиторій. З використанням мережі перетину аудиторій, було досліджено динаміку еволюції спільнот соціальної мережі *Telegram* та процес розповсюдження впливу.

Для розв'язання задачі оптимального поширення впливу було запропоновано модель з використанням агентів, а мережу перетину аудиторій використано для вибору оптимальних цільових аудиторій для агентів. Для ефективного застосування агентів було проведено моделювання поширення інформації для випадків з різною кількістю агентів.

Було проведено збір публічних даних як основу для побудови синтетичної популяції населення штату Гаваї. Для побудови синтетичної популяції було реалізовано алгоритм імітації відпалу з перевіркою результатів на основних демографічних характеристиках населення.

Для моделювання процесів в економічних мережах було розроблено агентну модель взаємодії населення, в якій кожному агенту було призначено набір правил поведінки. Для перевірки коректності агентної моделі було реалізовано процес калібрування з використанням еволюційних стратегій. Розроблено програмний комплекс застосування агентної моделі економічних зв'язків для побудови прогнозів стану економічної мережі в випадку наявності зовнішніх факторів.

Ключові слова: складні системи, складні мережі, агентне моделювання, синтетичні популяції, динамічні моделі, соціальні мережі, моделі соціальних мереж, розповсюдження інформації, розповсюдження впливу, оптимальне розповсюдження впливу, лідери думок та аудиторії в соціальних мережах, економічні мережі, моделі динаміки економік.

ANNOTATION

Vavryk P.R. Models and methods for analyzing evolutionary processes of information transmission in social networks. – Manuscript.

The thesis for a PhD degree, specialty 124 – System analysis. Taras Shevchenko National University of Kyiv MES of Ukraine, Kyiv, 2023.

The preparation of the dissertation was carried out at Department of System Analysis and Decision-Making Theory of Computer Science and Cybernetics of Taras Shevchenko National University of Kyiv.

The PhD thesis is dedicated to the research and formalization of evolutionary processes of information propagation in social networks, based on the construction of effective models for representing complex networks and modeling interactions between entities in complex networks.

With the rapid development of information technologies and the general increase in the size of technological processes, there is a growing need to conduct systematic research of such complex systems. During research, scientists most often face two

fundamental problems: collecting relevant data about systems and building mathematical models to describe the processes occurring within them. In the case of simple systems, the mathematical apparatus has already been formed, but as soon as the system begins to exhibit dynamic characteristics, the existing mathematical apparatus is not able to provide researchers with up-to-date information about the state of the system.

Frequently, interdisciplinary approaches based on separate methods of physics, biology, and sociology are used to formalize complex systems: the theory of complex networks, graph theory, chaos theory, cellular automata, information theory, game theory, diffusion models, fractal analysis, time series analysis, differential equations, partial differential equations, difference equations. However, with the growing complexity of systems and the need for their complex analysis, agent-based modeling methods are increasingly applied. The work of Albert-László Barabási, Stephen Wolfram, Stuart Kauffman, Robert Axtell, Joshua M. Epstein and others is devoted to the study of complex systems. It is worth noting that complex systems are usually differ significantly and depend on the modeling domain.

Social networks are a vivid example of complex systems in which users are constantly interacting, generating emerging complex behaviors related to the creation of publications, the distribution of information, and the formation of the influence propagation.

Relationships between entities in social networks form complex networks, which were actively researched by Albert-László Barabási, advancing the theory of random graphs and preferential attachment algorithms.

Therefore, the PhD thesis is focused on a comprehensive study of social networks and networks of economic activities as complex systems. At the heart of the study of complex systems is the construction of dynamic models that formalize the processes, characteristics, and behavior of individual components of complex systems and the system as a whole, allowing its simulation considering external factors. An important aspect of building models is the preservation of quantitative and qualitative characteristics of input complex systems.

The most effective approaches for mathematical modeling of processes, information propagation in social networks, as well as economic relations between people are agent-based models of complex networks using synthetic populations. In recent years, in the context of the COVID-19 pandemic and the active use of social networks for political advertising many studies have appeared that use a similar methodology. However, most of them encountered various nuances in applying them to solve real-world problems.

Regardless of the domain of complex systems modeling, researchers often overlook the importance of methodology. Among the frequent cases of using incorrect approaches: overfitting of models, oversimplification of models, ignoring the dynamic nature of complex systems, relying solely on historical data, ignoring feedback mechanisms, lack of the calibration and verification processes, ignoring the heterogeneity of complex systems, using statistical models when describing dynamic systems, ignoring external factors, ignoring the uncertainty of processes in complex systems, and narrowing the scope of complex systems. Each of these approaches inevitably makes it impossible to use such models for systematic analysis of complex systems, and their application during decision-making support.

The very effectiveness of using agent-based modeling for the study of dynamic interaction processes in social networks and networks of economic activity that made it the primary methodology of this thesis.

A review of current research on the spread of influence and information in social networks was conducted. The main features of modeling the information propagation process in the context of data sets on which the research was conducted, the chosen methodology and solving the problem of choosing the optimal target audiences for optimizing the impact are analyzed in detail. Based on the results of the reviewed studies, the fundamental shortcomings of approaches to the study of information propagation were established, which are characterized by the systematic lack of realistic models of social networks based on public data and the application of a methodology that is unable to describe the dynamic nature of processes and interactions in social media networks and social networks in general.

Current research on social relations in economic networks was considered as an important aspect of social interaction. The main features of modeling processes in economic networks are analyzed in detail. Based on the results of the reviewed studies, similar shortcomings of modeling were established, based on the lack of models of real economic networks, as well as a methodology that does not consider the behavioral aspect.

To model the spread of information in social networks and modeling of processes in economic networks, it was proposed to use the methodology of analyzing complex networks using agent-based modeling and synthetic populations.

The processes of information propagation in society were formalized, which were later scaled as the main component of modern social networks. The main mechanisms and models of information propagation in social networks were considered.

The dynamic factors influencing the propagation of information in social networks were investigated, and the application of modeling of social networks, which is based on the researched user behavior, was proposed. A study of the features of various social networks was conducted, their main essences and methods of interaction were formalized.

To effectively model the information propagation process, a model of the audience overlap network based on public data was proposed. In conditions of restriction of access to data by social networks, it was shown how the model of the network of similarity of publications and the network of interactions allow to evaluate the audience overlap. Using the audience overlap network, the dynamics of the evolution of Telegram communities and the process of spreading influence were investigated.

To solve the problem of optimal spread of influence, a model using agents was proposed, and the audience overlap network was used to guide the selection of target audiences for agents. For the effective application of agents, simulations of information propagation were carried out for cases with different numbers of agents.

Public data was collected as a foundation for constructing a synthetic population of the state of Hawaii. To create the synthetic population, a simulated annealing

algorithm was implemented with verification of the results based on the primary demographic characteristics of the population.

For modeling processes in economic networks, an agent-based model of population interaction was developed, where each agent was assigned a set of behavioral rules. To verify the correctness of the agent-based model, a rigorous calibration process was implemented using evolutionary strategies.

A set of tools were developed to utilize the agent-based model of economic activity, facilitating forecasting of the network's state when accounting for external factors.

Keywords: complex systems, complex networks, agent-based modeling, synthetic populations, dynamic models, social networks, models of social networks, information propagation, influence propagation, optimal influence propagation, influencers and audiences in social networks, networks of economic activity, dynamics economic models.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ПО ТЕМІ ДИСЕРТАЦІЇ

1. Ivohin E., Vavryk P. Building a graph of intersection of social network audiences on alternative data // XXXV International Conference PROBLEMS OF DECISION MAKING UNDER UNCERTAINTIES (PDMU-2020). May 11 – 15, 2020, Baku-Sheki, Republic of Azerbaijan, p.46.

ТЕЗИ

2. Vavryk P. Applying calibrations for agent-based simulations of economies in distress// Proceedings of XIX International Scientific-Practical Conference "Shevchenkivska Vesna – 2021", April 2021, Kyiv, Ukraine, p.29-30.

ТЕЗИ

3. Vavryk P. To ban or not to ban: fighting a trench war in comment sections // XXXVI International Conference PROBLEMS OF DECISION MAKING UNDER UNCERTAINTIES (PDMU-2021). May 11 – 14, 2021, Skhidnytsia, Ukraine, p.112.

ТЕЗИ

4. Ivohin, E., Vavryk, P., Rudoman, N. About audience overlaps in the social media. // Bulletin of Taras Shevchenko National University of Kyiv. Series: Physics and Mathematics. 2021, P. 69-73. DOI: 10.17721/1812-5409.2021/1.8.

СТАТТЯ у журналі категорії «Б»

5. E. Ivohin, V. Gavrilenko, і P. Vavryk, Про один метод оцінки перетину аудиторій у соціальних мережах, Системи управління, навігації та зв'язку. Збірник наукових праць, т. 4, вип. 66, с. 48-51, Груд 2021. DOI: <https://doi.org/10.26906/SUNZ.2021.4.048>

СТАТТЯ у журналі категорії «Б»

6. Vavryk P. et al. Forecasting Unemployment Dynamics in Response to COVID-19: A Case Study. Issues of support and analysis of decisions / ed. by Jakubczyk M, Kamiński B., Kuszewski T. Warsaw, 2021. P. 254-266 ISBN: 978-83-8030-443-7

(<https://24naukowa.com.pl/zagadnienia-wspomagania-i-analizy-decyzji,id17138.html>)

РОЗДІЛ ЗАКОРДОННОЇ (Польща) МОНОГРАФІЇ

7. Vavryk P. A method for estimating audience overlaps in the social media. // On-line conference: NETWORKS 2021: A JOINT SUNBELT AND NETSCI CONFERENCE, July 5–10, 2021.

Conference Program, p.444. за посиланням:
<https://drive.google.com/file/d/1gNxVAMD56kfvzehowjw6b0irKUOIM51o/view>

Участь у закордонній (он-лайн, США) конференції, тези не друкувалися.
ДОПОВІДЬ 07.07.2021.

8. Ваврик П.Р. Про один підхід до формалізації процесу поширення інформації на основі агрегації, що обмежена дифузією.// Вісник Київського національного університету імені Тараса Шевченка. Серія фізико-математичні науки. 2022, №1, С. 61-66. DOI: 10.17721/1812-5409.2022/1.8

СТАТТЯ у журналі категорії «Б»

9. Івохін Є.В., Аджубей Л.Т., Ваврик П.Р. Про один метод розв'язання задачі розподілу потужності каналів передачі даних з урахуванням нечітких обсягів споживання. // Інформаційні технології в культурі, мистецтві, освіті, науці, економіці та бізнесі. VII Міжнародна науково-практична конференція. 20-21 квітня 2022 р., Київ, Україна, ст. 45-47. УДК 519.87:004.02

ТЕЗИ

10. Vavryk P. Mapping growth of the Russian domestic propaganda apparatus on Telegram. // The 3rd International Conference Challenges to National Defence in Contemporary Geopolitical Situation (CNDCGS'2022), 15 – 16 September 2022, Vilnius, Lithuania, p.51-52.

ТЕЗИ, Участь у закордонній (он-лайн, Литва) конференції, ДОПОВІДЬ 15.09.2022.

11. Ivohin E., Adzhubey L., Vavryk P., Makhno M. On some methods for solving the problem of power distribution of data transmission channels taking into account

fuzzy constraints on consumption volumes. 2022. System research and information technologies. 88-99. DOI: 10.20535/SRIT.2308-8893.2022.4.08.

Стаття (Scopus)

- 12.Vavryk P. Mapping Growth of the Russian Domestic Propaganda Apparatus on Telegram. Challenges to national defence in contemporary geopolitical situation. 2022. 227-231. DOI: 10.47459/cndcgs.2022.29.

Стаття (Scopus, Conference paper)

- 13.Ivohin E., Adzhubey L., Vavryk P., Rets V. On the influence of fuzzy perception of the flow of time on the solution of optimization problems of planning // XXXVII International Conference PROBLEMS OF DECISION MAKING UNDER UNCERTAINTIES (PDMU-2022), November 23-25, 2022, p.56.

ТЕЗИ

- 14.Е.В. Івохін, д.ф.-м.н., Л.Т. Аджубей, к.ф.-м.н., П.Р.Ваврик Про один метод розв'язання задачі розподілу потужностей каналів даних з урахуванням нечітких обмежень обсягів споживання // Матеріали XVI міжнародної науково-технічної конференції «АВІА-2023», 18-20 квітня 2023 року. –К.: НАУ, с. 15.38 – 15.40.

ТЕЗИ

ЗМІСТ

АНОТАЦІЯ	2
ANNOTATION.....	6
СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ПО ТЕМІ ДИСЕРТАЦІЇ.....	11
ЗМІСТ	14
ВСТУП.....	17
РОЗДІЛ 1: ОГЛЯД СУЧАСНОГО СТАНУ ДОСЛІДЖЕНЬ СОЦІАЛЬНИХ ТА ЕКОНОМІЧНИХ МЕРЕЖ.....	24
1.1 Основні аспекти дослідження складних мереж	24
1.2 Моделювання та оптимізація розповсюдження впливу в соціальних мережах	28
1.2.1 Критерії пошуку актуальних досліджень	28
1.2.2 Порівняльна характеристика актуальних досліджень	30
1.2.3 Ключові характеристики актуальних досліджень	35
1.3 Моделювання наслідків пандемій на економічні мережі	36
1.3.1 Критерії пошуку актуальних досліджень	36
1.3.2 Порівняльна характеристика актуальних досліджень	38
1.3.3 Ключові характеристики актуальних досліджень	42
1.4 Висновки по першому розділі	44
РОЗДІЛ 2: ФОРМАЛІЗАЦІЯ ПРОЦЕСУ РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ В СОЦІАЛЬНИХ МЕРЕЖАХ.....	45
2.1 Механізми розповсюдження інформації	46
2.2 Моделі розповсюдження інформації.....	47
2.3 Динамічні фактори впливу на розповсюдження інформації	49
2.4 Висновки по другому розділі.....	50

РОЗДІЛ 3: АНАЛІЗ РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ В СОЦІАЛЬНИХ МЕРЕЖАХ..... 51

3.1	Підходи до побудови моделей соціальних мереж.....	52
3.1.1	Набір даних для побудови моделей соціальних мереж	58
3.1.2	Алгоритми побудови моделей соціальних мереж	59
3.1.3	Калібрування моделі перетину аудиторій.....	64
3.1.4	Аналіз отриманих результатів	66
3.2	Дослідження перетину аудиторій в соціальній мережі Telegram	68
3.2.1	Набір даних для побудови мережі перетину аудиторій.....	70
3.2.2	Дослідження динаміки еволюції мережі	72
3.2.3	Побудова мережі перетину аудиторій та пересилань	75
3.2.4	Аналіз отриманих результатів	79
3.3	Застосування перетину аудиторій для розповсюдження впливу.....	80
3.3.1	Концепція впливу в контексті користувача	81
3.3.2	Набір даних для моделювання розповсюдження впливу	82
3.3.3	Моделювання розповсюдження впливу	87
3.3.4	Аналіз отриманих результатів	90
3.4	Висновки по третьому розділі	91

РОЗДІЛ 4: АНАЛІЗ СОЦІАЛЬНИХ МЕРЕЖ ЕКОНОМІЧНИХ ВІДНОСИН.....92

4.1	Побудова синтетичної популяції.....	94
4.1.1	Набір даних для побудови синтетичної популяції	95
4.1.2	Синтез господарств та осіб	98
4.1.3	Синтез підприємств та робочих місць	100
4.1.4	Синтез мереж постачання продуктів та послуг	103
4.1.5	Аналіз отриманих результатів	106
4.2	Побудова агентної моделі соціальних взаємодій	107
4.2.1	Формалізація поведінки агентів	109
4.2.2	Особливості агентів	115

4.3	Моделювання сценаріїв зовнішнього впливу	115
4.3.1	Формалізація сценаріїв зовнішнього впливу	115
4.3.2	Калібрування агентної моделі	116
4.3.3	Прогнозування стану мережі	120
4.3.4	Аналіз отриманих результатів	122
4.4	Висновки по четвертому розділі.....	123
	ВИСНОВКИ.....	124
	СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	127
	ДОДАТОК 1.....	134
	ДОДАТОК 2.....	136
	ДОДАТОК 3.....	137

ВСТУП

Актуальність теми. З швидким розвитком інформаційних технологій та загальним збільшенням розміру технологічних процесів, все частіше виникає необхідність в проведенні систематичних досліджень таких складних систем. Під час проведення досліджень, науковці найчастіше стикаються з двома принциповими проблемами: збором актуальних даних про системи та побудовою математичних моделей для опису процесів, що відбуваються в них. У випадку простих систем, математичний апарат вже сформований, проте як тільки система починає проявляти динамічні характеристики, існуючий математичний апарат не здатний забезпечувати дослідників актуальною інформацією про стан системи.

Для прикладу, зі збільшенням кількості населення різних країн, та швидкими темпами розвитку міст, виникає проблема з автомобільним трафіком. Люди, що проживають в спальних районах, за відсутності якісного громадського транспорту, використовують власні автомобілі для поїздок до місць працевлаштування. З першого погляду, така складна система є тривіальною. Вона може бути описана у вигляді графу, в якому ребра представляють дороги, з вагою пропускної здатності, а далі з використанням математичного апарату теорії графів, можна моделювати різноманітні сценарії часу реакції світлофорів або проектування нових доріг. Проте, незважаючи на таку простоту, дослідники часто виявляються неспроможними провести моделювання в реальних умовах.

Перш за все, це спричинено відсутністю даних про подорожі автомобілістів. Для цього необхідно мати системи збору інформації про кількість автомобілів в довільний момент часу на кожній з доріг. Навіть у випадку наявності таких даних, дослідники застосовують моделі, що не враховують поведінковий аспект автомобілістів. Ці моделі, не враховують на які саме робочі місця прямують автомобілісти, скільки часу вони проводять за роботою та які сторонні візити здійснюються впродовж дня. Також, такі моделі не враховують

непередбачувані події, наприклад, аварії, пошкодження дорожнього покриття, ремонт доріг тощо. Тому, не зважаючи на наявність математичного апарату, що здатен описати складність системи в комбінації з відсутністю даних, робить таку проблему досі актуальною для дослідників. Важливим аспектом такої системи також є зв'язок між дослідниками та особами, що приймають рішення. Дослідники часто зосереджуються на математичному апараті, що призводить до непорозумінь. Моделювання такої системи, повинне надавати особам, що приймають рішення чіткі настанови в області наявних для них рішень.

Схожа ситуація виникає і в багатьох інших проблемних областях. Особливо це проявляється при дослідженні соціальних мереж, які активно розвиваються в сучасних реаліях, стаючи основним джерелом інформації для людей по всьому світу, а також при дослідженні мереж економічних зв'язків між людьми. Дослідники застосовують нейронні мережі або дифузійні методи для моделювання розповсюдження інформації в соціальних мережах, проте дуже часто не в змозі побудувати відображення самої мережі, на основі використання лише публічних даних. Тому, які б результати такі дослідження не генерували, вони приречені бути виключно теоретичними. Схожа ситуація відбулася під час пандемії *COVID-19*, під час якої дослідники були не в змозі прогнозувати розповсюдження захворювання, не зважаючи на те, що така система принципово не відрізняється від процесу розповсюдження інформації в соціальних мережах.

Саме тому, дисертаційна робота зосереджена на всебічному дослідженні таких складних систем. В основі дослідження лежить побудова динамічних моделей що формалізують процеси, характеристики та поведінку окремих складових складних систем та системи в цілому, що дозволяє моделювати її з врахування зовнішніх чинників. Важливим аспектом побудови моделей, є збереження кількісних та якісних характеристик вхідних складних систем. Побудова моделі найчастіше супроводжується виділенням основних структурних елементів системи (користувачі соціальних мереж), зв'язків між ними (дружба, підписки), та множини можливих взаємодій (коментар, вподобання, тощо). В залежності від області моделювання, важливість

структурних елементів може змінюватися, проте множина можливих взаємодій, залишається ключовою.

В загальному, моделі складних систем можуть бути як динамічними так і статичними, проте сама суть складної системи передбачає наявність динамічної поведінки. Проте, під час проведення досліджень, на кожному кроці модель може бути приведена до статичного вигляду, дозволяючи проводити аналіз окремих функціональних співвідношень між її складовими.

Необхідною складовою моделей є можливість появи непередбачуваних факторів. Наприклад, моделі фінансових ринків, що базуються на аналізі часових рядів, часто діють на основі припущення про експоненційне зростання економік. Це потенційно дозволяє отримувати прибутки на коротких проміжках часу, проте не дозволяє передбачувати події, такі як Світова фінансово-економічна криза 2008 року, або ж стрімке падіння глобальної економіки під час пандемії *COVID-19*.

Важливим аспектом побудови моделей є параметризація її складових елементів. Для прикладу, при побудові моделі аеропорту, кожен з типів літаків, що може приземлитися, має обмеження на допустиму швидкість. Проте, яку саме швидкість літак вибере для посадки описується параметром з інтервалу $[0, 1]$ від допустимої швидкості. Один і той самий літак при повторному приземленні, за незмінної допустимої швидкості, обиратиме різне значення параметру. В свою чергу, це впливає на час, необхідний для резервації посадкової смуги. Також природні умови як зовнішні чинники, що діють на систему, можуть впливати на значення параметру.

В загальному, для дослідження складних систем використовуються диференціальні і різницеві рівняння, клітинні автомати, агентне моделювання або їх комбінація в залежності від проблемної області.

Оскільки моделі складних систем є зазвичай динамічними, необхідною умовою їх використання є проведення процесу калібрування. Процес калібрування зосереджується на багаторазовому аналізі поведінки моделі на різних наборах

вхідних даних, для пошуку оптимальних значень параметрів моделі, Додаток 2, Додаток 3. В свою чергу це гарантує відсутність шуму та помилок в моделі.

Завершальним етапом побудови моделей є побудова механізму прогнозування. Модель аналізується з використання різних зовнішніх впливів та чинників, фіксуючи проміжні та кінцевий стан. Ці стани в агрегованому вигляді застосовуються для побудови аналітичних відображень, що дозволяють провести обробку та аналіз кількісних показників моделі.

Отримані результати моделювання динаміки складних систем породжують необхідність створення систем підтримки прийняття рішень. Використання таких систем є принциповим для осіб, що приймають рішення, оскільки ефективність моделі в кінцевому випадку визначається ефективністю її застосування для вирішення реальних проблем. Якість моделі також визначається її здатністю до пристосування до нових умов, у яких вона використовується, здатністю до інтеграції актуальних наборів даних, а також можливість реагувати на зворотній зв'язок з особами, що приймають рішення.

Найбільш ефективними підходами до математичного моделювання процесів, розповсюдження інформації в соціальних мережах, а також економічних зв'язків між людьми є агентне моделювання складних мереж з використанням синтетичних популяцій. Впродовж останніх років, в умовах пандемії *COVID-19* та активного використання соціальних мереж для політичної реклами, розповсюдження та пропаганди, з'явилося багато досліджень, які використовують схожу методологію, проте більшість з них зіткнулися з різноманітними особливостями їх застосування до вирішення реальних проблем. Саме ефективність застосування агентного моделювання для дослідження процесів в соціальних мережах та мережах економічних зв'язків визначили вибір теми, мети і задач дисертаційної роботи.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота є складовою частиною наукових робіт, що ведуться на кафедрі системного аналізу і теорії прийняття рішень факультету комп'ютерних наук та кібернетики та в науково-дослідному секторі "Проблем системного

аналізу" Київського національного університету імені Тараса Шевченка. Дослідження виконувалися в рамках науково-дослідної теми №22БФ015-02 «Розробка нових математичних методів теорії оптимальних рішень» (державний номер реєстрації 0122U001844, термін виконання 2022-2024г.г., в рамках програми "Інформатизація суспільства") та українсько-литовського науково-дослідного проекту «Моделювання ролі людського потенціалу для забезпечення оборони країни під час новітніх загроз» (у 2022 р. № держреєстрації 0122U002629, договір з МОНУ М/27-2022 від 23.05.2022 р.).

Мета і завдання дослідження. Метою роботи є систематичне дослідження складних систем у вигляді соціальних мереж та мереж економічних зв'язків, та формалізація процесів, що лежать в їх основі. Завдання дослідження полягає у побудові моделей відображення складних систем, розробці програмного комплексу збору даних для проведення досліджень, аналізі структури аудиторій соціальних мереж, дослідженні причин росту аудиторій, аналізі розповсюдження інформації, моделюванні розповсюдження впливу, моделюванні процесів економічних мереж, результати яких використовуються для розв'язання прикладних задач та підтримки прийняття рішень.

Досягнення мети дослідження пов'язано з розв'язанням наступних задач: побудова моделей мереж перетину аудиторій; розробка підходу ідентифікації ключових аудиторій соціальних мереж; розробка методу оцінювання ефективності рекламних кампаній; розробка моделі розповсюдження впливу на основі використання агентів; побудова синтетичної популяції; розробка агентної моделі; формалізація процесу калібрування агентної моделі; дослідження прогнозів агентної моделі.

Об'єктом дослідження є процеси взаємодії користувачів в соціальних мережах, що спричиняють розповсюдження впливу, процеси економічних відносини між населенням в умовах пандемії.

Предмет дослідження – моделі та методи дослідження розповсюдження інформації та впливу в соціальних мережах, моделі взаємодій в мережах економічних відносин.

Методи дослідження. Основу методології складає основний підхід системного аналізу, що полягає в розбитті об'єкту дослідження на незалежні компоненти, та застосування до них відповідних методів. Серед яких агентне моделювання, аналіз складних мереж, методи теорії оптимізації, методи теорії прийняття рішень, елементи теорії графів, алгоритм переважного приєднання, алгоритм імітації відпалу, еволюційні стратегії.

Наукова новизна отриманих результатів. У процесі розв'язання задач дослідження отримано нові наукові результати, які полягають в наступному:

- запропоновано підхід до визначення перетину аудиторій соціальних мереж, застосовано його до розв'язання проблеми побудови мережі перетину аудиторій на публічних даних.
- запропоновано підхід до визначення ефективності рекламних кампаній.
- вперше запропоновано підхід до розповсюдження інформації на основі агентів з використання оптимального розташування.
- розроблено підхід до побудови синтетичних популяцій на основі публічних даних.
- розроблено агенту модель мережі економічних зв'язків.
- розроблено програмний комплекс застосування агентної моделі економічних зв'язків для побудови точних прогнозів.

Практичне значення отриманих результатів полягає в можливості негайного впровадження результатів, отриманих в дисертаційній роботі, для застосування при аналізі оптимальної рекламної стратегії та розробки системи розповсюдження впливу. Результати дисертаційного дослідження використовувались у навчальному процесі на факультеті кібернетики Київського національного університету імені Тараса Шевченка при підготовці лекційних та спеціальних курсів «Сучасні інформаційні технології», «Невизначеність та кластеризація» та «Моделі оптимізації систем».

Особистий внесок здобувача. Дисертація є самостійною науковою працею, в якій висвітлені власні ідеї та розробки автора, що дозволили досягти поставленої мети. Наукові положення, пропозиції та рекомендації, що

вносяться на захист, отримані здобувачем самостійно. У спільних роботах із науковим керівником Є.В.Івохіним, особистий внесок полягає у наступному: у роботі [4], формалізовано процес побудови мереж перетину аудиторій на основі публічних даних, у роботі [11] досліджено математичне формулювання проблеми оптимального розподілу потужності каналів передачі даних.

Апробація результатів дисертації. Матеріали дисертаційної роботи доповідалися та обговорювалися на наукових конференціях та семінарах:

III International Conference Challenges to National Defence in Contemporary Geopolitical Situation (CNDCGS'2022) (*Vilnius, Lithuania*, вересень, 2022); міжнародній науково-технічна конференція “ABIA-2023” (Київ, Україна, квітень, 2023); Міжнародній науковій конференції “Problems of Decision Making under Uncertainties” (Baku-Sheki, Republic of Azerbaijan, травень, 2020; Skhidnytsia, Ukraine, травень, 2021; Київ, Україна, листопад, 2022); Міжнародній науково-практичній конференції «Інформаційні технології в культурі, мистецтві, освіті, науці, економіці та бізнесі» (Київ, Україна, квітень, 2022); Міжнародній науково-практичній конференції "Shevchenkivska Vesna – 2021" (Kyiv, Ukraine, квітень, 2021); на наукових семінарах факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка.

Публікації. Основні наукові твердження, висновки і результати дисертаційної роботи опубліковані в 14 наукових працях. З них – 6 наукових статей, у тому числі 4 у фахових виданнях [1-4], 1 стаття у виданні, яке входить до наукометричної бази даних [1], 1 публікація - розділ закордонної монографії (Польща) [6], 1 публікація у працях конференції, що входить до бази SCOPUS [5], 7 – праці та тези наукових конференцій [7-13].

Структура дисертації. Дисертаційна робота складається з вступу, чотирьох розділів, висновку, списку використаних джерел з 71 найменувань (на 7 сторінках) і трьох додатків (на 4 сторінках). Загальний обсяг роботи становить 137 сторінок.

РОЗДІЛ 1: ОГЛЯД СУЧАСНОГО СТАНУ ДОСЛІДЖЕНЬ СОЦІАЛЬНИХ ТА ЕКОНОМІЧНИХ МЕРЕЖ

Для проведення всебічного огляду літератури, було використано низку спеціалізованих баз даних та ресурсів для збору дослідницьких робіт і наукових статей. Ці ресурси відомі своїми значними колекціями академічних матеріалів, що охоплюють різні дисципліни, в тому числі й предмет дослідження дисертаційної роботи:

- Google Scholar
- ResearchGate
- arXiv
- ScienceDirect
- The National Bureau of Economic Research

Також було використано пошукову систему Google разом із вищенаведеними ресурсами, для того, щоб забезпечити включення широкого спектру досліджень та точок зору, роблячи комплексний і вичерпний огляд літери для дисертаційної роботи.

1.1 ОСНОВНІ АСПЕКТИ ДОСЛІДЖЕННЯ СКЛАДНИХ МЕРЕЖ

Складні мережі

Складні мережі стали важливим об'єктом наукових досліджень, охоплюючи ряд дисциплін, таких як фізика, комп'ютерні науки, біологія то соціологія [1] [2]. Відображаючи системи як мережі зв'язаних вузлів, складні мережі надають основу для глибокого розуміння динаміки процесів та структури різних складних

систем, від біологічних до соціальних мереж. Метою цих досліджень є виявлення та ідентифікація закономірностей та принципів на яких базуються такі системи [1].

Основними елементами складної мережі є вершини (вузли) та ребра (зв'язки) [1]. Складні мережі володіють низкою характеристик, що роблять їх відмінними від більш простих мереж. Наприклад, в багатьох мережах, що описують реальні системи, ступінь розподілу вершин відповідає степеневому розподілу (що робить їх безмасштабними), що вказує на наявність в мережі кількох тісно пов'язаних вершин. Інші важливі характеристики включають коефіцієнт кластеризації та середню довжину шляхів, які дозволяють ідентифікувати мережу "Тісний світ" з високою кількістю взаємопов'язаних вершин та короткою відстанню між ними [3].

Існує багато моделей складних мереж, розроблених для відображення різних реальних мереж. Модель Ердеша - Реньї формує випадкову мережу з пуассонівським розподілом вершин [4], в той же час модель Барабаші використовує принцип переваги зв'язків (переважного приєднання), створюючи безмасштабні мережі. [5]. Модель Ваттса - Строгаца допомагає генерувати мережі, що показують властивості "Тісний світ", типові для багатьох реальних мереж [3] [2].

Для аналізу складних мереж застосовуються різноманітні математичні методи, теорія графів та комп'ютерне моделювання. Застосування теорії графів дозволяє зрозуміти базову структуру мережі, наприклад центральність вершин, що є одним з початкових критеріїв оцінки його впливу на мережу [6]. Алгоритми виявлення спільнот допомагають розкривати модульну структуру мережі, групуючи вершини з більшою кількістю спільних зв'язків [2] [7]. Використання комбінації цих інструментів дозволяє розкрити природу зв'язків між вершинами та прогнозувати процеси, що відбуваються в складних мережах.

Аналіз складних мереж має систематичне практичне застосування. У сфері охорони здоров'я моделі мереж описують динаміку поширення захворювань,

дозволяючи ефективно приймати рішення для запобігання їх розповсюдження [1] [8]. Аналіз соціальних мереж кардинально покращив розуміння соціальних структур та людської поведінки [9]. Більше того, теорія складних мереж знайшла ідеальне застосування при дослідженні мережі Інтернет та онлайн-зв'язків між різними її сутностями, що надало уявлення про їхню ефективність, стабільність та вразливість [2] [10].

Синтетичні популяції

Синтетичні популяції є методом створення реалістичних відображень популяцій без розкриття особистих даних населення [11]. Синтетичні популяції застосовуються в ряді галузей, включаючи охорону здоров'я для вивчення поширення захворювань, транспортне моделювання для моделювання маршрутів [12], та маркетинг для розуміння споживчої поведінки, аналізуючи складні зв'язки та тенденції серед різноманітних груп населення. Основна мета синтетичної популяції полягає в імітації реальних демографічних характеристик і звичок, при цьому забезпечуючи анонімність, що робить їх корисними для аналізу сценаріїв та підтримки прийняття рішень [13].

Для створення синтетичної популяції використовуються агреговані дані та мікро джерела інформації. Зазвичай це включає у себе створення моделей осіб та господарств, базуючись на даних перепису та інших ресурсах [11]. Створена таким чином синтетична популяція відповідає загальним характеристикам реального населення, в той час як детальні зв'язки формуються за допомогою спеціалізованих статистичних методів [13]. Отриманий результат представляє собою групу синтетичних сутностей, які відображають реальне населення, але не є його точною копією.

Для створення синтетичних популяцій використовуються різноманітні моделі та техніки. Процедура ітеративного пропорційного пристосування (*IPF*) є одним з поширених підходів, яка коригує спільний розподіл змінних так, щоб

відповідати відомим маржинальним характеристикам [14]. Однак *IPF* не завжди зберігає високопорядкові взаємозв'язки між змінними. Останнім часом були запропоновані такі методи як комбінаторна оптимізація, максимізація ентропії та машинне навчання для вирішення цієї проблеми [15].

Синтетичні популяції використовуються як вхідні дані для агентних моделей, в яких симулюються поведінка окремих агентів для аналізу результатів на системному рівні. За допомогою агентних моделей аналітики можуть досліджувати різні сценарії, розуміти виникаючі явища та тестувати втручання у контрольованому середовищі [16]. Крім того, використання синтетичних популяцій в агентних моделях дозволяє досліджувати гетерогенність у поведінці та результатах, що не завжди можливо при традиційних методах моделювання.

Агентне моделювання

Агентне моделювання є високоефективним обчислювальним підходом для аналізу кожної з складових складної систем [17]. Агентне моделювання базується розбитті складної системи на індивідуальні сутності – агенти, які взаємодіють між собою на основі наперед встановлених правил (поведінок) [18]. Хоча поведінка окремого агента може здаватися простою, їх колективна взаємодія може породжувати нові явища на макрорівні, допомагаючи розкрити приховані аспекти взаємодії у складних системах [19]. Іншим важливим структурним елементом агентної моделі є середовище, в якому оперують агенти. Середовище може бути конкретним (як от географічна зона) або більш узагальненим (як мережева структура). Агенти розуміють зміни в власному середовищі та адаптують поведінку, що відображається під час прийняття ними рішень. Поведінки агентів можуть бути визначені як детерміновані або ж в випадкові, регулюючи взаємодію агентів між собою та середовищем.

Агентне моделювання може бути застосоване в широкому спектрі дисциплін, таких як економіка, соціологія, екологія та епідеміологія. Класичний

приклад застосування такого моделювання - модель сегрегації Шилінга, в якій простий набір поведінок агентів може призвести до виникнення систематичної сегрегації на рівні всієї системи [20]. Аналогічно, у епідеміології агентне моделювання використовується для моделювання розповсюдження захворювань, де кожен агент може інфікувати іншого на основі їх мобільності та положення в середовищі [21]. В економіці, агентне моделювання пояснює роботу фінансових ринків [22], а в транспортному плануванні – поведінку людей під час подорожей.

Аналіз результатів агентного моделювання зазвичай включає вивчення виникаючих явищ на макрорівні породжених взаємодіями на мікро-рівні [16]. Поширеними методам для вивчення поведінки моделі в різних сценаріях є аналіз чутливості, стабільності параметрів системи, та симуляції Монте - Карло, статистичні методи та методи візуалізації.

1.2 МОДЕЛЮВАННЯ ТА ОПТИМІЗАЦІЯ РОЗПОВСЮДЖЕННЯ ВПЛИВУ В СОЦІАЛЬНИХ МЕРЕЖАХ

1.2.1 Критерії пошуку актуальних досліджень

Огляд літератури було проведено з використанням наступних ключових слів та виразів:

- Симуляції / моделювання впливу.
- Оптимізація / максимізація впливу.
- Моделі розповсюдження трендів / віральність / поширення чуток / дезінформації.
- Агенти моделі / кількісні моделі / моделі симуляції впливу / розповсюдження інформації / маніпуляції думок.

- Вплив / поширення інформації / дифузія / інформаційні каскади.
- Вимірювання / прогнозування впливу.

Кожен пошук, був додатково уточнений з використанням наступних модифікаторів в комбінації з ключовими словами:

- Соціальна мережа.
- Онлайн-мережа.
- Онлайн-спільнота.
- Facebook / Instagram / Twitter / Telegram / YouTube.

1.2.2 Порівняльна характеристика актуальних досліджень

Для кращого розуміння відмінностей у підходах використаних дослідниками, для кожного дослідження наведеного в Табл. 1. було виділено три основні компоненти:

- **Набір даних:** конкретний набір даних на якому було проведено дослідження.
- **Методологія:** алгоритми та підходи використані в дослідженні.
- **Оптимізація впливу:** підхід, використаний дослідниками, для розв'язання проблеми максимізації та оптимізації впливу в мережі.

Табл. 1. Порівняльний аналіз досліджень, зосереджених на моделюванні та оптимізації впливу в соціальних мережах.

Дослідження	Набір даних	Методологія	Оптимізація впливу
Simulating the spatial diffusion of memes on social media networks [23]	Мережі “Тісний світ”, випадкові мережі, та мережі переважного приєднання.	Агентне моделювання дифузії інформації	Випадкова вибірка цілей, на основі центральності
Opinion mining in social media: Modeling, simulating,	Часткові дані соціальних мереж	Агентне моделювання, виявлення думок, симуляції	Відсутня

and forecasting political opinions in the web [24]		розповсюдження думок, аналіз тональності	
Polarization in social media assists influencers to become more influential: analysis and two inoculation strategies [25]	Довільний набір даних	Модель динамічної комунікації, моделювання розповсюдження публікацій	Відсутня
The Effect of People Recommenders on Echo Chambers and Polarization [26]	Випадкові мережі	Метод Монте-Карло, симуляції рекомендацій, модель розповсюдження думок, динамічні моделі	Відсутня
A two-stage VIKOR assisted multi-operator differential evolution approach for Influence Maximization in social networks [27]	Дві синтетичні мережі та вісім реалістичних мереж	Багатокритеріальний підхід прийняття рішень, модифікований диференціальний еволюційний алгоритм, евристичний жадібний алгоритм	Мінімізація простору пошуку цілей, оптимізація впливу, шляхом вибору цілей-лідерів думок на основі їх аудиторій

Network segregation and the propagation of misinformation [28]	Краудсорсингові дані платформи <i>Amazon Mechanical Turk</i>	Агентне моделювання дифузії інформації	Випадкова вибірка цілей
Social Media Networks, Fake News, and Polarization [29]	Синтетична мережа, що імітує соціальну мережу <i>Twitter</i>	Агентне моделювання, метод Монте-Карло, Регресійний аналіз	Рівномірний розподіл агентів-цілей за спектром їх переконань
A Framework for Analyzing Influencer Marketing in Social Networks: Selection and Scheduling of Influencers [30]	Дані соціальної мережі <i>Twitter</i> , що складається з публікацій 37 лідерів думок	Поліноміальні моделі	Оптимізаційна модель вибору цілей-лідерів думок для ефективної рекламної кампанії
Influence maximization on temporal networks: a review [31]	Довільний набір даних	Дифузійні моделі, Жадібний алгоритм максимізації впливу	Вибірка цілей на основі жадібного алгоритму, ранжування цілей, послідовний підбір цілей

Evaluating Stochastic Seeding Strategies in Networks [32]	Експериментальний набори даних соціальних зв'язків у сільській місцевості	Метод Монте-Карло	Стохастичний вибір цілей
Ambassadors of ideology: A conceptualization and computational investigation of far-right influencers, their networking structures, and communication practices [33]	Набір даних з 700 лідерів думок в <i>Telegram</i>	Аналіз структури мережі	Попередньо обрані цілі
Deep Graph Representation Learning and Optimization for Influence Maximization [34]	Синтетичні мережі, мережі співпраці джазових, мережі науково-дослідних робіт в сфері інформатики, мережа енергосистеми США, мережа співавторства	Система на основі дифузії інформації	Оптимальний вибір цілей на основі центральності цілей

Soft rumor control in mobile instant messengers [35]	Синтетична мережа повідомлень мобільних месенджерів	Марківський процес прийняття рішень	Відсутня
Evaluating the Influence of Twitter Bots via Agent-Based Social Simulation [36]	Набір даних соціальної мережі <i>Twitter</i>	Агентне моделювання	Випадкова вибірка цілей

1.2.3 Ключові характеристики актуальних досліджень

Висновок 1

Дослідники, як правило, зосереджуються на методології та алгоритмах більше, ніж над якістю даних. Це не дивно, адже отримувати дані від соціальних мереж стає дедалі складніше, зокрема через постійні зміни в їхній політиці конфіденційності (див. зміни до Facebook API після скандалу з *Cambridge Analytica* в 2020 році [37]), та змінах в бізнес вимогах (див. соціальна мережа Twitter обмежила доступ до API для академічних потреб в 2023 році [38]) щодо збору даних.

Тому, в дисертаційній роботі було поставлено за мету побудувати повнорозмірну структуру соціальних мереж з реалістичною топологією, що стала базою для проведення всіх досліджень представлених в дисертаційній роботі.

Висновок 2

В більшості випадків, дослідники використовували епідеміологічні моделі або “фізичні” моделі (Модель Ізінга, модель *SIR*, моделі дифузії) для опису користувачів в соціальних мережах. В той же час, дослідники рідко збирають та аналізують мікро-дані соціальних мереж, та не використовують їх при моделюванні поведінки окремих користувачів. В свою чергу такі моделі стають абстрактними, та ігнорують специфіку та особливості функціонування різних соціальних мереж.

Тому, модель описана в дисертаційній роботі є орієнтованою на конкретну соціальну мережу, і повністю враховує способи комунікації і взаємодії між користувачами, в межах системних обмежень.

Висновок 3

Розглянуті моделі моделювання впливу припускають, що цілі (користувачі, лідери думок) можуть бути переконаними до зміни ставлення на бажане, що не завжди відображає реальність. Для прикладу, рекламні кампанії з використанням лідерів думок зазвичай потребують довгих переговорів, тільки після яких, може бути впевненість про підтримку того чи іншого лідера думок. Аналогічно, перед тим як бот-мережі можуть стати ефективними, вони потребують часу для “акліматизації” в цільових аудиторіях.

Тому, модель впливу запропонована в дисертаційній роботі буде реалістичною в цьому аспекті. В свою чергу, механізм моделі описаний не тільки з точки зору кінцевого результату, але і проміжних кроків, необхідних для його досягнення.

1.3 МОДЕЛЮВАННЯ НАСЛІДКІВ ПАНДЕМІЙ НА ЕКОНОМІЧНІ МЕРЕЖІ

1.3.1 Критерії пошуку актуальних досліджень

Огляд літератури було проведено з використанням наступних ключових слів та виразів:

- Симуляції / моделювання пандемій / епідемій.
- Економічні наслідки / мережі.
- Законодавчі заходи.
- Вимірювання/ прогнозування економіки.

- Агентні моделі динаміки економіки.
- Прогнозування ланцюгів поставок.

Кожен пошук, був додатково уточнений з використанням наступних модифікаторів в комбінації з ключовими словами:

- Світова / національно / регіональна / місцева економіка.
- Специфічні сектори економіки (охорона здоров'я, туризм, роздрібна торгівля, тощо.).
- Економічні показники (ВВП, рівень безробіття, інфляція, тощо.).
- 2020 / пандемія COVID-19.
- Пост пандемічна економіка.

1.3.2 Порівняльна характеристика актуальних досліджень

Для кращого розуміння відмінностей у підходах використаних дослідниками, для кожного дослідження наведеного в Табл. 2. було виділено п'ять основних компонент:

- Рік: рік у якому дослідження було опубліковане.
- Набір даних: конкретний набір даних на якому було проведено дослідження.
- Методологія: алгоритми та підходи використані в дослідженні.
- Прогнозування: який підхід автори використали для прогнозування.
- Валідація: чи проводили автори перевірку результатів власних прогнозів.

Табл. 2. Порівняльний аналіз досліджень, зосереджених на аналізі впливу пандемій на економічні мережі.

Дослідження	Рік	Набір даних	Методологія	Прогнозування	Валідація
Predicting the economic impact of the <i>COVID-19</i> pandemic in the United Kingdom using time-series mining [39]	2021	Щомісячні, щоквартальні та щорічні дані ВВП Великої Британії за секторами економіки (сфера послуг,	Модель авторегресійної інтегрованої ковзної середньої	Відсутнє	Відсутня

		виробництво та будівництво)			
Self-correcting error-based prediction model for the <i>COVID-19</i> pandemic and analysis of economic impacts [40]	2021	Витрати за кредитними та дебетовими картками за секторами економіки в деяких штатах США	Рекурентна нейронна мережа	Відсутнє	Відсутня
Forecasting tourism recovery amid <i>COVID-19</i> [41]	2021	Щоквартальна кількість туристів, ВВП, індекс споживчих цін, валютні курси.	Панель експертів	Квартальна динаміка відновлення економіки	Панель експертів
Economic recovery forecasts under impacts of <i>COVID-19</i> [42]	2022	ВВП, індекс споживчих цін, об'єми і вартість експортованих та імпортованих товарів, рівень безробіття, валютні курси (долар США до місцевої	Байєсова авторегресивна модель, модель SIR	Прогноз на 5 місяців	Відсутня

		валюти), та міжнародні резерви			
Epidemiological and economic impact of <i>COVID-19</i> in the US [43]	2020	Дані перепису населення США, дані про щоденну активності національного опитування, дані мобільності населення, параметри моделі вірусу від Центру контролю та профілактики захворювань США, дані про взаємозалежність між секторами економіки США, телекомунікаційні дані	Аналіз мережі соціальних контактів, модель <i>SIR</i>	400-денний прогноз втрат окремих секторів економіки	Відсутня

<i>COVID-19</i> and trade: Simulated asymmetric loss [44]	2021	Дані секторів економіки, ВВП, заробітні плати, та вартості тарифів	Обчислювальна модель загальної	Вплив на міжнародну торгівлю	Відсутня
Socioeconomic Network Heterogeneity and Pandemic Policy Response [45]	2020	Штучна популяція, телекомунікаційні дані	Агентна епідеміологічна модель	Дані перебігу пандемії (інфекції, госпіталізації, реанімації, та смерті)	Відсутня
The economic impact of <i>COVID-19</i> interventions: A mathematical modeling approach [46]	2022	Дані про соціальне дистанціювання в Південній Кореї	Математична модель з віковою структурою динаміки передачі <i>COVID-19</i>	Макроекономічні оцінки стану економіки	На основі агрегованих національних даних

1.3.3 Ключові характеристики актуальних досліджень

Висновок 1

Дослідники свідомо ігнорують динамічну природу економічних мереж обираючи хибну методологію моделювання. Пандемія та хаотичні дії урядів для подолання її наслідків стали непередбачуваними подіями, для яких відсутні історичні аналоги. Незважаючи на це, дослідники все ж вдалися до моделювання використовуючи підходи економетрії, авторегресивні та моделі рівноваги.

Тому, в межах дисертаційної роботи було використано агентне моделювання для опису впливу пандемії на економічні мережі, що дозволить:

- a) Аналізувати перехідні процеси, враховуючи динамічну природу пандемії.
- b) Прогнозувати динаміку економічних мереж, враховуючи відсутність історичних даних.
- c) Врахувати поведінку окремих агентів базуючись на базових принципах функціонування економіки (максимізація прибутків компаніями).

Запропонована агентна модель включаю не лише перехідну динаміку, але й враховує обмеження, які відповідають принципам ринково-обмеженої раціональності. Крім того, мікроекономічні вихідні дані моделі використовують під час процесу калібрування.

Висновок 2

Незважаючи на динамічну природу розглянутих моделей, багато з них здані прогнозувати динаміку системи лише на довготривалих часових інтервалах. Внаслідок цього, автори часто зіштовхуються з проблемою перевірки якості прогнозів.

Тому модель запропонована в дисертаційній роботі характеризується здатністю до генерації прогнозів з високою частотою. Така модель може бути перевіреною на реальних макроекономічних даних, що надходять до публічного доступу без суттєвих затримок часу. З кожною наступною ітерацією, модель буде включати дані згенеровані на попередніх ітерація, дозволяючи проводити динамічні калібрування за допомогою даних про безробіття, одночасно генеруючи нові прогнози. Така характеристика є критично важливою для моделей, призначених для використання в динамічних та хаотичних умовах, прикладом чого є пандемія *COVID-19* 2020 року.

Висновок 3

Використання синтетичної популяції в комбінації з агентним моделюванням дозволило запропонованій моделі стати корисною для прийняття рішень, особливо на місцевому рівні. Це пояснюється тим, що як правило, дії на покращення економічної ситуації спрямовані на конкретні соціальні та демографічні групи. Всі розглянуті моделі виявилися неефективними при комунікації з органами прийняття рішення, через відсутність конкретної та зрозумілої термінології. Наприклад, місцева влада хоче зрозуміти скільки домогосподарств, що складаються з п'яти осіб, повинні отримувати продовольчі талони в окрузі Гаваї. У той час, як моделі рівноваги не можуть надати такої інформації, запропонована модель сприяє ефективному спілкуванню з політиками та органами прийняття рішень, використовуючи зрозумілу для них термінологію.

Розглянуті моделі не включали або ж унеможлилювали врахування реальних дій, що приймалися для впливу на перебіг епідемії. Важливо зазначити, що ці дії були спрямовані на осіб, компаній, та домогосподарства на основі демографічних, географічних, економічних і соціальних факторів, які не були явно представлені в моделях. Тому, запропонована в дисертаційній роботі модель

базується на використанні синтетичної популяції. Це дозволяє не тільки моделювати рішення, але й відобразити їх наслідки на існуючу синтетичну популяцію. Для того, щоб модель була актуальною з точки зору прийняття рішень, вона була розроблена як повністю рекурсивна.

1.4 ВИСНОВКИ ПО ПЕРШОМУ РОЗДІЛІ

У першому розділі дисертаційної роботи було проведено огляд актуальних досліджень розповсюдження впливу та інформації в соціальних мережах. Детально проаналізовано основні особливості моделювання процесу розповсюдження інформації в контексті наборів даних на яких проводилося дослідження, обраної методології та розв'язання проблеми вибору оптимальних цільових аудиторій для оптимізації впливу. Виходячи з результатів розглянутих досліджень, було встановлено принципові недоліки підходів до дослідження розповсюдження інформації, що характеризуються системною відсутністю реалістичних моделей соціальних мереж на основі публічних даних та застосуванню методології, що не здатна описувати динамічну структуру процесів та взаємодій в соціальних мережах та соціальних мереж загалом.

Було розглянуто актуальні дослідження соціальних зв'язків в економічних мережах, як важливого аспекту соціальної взаємодії. Детально проаналізовано основні особливості моделювання процесів в економічних мережах. Виходячи з результатів розглянутих досліджень, було встановлено схожі недоліки моделювання, в основі яких – відсутність моделей реальних економічних мереж, а також методології, що не враховує поведінковий аспект.

Для моделювання розповсюдження інформації в соціальних мережах та моделювання процесів в економічних мережах, було запропоновано використати методологію аналізу складних мереж з використанням агентного моделювання та синтетичних популяцій.

РОЗДІЛ 2: ФОРМАЛІЗАЦІЯ ПРОЦЕСУ РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ В СОЦІАЛЬНИХ МЕРЕЖАХ

For millions of years, mankind lived just like the animals. Then something happened which unleashed the power of our imagination. We learned to talk and we learned to listen. Speech has allowed the communication of ideas, enabling human beings to work together to build the impossible. Mankind's greatest achievements have come about by talking, and its greatest failures by not talking. It doesn't have to be like this. Our greatest hopes could become reality in the future. With the technology at our disposal, the possibilities are unbounded. All we need to do is make sure we keep talking. - Stephen Hawking

З самого початку свого існування люди займаються створенням та розповсюдженням інформації, використовуючи її для різноманітних цілей: розуміння свого оточення, розуміння природних явищ, покращення власного добробуту, тощо. Передача інформації між людьми відіграла ключову роль під час знакових подій в світі. На кожному етапі розвитку людства, найактуальніші технологічні розробки забезпечували масове розповсюдження інформації. Для прикладу, основним способом розповсюдження інформації впродовж сотень років, були друковані джерела, а згодом телебачення.

Розповсюдження інформації базується на глибинних потребах людей обмінюватися інформацією, а тому завжди існували особи, що прагнули її спотворити та використати для власних цілей. Тому, власники друкарень а згодом і власники радіо-станцій та телевізійних каналів володіли надзвичайними можливостями впливу на сприйняття людьми інформації про події, таким чином маніпулюючи виборами та починаючи війни. Соціальні мережі стали черговою технологічною еволюцією пропонуючи небачену швидкість розповсюдження інформації. Більшість соціальних мереж зароджувалися як засоби обміну миттєвими повідомленнями та майданчиками для обговорень, переростаючи в

незамінний атрибут повсякденного життя, що дозволяє слідкувати за подіями в світі в режимі реального часу. Такі особливості соціальних мереж, а також їх потенційне використання для розповсюдження хибної інформації породжують необхідність розуміння фундаментальних механізмів їх функціонування.

2.1 МЕХАНІЗМИ РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ

Основою соціальних мереж є обмін публікаціями (контентом) користувачів. Соціальні мережі пропонують різні механізми обміну публікаціями: *retweet* в *Twitter*, *share* в *Facebook*, *repost* в *Instagram*, які в комбінації дозволяють інформації розповсюджуватися в мережі. Коли один користувач в *Facebook* робить публікацію, всі його друзі бачать цю публікацію в своїй стрічці, і можуть залишати коментарі або вподобання. Проте, користувачі можуть блокувати інших користувачів, зокрема і друзів. В такому випадку зв'язок дружби зберігається, проте користувач не бачить нових публікацій.

В інших соціальних мережах, на рівні з користувачами існують сторінки, які також мають можливість створювати публікації. Користувачі можуть слідкувати за сторінками та отримувати їх публікації.

Разом з тим, соціальні мережі володіють механізмом реклами. Реклама в соціальних мережах, зазвичай базуються на ключових характеристиках користувачів, наприклад: поточне місто проживання, інтереси, вік, стать, відвідані місця. Таким чином, рекламодавці можуть збудувати цільову аудиторію, для показу специфічної реклами.

Безпосередні дії користувачів відіграють важливу роль у процесі розповсюдження інформації, проте алгоритми рекомендацій соціальних мереж відіграють ключову роль. Такі алгоритми функціонують подібно до реклами, підбираючи публікації схожі за інтересами користувачів, та саме вони підтримують еволюцію соціальної мережі.

Окремі публікації в соціальних мережах універсально резонують з інтересами користувачів стаючи вірусними, водночас певні об'єднання користувачів за інтересами стають нішевими, та ніколи не рекомендуються іншим користувачам.

2.2 МОДЕЛІ РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ

Базуючись на хаотичній структурі появи публікацій в соціальних мережах, можна припустити, що процес розповсюдження інформації в соціальних мережах може бути описаний з використанням моделей дифузії. Відповідно до [47] поширення інформації в суспільстві базується на категоризації людей на основі їх ймовірності сприйняття та прийняття інновацій, та може бути представлене у вигляді *S*-кривої. В такому випадку інформація поширюється від новаторів до перших послідовників, потім до ранньої та пізньої більшості, і в кінці досягаючи відстаючих. Така модель пропонує основу для розуміння динаміки дифузії, проте не враховує особливості функціонування соціальних мереж як складних мереж.

Враховуючи складну структуру зв'язків в соціальних мережах, необхідно також враховувати динамічний поріг сприйняття інформації користувачами. Порогова модель [48], пропонує розширення класичної моделі дифузії, визначаючи поріг сприйняття як властивість соціального кола користувача, в якому він адаптує його інтереси. Таким чином, користувач ймовірніше прийме позицію своїх найближчих зв'язків, ніж нову динамічну позицію мережі.

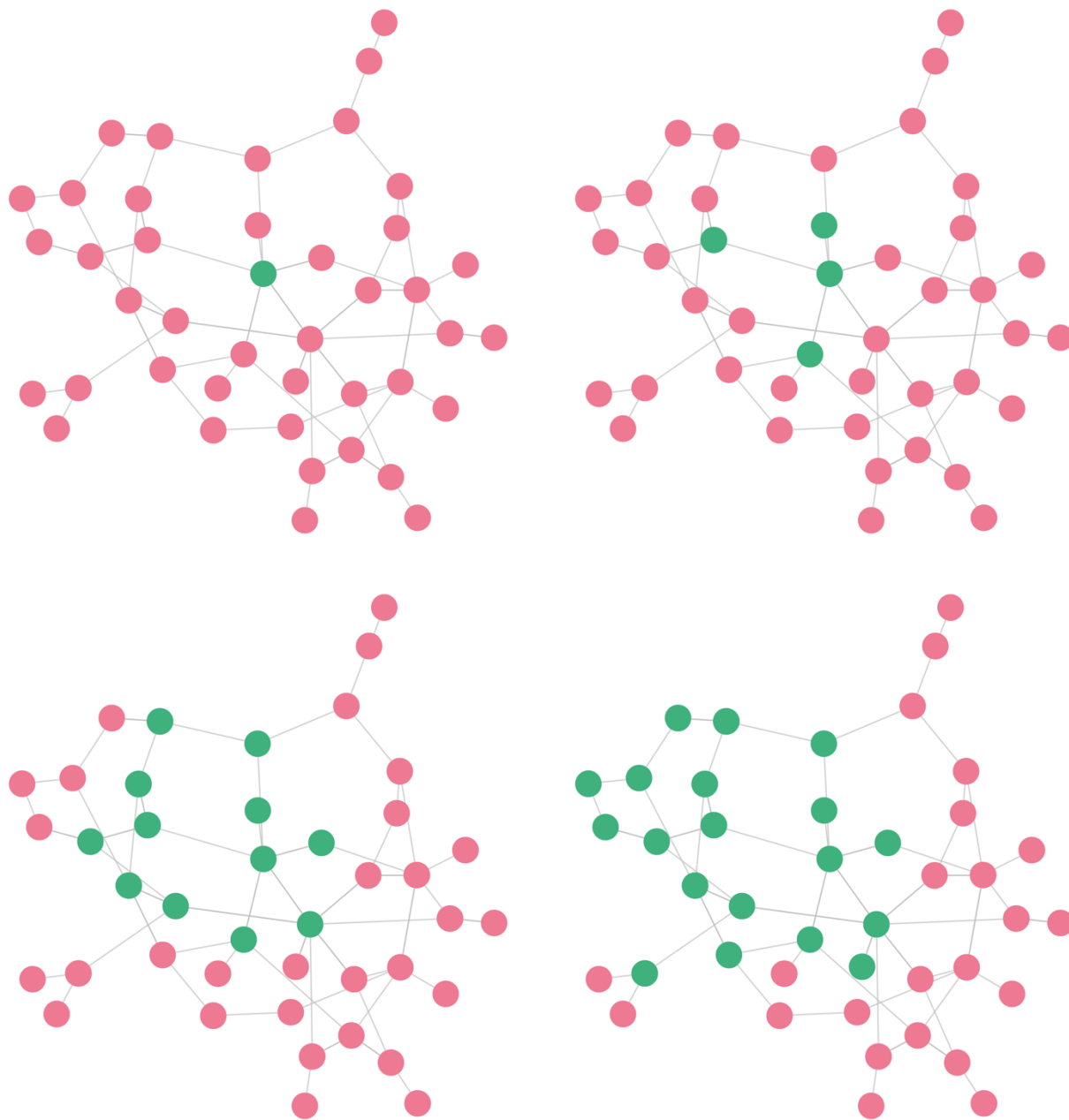


Рис. 2.1. Приклад дифузії інформації на випадковій мережі дружніх зв'язків.

Моделі дифузії розповсюдження інформації в соціальних мережах дають можливість моделювати ефективність рекламних кампаній, інформованості та актуальності інформації [49].

На Рис. 2.1 зображено еволюцію випадкової мережі, в контексті поширення інформації (публікації) з використанням дифузії та випадкових порогів сприйняття. Встановлення значень порогів в реальних соціальних мережах не є

реалістичним в умовах обмеженої кількості даних, неоднорідності поглядів користувачів та динамічної природи мереж.

Для проведення моделювання розповсюдження інформації в соціальних мережах доцільно використовувати агентне моделювання, в якому система агентів що діють на основі простих правил, дозволяє виявляти складну поведінку та прогнозувати еволюцію мережі в цілому [50].

2.3 ДИНАМІЧНІ ФАКТОРИ ВПЛИВУ НА РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ

Соціальні мережі налічують рекорду кількість користувачів, тільки в соціальній мережі *Facebook* станом на 2023 рік, налічується 2.989 млрд. унікальних користувачів [51]. Кожен з них регулярно використовує і інші соціальні мережі, читаючи публікації, реагуючи на новини. Незважаючи на таку кількість користувачів, більшість публікацій в соціальних мережах рідко переглядаються іншими користувачами. Лише обрані публікації отримують загально-мережеве визнання та впливають на погляди користувачів.

Існує багато факторів які можуть вплинути на популярність публікацій:

- **Актуальність:** новизна та актуальність вмісту публікацій часто спричиняють високий рівень зацікавленості серед користувачів.
- **Емоційний резонанс:** публікації які апелюють до емоцій користувачів мають більшу ймовірність бути пересланими. Такі публікації резонують з аудиторіями, змушуючи користувачів створювати власні публікації на дану тему [52].
- **Візуальні засоби:** зображення, відео та інші візуальні елементи збільшують зацікавленість аудиторії в публікації.

- Інтерактивність: публікації, які заохочують користувачів до дії, такі як опитування чи виклики, сприяють більшому поширенню.
- Лідери думок: окремі користувачі завдяки автентичності, швидкості реакції на події та якості публікацій мають значний вплив на своїх послідовників, які реагують на нові публікації [53].
- Алгоритми рекомендацій: кожна соціальна мережа володіє власним алгоритмом рекомендацій, виділяючи пріоритетні публікації для користувачів.
- Спільноти: користувачі окремих спеціалізованих спільнот взаємодіють з публікаціями, збільшуючи його популярність.

Використання комбінації таких факторів, з використанням психологічних та культурних особливостей користувачів дозволяє перетворювати публікації у засіб впливу на користувачів в соціальних мережах.

2.4 ВИСНОВКИ ПО ДРУГОМУ РОЗДІЛІ

У другому розділі дисертаційної роботи було проведено формалізацію процесів розповсюдження інформації як основної складової соціальних мереж. Було розглянуто основні механізми та моделі розповсюдження інформації в соціальних мережах.

Було досліджено динамічні фактори впливу на розповсюдження інформації в соціальних мережах, та запропоновано застосування моделювання соціальних мереж, яке базується на дослідженні поведінки користувачів.

РОЗДІЛ 3: АНАЛІЗ РОЗПОВСЮДЖЕННЯ ІНФОРМАЦІЇ В СОЦІАЛЬНИХ МЕРЕЖАХ

Впродовж останніх років соціальні мережі зазнали значної трансформації, перетворюючись з простих сервісів повсякденного спілкування в основне джерело новин та інформації для людей по всьому світу. Соціальні мережі відіграли ключову роль у формуванні значущих подій та явищ останніх років.

Наприклад, під час президентських виборів у США 2016 року соціальні мережі перетворилися на впливові платформи для політичного дискурсу та агітації. Кандидати використовували соціальні мережі для зв'язку з виборцями, розповсюдження своїх ідей, дискредитації опонентів та мобілізації підтримки. Широке охоплення та залученість аудиторій в соціальних мережах дозволило кандидатам швидко поширювати політичні новини, породжувати обговорення та вести кампанії з дезінформації, перетворюючи соціальні мережі на центральне поле битви політичних наративів.

Аналогічно, під час спалаху пандемії *Covid-19*, соціальні мережі стали важливими джерелами інформації для людей у всьому світі. Новини, чутки, медичні рекомендації урядів та особистий досвід людей, пов'язаний з пандемією, швидко поширювалися через соціальні мережі. Вони стали платформою для обміну важливою інформацією, підвищення обізнаності та участі у обговореннях про вірус, його наслідки та необхідні заходи для його локалізації.

Більше того, соціальні мережі відіграли значущу роль в контексті російського вторгнення в Україну, підкреслюючи їх потенційний вплив на геополітичні події в цілому світі. У моменти найбільшої невизначеності та браку достовірної інформації, під час початку конфліктів та криз, соціальні мережі стали платформою для поширення оперативних повідомлень, свідчень очевидців та підвищення міжнародної уваги. Користувачі по всьому світу могли слідкувати за розвитком подій, допомагати потребуючим, висловлювати солідарність та брати участь в обговореннях навколо вторгнення. Соціальні мережі сприяли

поширенню новин, зображень та відео, формуючи громадське сприйняття та реакцію на ситуацію.

Ці приклади є лише окремими випадками та ілюструють, як соціальні мережі все більше стають невід'ємними засобами для спілкування під час критичних моментів, впливаючи на політику, громадське здоров'я та міжнародні відносини. Соціальні мережі продовжують активно розвиватися, що безумовно призведе до збільшені їх впливу на повсякденне життя людей, формування їх думок, громадського обговорення.

3.1 ПІДХОДИ ДО ПОБУДОВИ МОДЕЛІЙ СОЦІАЛЬНИХ МЕРЕЖ

Загалом, аналіз соціальних мереж можна розділити на дві категорії, кожна з яких зосереджується на окремих принципових аспектах функціонування соціальних мереж.

Структурний аналіз соціальних мереж

Структурний аналіз зосереджується на виявленні базових структурних особливостей відносин між користувачами. Використання структурного аналізу дозволяє зрозуміти зв'язки, взаємодію користувачів та процес розповсюдження інформації в соціальних мережах. Серед основних елементів структурного аналізу доцільно виділити:

- Структура мережі: передбачає встановлення і аналіз відносин та взаємодій між користувачами, що дає змогу ідентифікувати потенційно впливових користувачів та спільнот, а також відстежувати розповсюдження інформації.
- Центральність: показники центральності в мережі, такі як, ступінь вершин, ступінь посередництва, та ступінь впливу, кількісно ідентифікують важливість та міру впливу окремих користувачів в межах соціальної

мережі. Ці показники дозволяють ідентифікувати ключових впливових користувачів.

- Виявлення спільнот: методи виявлення спільнот виявляють кластери або групи користувачів які мають найсильніші зв'язки між собою. Ці методи дозволяють виявляти тісно-зв'язані групи користувачів за інтересами, моделями і закономірностями їх поведінки.
- Аналіз розповсюдження інформації: досліджує, як інформація, тренди та ідеї поширюються в соціальних мережах. Зокрема, шляхи та динаміку поширення нової інформації, її сприйняття та роль впливових користувачів.

Аналіз контенту соціальних мереж

Аналіз контенту полягає в дослідженні публікацій, який розміщують користувачі в соціальних мережах, включаючи текст, зображення, відео та інші форми медіа. Використання аналізу контенту дає уявлення про особливості, тренди та емоційне забарвлення, виражене в соціальних мережах. Основні аспекти аналізу контенту включають:

- Текстовий аналіз та обробка природної мови (*NLP*): техніки текстового аналізу, у поєднанні з *NLP*, дозволяють виділяти цінну інформацію з текстових даних, що включає в себе аналіз емоційного забарвлення, моделювання тем та виявлення ключових слів.
- Аналіз зображень і відео: аналіз візуального контенту в соціальних медіа включає в себе такі техніки, як розпізнавання зображень, виявлення об'єктів, аналіз емоційного забарвлення візуального контенту.
- Аналіз трендів: полягає у виявленні трендів, популярних теми та домінуючих напрямів розмов між користувачів, що дозволяє розуміти динаміку та інтереси користувачів.
- Аналіз емоційного забарвлення: зосереджений на розпізнаванні емоційного тону або настрою, вираженого в контексті публікацій

користувачів. Він дозволяє зрозуміти громадську думку, сприйняття реклами і брендів, та ставлення користувачів до конкретних тем.

Поєднання структурного аналізу та аналізу контенту забезпечує всебічне розуміння поведінки користувачів, соціальних динамік, поширення інформації та виникаючих трендів в соціальних мережах. Використання двох категорій надає цілісне уявлення про складні відносини між сутностями які розповсюджують, та контентом який розповсюджується в соціальних мережах. Кожна з категорій аналізу соціальних мереж вимагає попереднього етапу у вигляді збору даних та відновлення структури відносин в соціальних мережах.

Сутності в соціальних мережах

Базис для кожної соціальної мережі складають користувачі, проте кожна з них запроваджує власний підхід до трактування користувача як сутності з певним набором правил та функцій. В більшості випадків користувачі можуть бути представлені як наступні сутності в межах соціальних мереж:

- **Особисті профілі:** більшість соціальних мереж надають користувачам особистий профіль, де вони можуть представити інформацію про себе. Профілі часто включають фотографію, біографію або опис, місцезнаходження, інтереси, вподобання, освіту та професійну діяльність. Користувачі можуть налаштовувати свої профілі так, щоб вони відображали їхню особистість і створювали цифрову присутність в соціальній мережі. Користувачі можуть керувати налаштуваннями приватності свого профілю, визначаючи хто може мати доступ до інформації про них.
- **Сторінки:** деякі соціальні мережі пропонують концепцію сторінок, які є спеціалізованими сутностями доступними для бізнесу, організацій, громадських діячів. Сторінки дають змогу демонструвати бренд, продукцію чи послуги та взаємодіяти із зацікавленою аудиторією. Профілі можуть слідкувати за сторінками та залишати вподобання, що буде

використано соціальною мережею для ідентифікації інтересів профілю. Ті платформи, в яких відсутня сутність сторінки, зазвичай пропонують спеціальну відмітку про автентичність певного профілю.

- Друзі: профілі можуть створювати зв'язки між собою у вигляді друзів у соціальній мережі. Це дозволяє їм отримувати публікації та оновлення від інших профілів.
- Послідовники та підписники: деякі соціальні мережі акцентують увагу на створенні контенту, і пропонують профілям та сторінкам можливість залучати інші профілі у вигляді послідовників та підписників. Ця функція особливо поширена в соціальних мережах, таких як *Instagram*, *Twitter* та *YouTube*.
- Взаємодії: профілі та сторінки мають можливість здійснювати різні взаємодії в межах соціальної мережі, серед яких: вподобання, коментарі, можливість поділитися, підписатися, та безпосередньо обмінюватися приватними повідомленнями.

Сторінки та профілі природньо взаємодіють між собою, і в більшості випадків переймають роль поведінки або послідовника або лідера думок:

- Послідовники: профілі, які зацікавлені в думці інших профілів та сторінок. Вони проявляють своє зацікавлення у вигляді підписування та стеження за сторінками на основі свої інтересів та поглядів. Вони першими отримують сповіщення про нові публікації, першими реагують на нього, залишаючи вподобання, коментарі або ж діляться ними зі своїми друзями.
- Лідери думок: це профілі або сторінки які мають значний вплив в соціальній мережі. Вони мають переважаючу кількість підписників, часто в специфічних сферах. Лідери думки, часто характеризуються своєю креативністю, яка здатна збільшувати аудиторію та кількість реакцій. Саме вони впливають на думки профілів, пропонуючи рекомендації та рекламу. Лідери думки часто співпрацюють з брендами, створюють спонсорований контент або укладають партнерські угоди для монетизації свого впливу.

Хоча не існує універсально точного визначення, яке відрізняє послідовників від лідерів думки, різниця між ними стає більш очевидною при вивченні контексту тем або конкретних предметних областей обговорень в соціальних мережах, де лідери думки зароджують ідеї та починають спілкування між послідовниками.

Важливо розуміти, що ролі послідовників та лідерів думки можуть варіюватися в залежності від різних тем та обговорень. Ті сутності, що в одній предметній області вважаються лідерами думок, можуть вважатися послідовником у іншій, залежно від рівня знань та активності у кожній сфері.

Розповсюдження впливу в соціальних мережах

Виходячи з основних принципів функціонування соціальних мереж, слідує, що розповсюдження інформації в них відбувається завдяки активному створенню публікацій лідерами думок та взаємодії їх послідовників. У контексті взаємодії користувачів, без врахування вбудованих механізмів реклами соціальних мереж, така діяльність призводить до розповсюдження впливу. Саме тому, ідентифікація лідерів думок в межах спеціалізованих тем, а також їх аудиторій у вигляді послідовників є ключовим аспектом дослідження розповсюдження впливу у соціальних мережах. Взаємодії між лідерами думок та їх послідовниками необхідно представити у вигляді складної мережі.

Певні соціальні мережі, такі як *Twitter* або *Reddit* надають необмежений доступ до списків послідовників сторінок, інші такі як *YouTube* або *Facebook* не надають схожого доступу, або ж обмежують його деталізацію, що унеможлиблює побудову складної мережі. В цьому розділі було досліджено методи побудови математичної моделі відображення соціальних мереж, їх недоліки, та оптимізую їх роботу в умовах відсутності достатньої кількості вхідних даних. Зокрема, для ідентифікації лідерів думок та їх аудиторій, було досліджено статистичне відношення між їх публікаціями та відповідним перетином аудиторій. В такому випадку, чим більша подібність публікацій двох

лідерів думок, тим більша імовірність що їх аудиторія складається з тих самих послідовників.

Що робить зв'язок між подібністю публікацій та перетином аудиторій можливим? По-перше, механіка використання соціальних мереж спонукає до такого зв'язку. Наприклад, алгоритми рекомендацій соціальних мереж будують стрічки новин для користувачів на основі подібності публікацій (згадки про актуальні події, використання спільних ключових слів, використання хештегів, прив'язка до локації, тощо). По-друге, поведінка аудиторій також сприяє створенню таких зв'язків. Наприклад, зацікавленість користувачів у певній, специфічній тематиці, призводить до пошуку ним лідерів думок з подібними інтересами що збільшує імовірність підписування на лідерів думок зі захожими публікаціями. Зрештою, імітація контенту лідерами думок, їх співпраця та реклама при залученні аудиторій (кампанії типу "підпишись на друга"), або ж реклама для підписників конкуруючого лідера думок спонукають аудиторію підписуватися на лідерів думок із подібними публікаціями.

В цьому розділі, було продемонстровано як прості алгоритми для оцінки подібності публікації на основі виключно публічних даних дають статистично значущі результати. Це дає можливість використати в подальшому подібність публікацій як вхідні дані для більш загальних статистичних моделей, що враховують інші відносини між лідерами думок, такі як їх пряму взаємодію в процесі розповсюдження інформації, їх поведінку у спільних розмовах, розміри їх аудиторій. Такі моделі в свою чергу можуть бути використані для розуміння властивостей та ідентифікації структур мережі що зазнали цензурування.

Для перевірки гіпотези було використано стратегію побудови еталонної вибірки як критерій ефективності: спочатку використовую публічні дані (публікації та послідовники лідерів думок) соціальної мережі *Twitter* для якої можливо обчислити подібність публікацій та перетин аудиторій, далі використовую лише публікації з соціальної мережі *Facebook* для перехресної перевірки.

3.1.1 Набір даних для побудови моделей соціальних мереж

Для проведення комплексного дослідження було виділено чотири теми, що принципово відрізняються між собою. Кожна тема характеризується власними ключовими словами, виразами, інтересами та публікаціями розділяючи користувачів соціальних мереж на окремі аудиторії:

1. T_1 . Вибори в США. Публікації про вибори кандидата від Демократичної партії на Президентських виборах США в 2020 році. Лідерами думок виступають найкращі номінанти, що активно користуються соціальними мережами.
2. T_2 . Авіаційні ентузіасти. Публікації пов'язані з авіа-індустрією та бізнес-подорожами. До лідерів думок відносяться рецензенти програм постійних пасажирів та програм лояльності, журналісти які описують власний досвід подорожей, мандрівники та спеціалісти в авіаційній галузі.
3. T_3 . Глобальні проблеми. Публікації стосовно звітування, аналізу та думок про глобальні екологічні, кліматичні, та енергетичні проблеми. До лідерів думок відносяться впливові журнали, так як *The Economist*, а також обрані журналісти та організації, такі як *The World Economic Forum*.
4. T_4 . Інвестиції в технології. Публікації про актуальні інвестиції в технологічні компанії, новітні технології та підприємництво. До лідерів думок відносяться інвестиційні фонди, приватні інвестори та спеціалізовані ресурси новин.

Лідери думок та їхні публікації є вхідними даними для алгоритмів оцінки подібності публікацій та побудови мережі перетину аудиторій.

Вибір лідерів думок базується на їх актуальності та зосередженості на темі, а також за їх здатністю до залучення послідовників до обговорення. Табл. 3.1 описує обраних лідерів думок в алфавітному порядку за назвою організації чи прізвищем. Кожна з тем є великою за розмірами, та в загальному налічує сотні лідерів думок, проте для дослідження обрано лиш найбільш впливових.

Табл. 3.1. Набір даних: теми та лідери думок.

Тема	Назва	Лідери думок
T_1	Вибори в США	<i>Joe Biden, Mike Bloomberg, Pete Buttigieg, Kamala Harris, Amy Klobuchar, Bernie Sanders, Elizabeth Warren</i>
T_2	Авіаційні ентузіасти	<i>Boarding Area, French Painter, One Mile at a Time, Runway Girl, Secret Flying, Wander Me</i>
T_3	Глобальні проблеми	<i>Inside Climate, Ro Khanna, Sierra Club, The Economist, World Bank, Yale E360, Chris J. Zullo</i>
T_4	Інвестиції в технології	<i>Adam Scrabble, Epsilon Theory, Eric R. Weinstein, Fast Company, Naval, Nick Timiraos, WIRED</i>

Для кожної теми, було використано набір інструментів *Twitter API* [54] для завантаження всіх актуальних публікацій за перші 6 місяців 2020 року, включно з публікаціями обраних лідерів думок, а також списки всіх підписників лідерів думок.

3.1.2 Алгоритми побудови моделей соціальних мереж

Побудова моделі соціальної мережі

Базовим представленням соціальної мережі є складна мережа $G(N, A)$. В якій множина вершин N описує набір сутностей соціальної мережі, a множина зв'язків A описує типи відносин між ними. Зв'язки між сутностями A та B в мережі де також присутня сутність C , можуть базуватися на наступних відносинах:

- A є другом B .
- A є послідовником B .
- A взаємодіє з публікаціями B .
- A опублікувала публікацію B .
- C одночасно є послідовником A та B .

В залежності від типу відносин між сутностями, мережа стає або орієнтованою або неорієнтованою, зваженою або незваженою. Розглянемо наступні приклади:

- Мережа друзів. Мережа такого типу є неорієнтованою та не зваженою, оскільки зв'язок дружби є симетричним. Якщо сутність A є другом B , це означає, що B також є другом A . Крім того, в такій мережі зв'язки не мають жодної значущості і відповідно ваги окрім самого відношення дружби.
- Мережа взаємодії. На відміну від мережі друзів, мережа взаємодії є орієнтованою та зваженою. Це означає, що зв'язки мають конкретний напрямок, вказуючи на напрям взаємодії між сутностями. Наприклад, якщо сутність A взаємодіє з сутністю B , коментуючи її публікації, зв'язок буде спрямованим від A до B . Крім того, зв'язки в такій мережі можуть мати ваги для представлення інтенсивності або частоти взаємодій між сутностями.

Перетин аудиторій в соціальних мережах

Для побудови мережі перетину аудиторій було використано базовий підхід [55] застосований для ідентифікації аудиторій новинних ресурсів, що базується на підрахунку кількості спільних послідовників.

Розглянемо Рис. 3.1., на якому зображені лідери думок A та B , та вісім їх послідовників. Визначаємо значення перетину аудиторій $O(A, B)$ між A та B як:

$$O(A, B) = \frac{|F_A \cap F_B|}{|F_A|} \quad (3.1)$$

де F_A та F_B позначає множину підписників лідерів думок A та B . У даному випадку, перетин аудиторій $O(A, B) = \frac{3}{6}$ and $O(B, A) = \frac{3}{5}$.

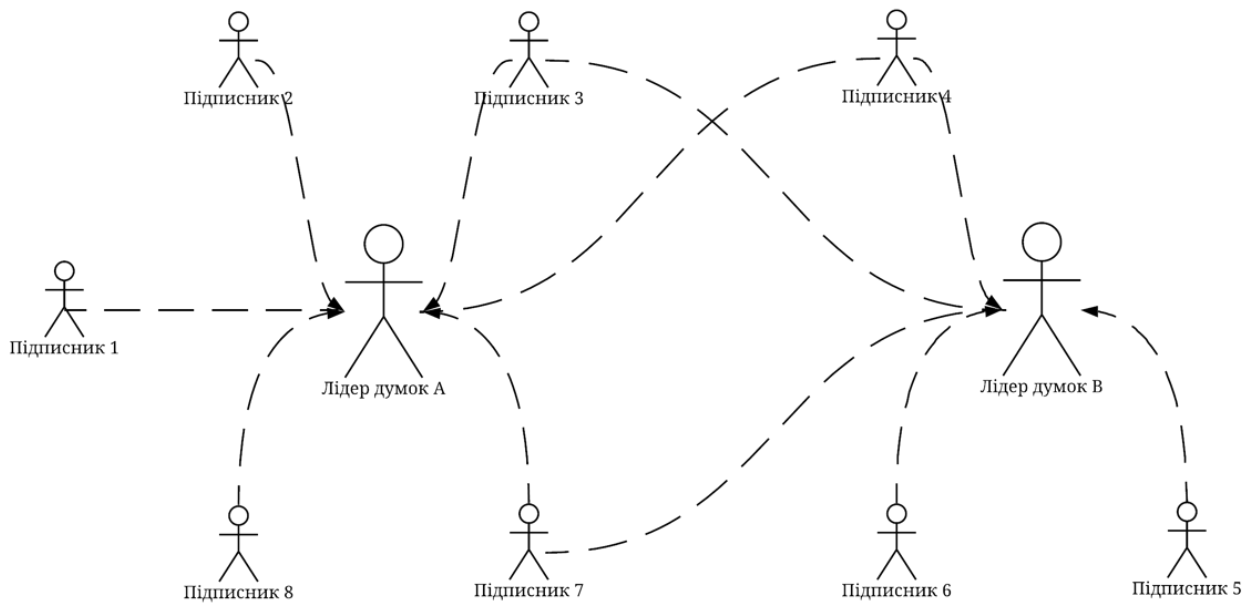


Рис. 3.1. Приклад перетину аудиторій лідерів думок A та B .

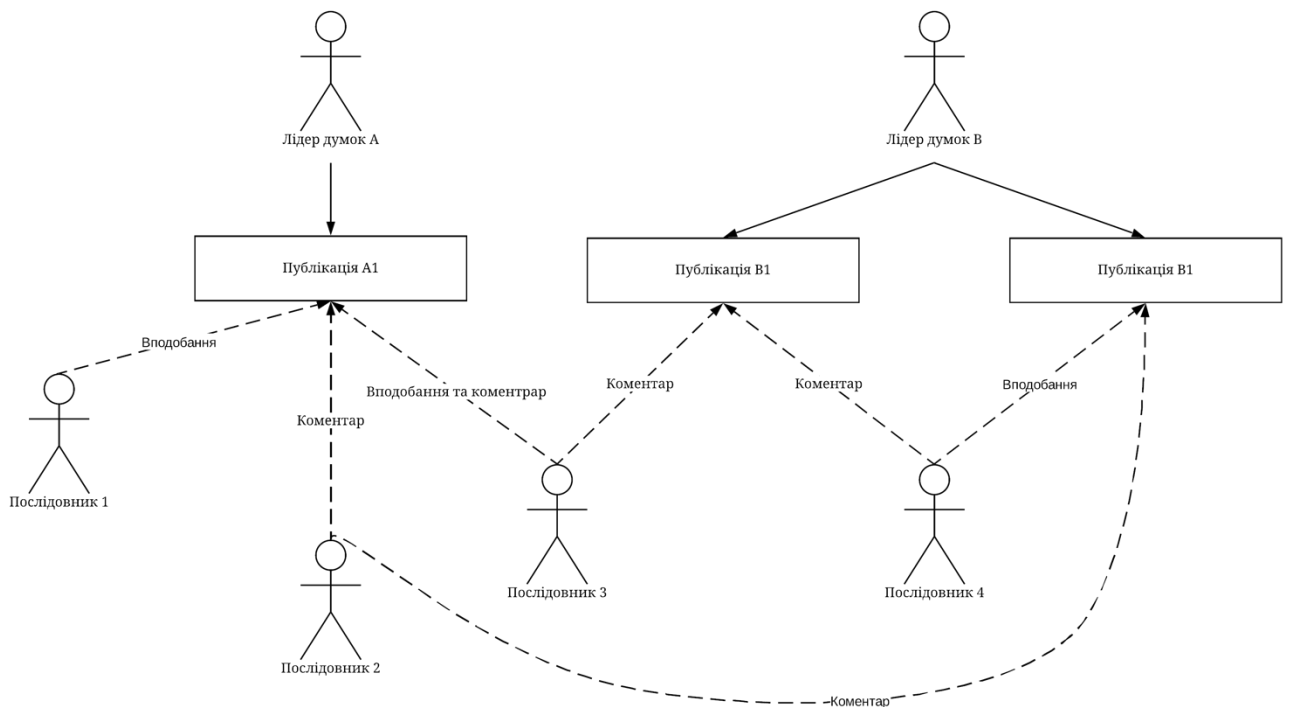


Рис. 3.2. Приклад взаємодії підписників з публікаціями лідерів думок.

В умовах обмеженості даних про послідовників в соціальних мережах, при побудові мережі перетину аудиторій можна використовувати мережу взаємодій послідовників, див Рис. 3.2. На основі публікацій лідерів думок, відбувається ідентифікація послідовників які залишили коментарі, вподобання або поділилися такими публікаціями. Однак, доступ до таких даних все частіше обмежується соціальними мережами.

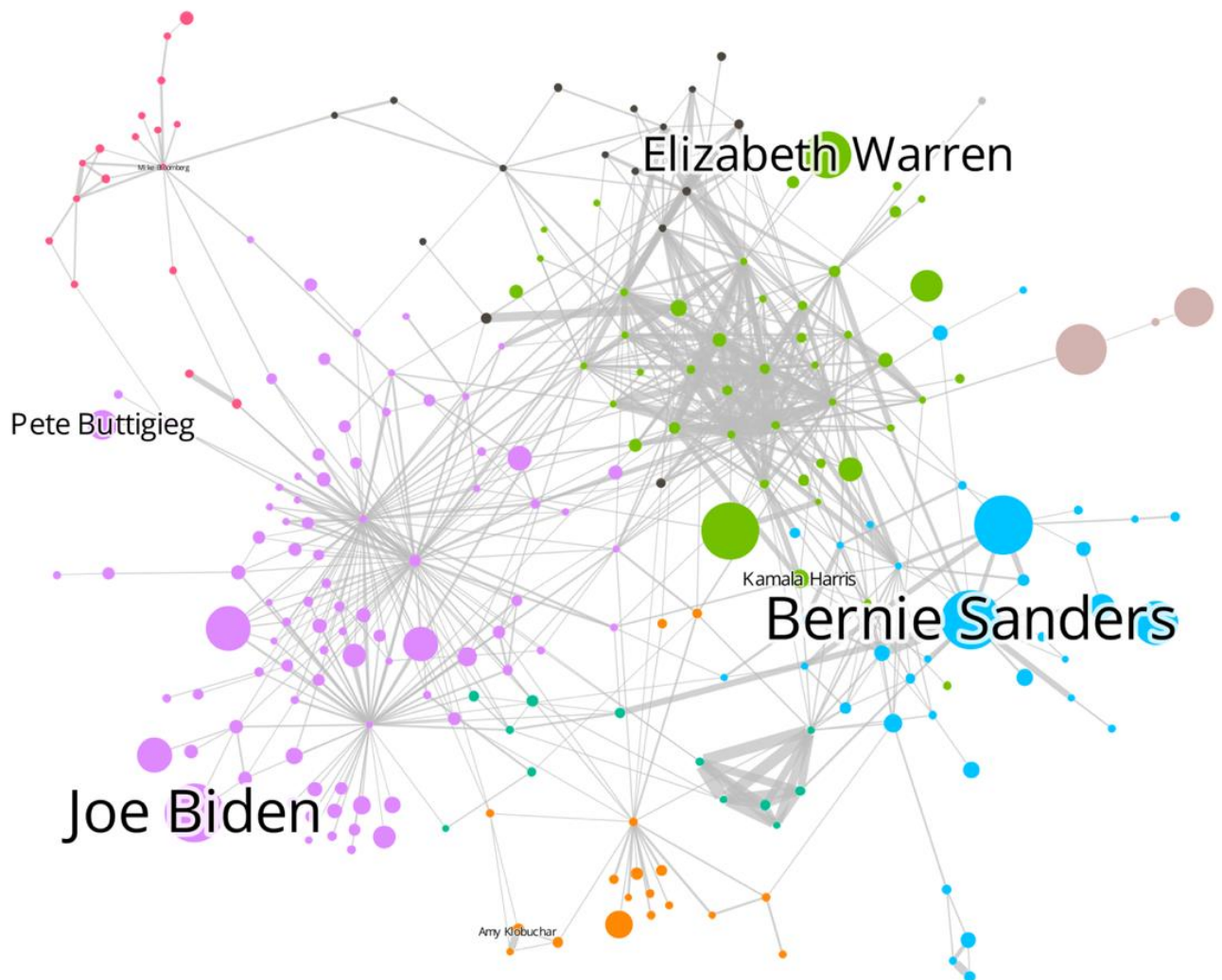


Рис. 3.3. Мережа перетину аудиторій всіх лідерів думок теми T_1 за перші 6 місяців 2020 року в соціальній мережі Twitter.

Подібність публікацій як індикатор перетину аудиторій

З активним розвитком машинного навчання, методи обробки природніх мов зазнали значної трансформації. Вони замінили неефективні методи, що

дозволяють порівнювати тексти лише на одній мові (*Bag-of-words* [56], *TF-IDF* [57]), на більш ефективні методи порівняння. Такі методи часто базуються на архітектурі *Transformers* [58] та зокрема дозволяють порівнювати тексти на довільних мовах. Для оцінювання семантичної подібності двох публікацій в *Twitter* було використано модель *Universal Sentence Encoder (USE)*. [59]. *USE* кодує текст у вигляді 512-вимірних векторів. Такі вектори надають простий спосіб вимірювання відстані між двома будь-якими публікаціями $d(\cdot, \cdot)$. На Рис. 3.4 зображено візуалізацію всіх публікацій лідерів думок теми T_1 після опрацювання з використанням моделі *USE*. Одним з недоліків моделі *USE*, є обмеження допустимих мов: серед яких арабська, спрощена китайська, традиційна китайська, англійська, французька, німецька, італійська, японська, корейська, нідерландська, польська, португальська, іспанська, тайська, турецька та російська.

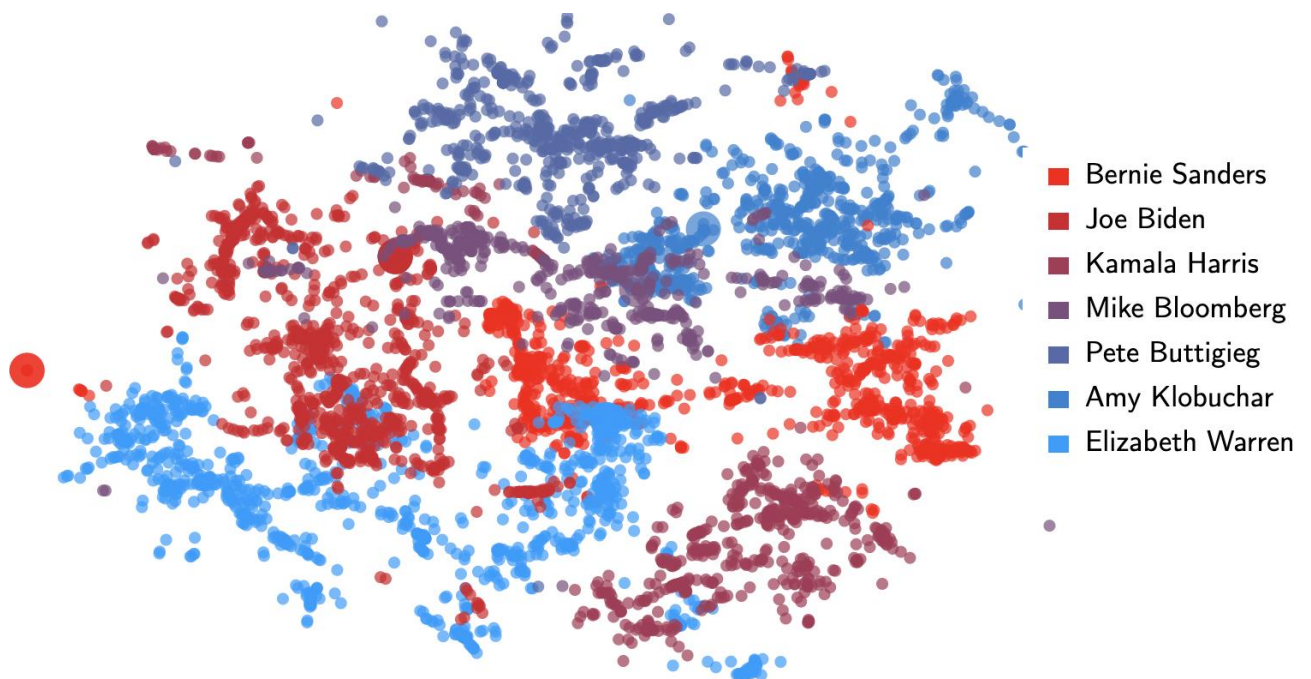


Рис. 3.4. Розподіл публікацій лідерів думок теми T_1 на основі *USE*. Для відображення векторів двовимірному просторі, було використано алгоритм зниження розмірності *t-SNE* [60]. Лідери думок як правило, публікують в своєму унікальному стилі окрім *Mike Bloomberg* який імітує всіх лідерів думок одночасно.

Для обрахунку подібності публікацій двох лідерів думок, позначимо $T = \{T_A, T_B, \dots\}$ як множину всіх публікацій лідерів думок, де T_A непорожня множина публікацій A . Для будь-якого $k = 1, 2, 3, \dots$ та ϵ , нехай $N_{k,\epsilon}(t, X)$ множина k -найближчих публікацій публікації t , обраної з множини X таким чином, що відстань $d(t, x) < \epsilon$ де $x \in X$. Тоді, подібність публікацій для лідерів думок A та B обраховуємо як:

$$C(A, B, T)_{k,\epsilon} = \sum_{a \in T_A} \frac{|N_{k,\epsilon}(a, T) \cap T_B|}{k|T_A|} \quad (3.2)$$

Важливо зазначити, що значення $C(A, A, T)_{k,\epsilon}$, визначає подібність власних публікацій лідера думок, та є характеристикою, що вказує на скільки різноманітними є його публікації.

3.1.3 Калібрування моделі перетину аудиторій

Кожна тема відрізняється специфікою термінів вжитих у публікаціях та загальною стилістикою написання. Тому необхідно на рівні кожної теми T означити два параметри k and ϵ для мінімізації абсолютної похибки між подібністю публікацій та перетином аудиторій лідерів думок:

$$(\hat{k}, \hat{\epsilon}) = \underset{k,\epsilon}{\operatorname{argmin}} \sum_{A \in I} \sum_{B \in I} |C(A, B, T)_{k,\epsilon} - O(A, B)| \quad (3.3)$$

де $k, \epsilon = 1, 2, 3, \dots, 25$.

На Рис. 3.5 зображено всі отримані значення похибки для теми T_l , в той же час в Табл. 3.2 наведено оптимальні значення для параметрів k та ϵ для всіх тем. Було встановлено, що значення $\hat{k} = 1$ є універсальним для всіх тем, в той самий час значення параметру ϵ є принципово важливим, та залежить від конкретної

теми. Такий результат, спричинений в першу чергу публікаціями лідерів думок: чим більш вузькоспеціалізована тема, тим більш схожі лідери думок, і відповідно менше значення k .

Табл. 3.2. Оптимальні параметри та показники похибки для кожної теми.

Тема	\dot{k}	$\dot{\epsilon}$	Абсолютна похибка	Коефіцієнт Спірмена ρ
T_1	1	0.92	14.08	0.57
T_2	1	1.50	22.94	0.80
T_3	1	0.20	15.50	0.71
T_4	1	1.02	12.75	0.83

В Табл. 3.2 наведено значення оптимальних параметрів \dot{k} та $\dot{\epsilon}$ для кожної теми. Калібрування параметрів відбувалася для мінімізації похибки, в Табл. 3.2 також наведено значення коефіцієнту кореляції рангу Спірмена [61]. Всі значення коефіцієнту кореляції є значущими на рівні 0.01 .

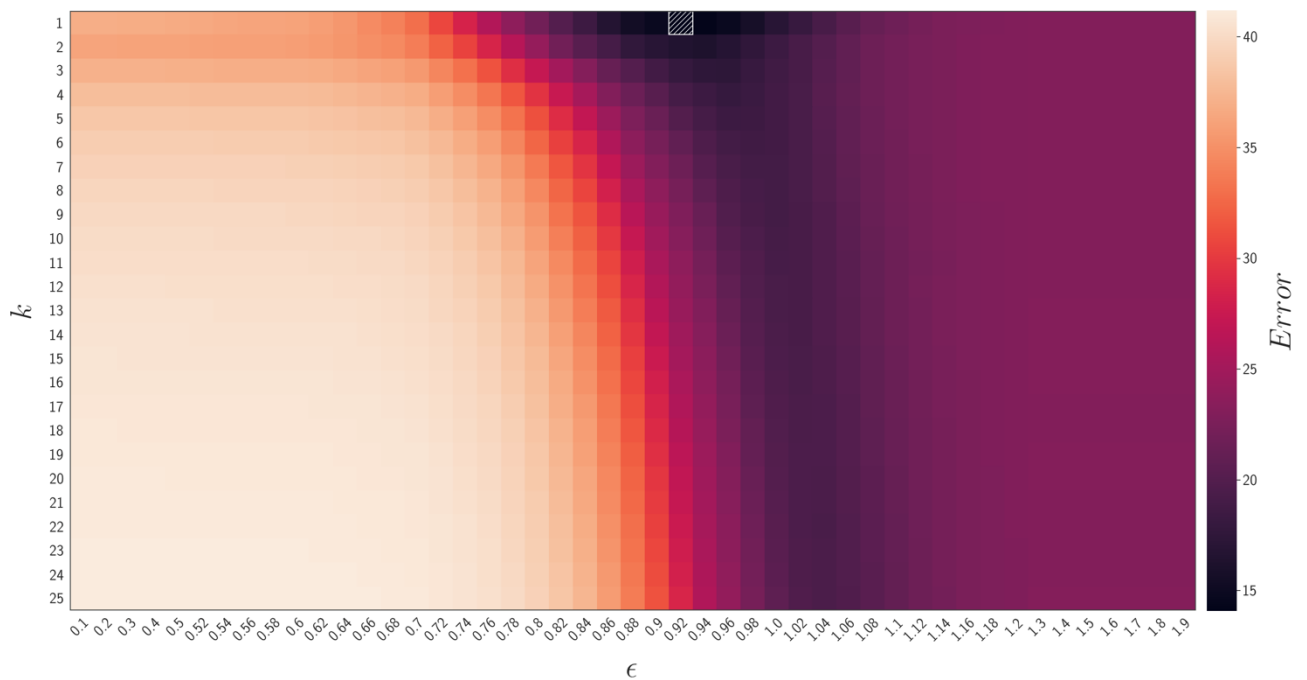


Рис. 3.5. Значення похибки для теми T_1 , з оптимальними параметрами $\dot{k} = 1$ and $\dot{\epsilon} = 0.92$.

3.1.4 Аналіз отриманих результатів

Табл. 3.3 наведено калібровані мережі подібності публікацій та перетину аудиторій для кожної теми, відображаючи значення параметрів та помилок наведених в Табл. 3.2. Мережі відображені у вигляді матриць, де кожне значення для рядка X та стовбця Y відповідає значенню коефіцієнтів $O(A, B)$ та $C(A, B, T)$.

Візуально, подібність публікацій відповідає перетину аудиторія для тем T_1 , T_3 та T_4 . В той же час, калібрування мереж для теми T_2 не дало бажаних результатів, і для неї мережі не співпадають. Такий результат, спричинений в першу чергу відсутністю достатньої кількості публікацій для теми Авіаційних ентузіастів. Незважаючи на це, відсутність тісних та явно виражених зв'язків в цій темі, також відображена в подібності публікацій.

Початкове припущення про тісний зв'язок між подібністю публікацій та перетином аудиторій в соціальних було кількісного підтверджено з використанням чотирьох теми. Побудовані мережі подібності публікацій візуально характеризуюся мережевою структурою мереж перетину аудиторій.

Хоча цей підхід є багатообіцяючим, побудова мереж перетину аудиторій не може покладатися лише на подібність публічно доступних даних. Найбільш ефективним застосуванням мереж подібності публікацій є їх використання як допоміжного фактору при побудові мереж перетину аудиторій. В комбінації з іншими характеристиками лідерів думок, а також інтеграцією мереж взаємодії створює необхідний набір інструментів для комплексного аналізу розповсюдження впливу в соціальних мережах. Важливим аспектом також залишається якісь результатів при еволюції та рості мережі, оскільки теми які виникають в соціальних мережах є тимчасовими. Це, у свою чергу, підкреслює необхідність оптимізації активного збору даних.

Табл. 3.3. Візуальне порівняння моделей мереж для тем T_1, T_2, T_3, T_4

Тема	Мережа подібності публікацій	Мережа перетину аудиторій
T_1		
T_2		
T_3		
T_4		

3.2 ДОСЛІДЖЕННЯ ПЕРЕТИНУ АУДИТОРІЙ В СОЦІАЛЬНІЙ МЕРЕЖІ TELEGRAM

Одним із аспектів міжнародних санкції щодо Росії, стала консолідація зусиль західних соціальних мереж в контексті обмеженні доступу для користувачів з Росії. Зокрема, соціальні мережі *Facebook*, *Twitter* та *YouTube*, заборонили рекламу, активно видаляли пропагандистські профілі лідерів думок та практично заборонили доступ до всіх послуг звичайним користувачам.

Багато російських пропагандистів, знаючи про можливі ризики залишитися без доступу до власних аудиторій, шукали альтернативні соціальні мережі з меншою кількістю модерації ще до 24 лютого 2022 року. Враховуючи популярність *Telegram* як засобу повсякденного зв'язку росіян, він став для пропагандистів очевидним вибором. З іншого боку, уряд Росії ще більше прискорив процес міграції користувачів в *Telegram* шляхом заборони західних соціальних мереж, та наказуючи всім лідерам думок мотивувати користувачів до використання *Telegram* для всіх можливих комунікацій.

В той час, коли публікації та сторінки в соціальних мережах *Facebook* та *YouTube* просуваються автоматично на основі попередніх інтересів користувача, в *Telegram* користувачі повинні власними силами шукати інформацію. Тому, незважаючи на численні переваги, в *Telegram* відсутні ефективні механізми пошуку нових публікацій та каналів. Функція пошуку в *Telegram* працює за принципом точного співпадіння пошукової фрази та назви каналу, уникаючи пошуку за публікаціями. В той же час, механізм реклами в *Telegram* знаходився на стадії тестування.

Для сприяння швидкому зростанню аудиторій державних пропагандистських каналів, росіяни використали декілька тактик. Найбільш видимою тактикою, стало обрання існуючих лідерів думок в *Telegram*, серед яких ведучий державних каналів Соловйов [62] та інші, та використання них для заміни відсутньої системи пошуку.

Telegram канал Соловйова активно поширювався та рекламувався традиційними засобами масової інформації, телевізійними каналами які транслювали його авторські програми, для максимізації притоку нових та існуючих користувачів у вигляді нових послідовників. Його канал регулярно публікував тематичні огляди на менші канали, які тільки розвивалися, підтримував їх, просував та рекламував їхні публікації. Таким чином його канал перетворився на найбільше джерело інформації для росіян, стаючи відсутнім механізмом генерації рекомендацій. Це призвело до появи величезного апарату внутрішньої пропаганди на базі соціальної мережі *Telegram*.

В цьому розділі було досліджено зростання аудиторій пропагандистських спільнот та каналів, що підтримують Російське вторгнення в Україну, та кількісно оцінюю вплив централізованого просування лідерами думок на розвиток менших спеціалізованих каналів в *Telegram*.

Російські соціальні мережі стали доволі різноманітними впродовж останніх 5 років. *Vkontakte*, *odnoklassniki*, *YouTube*, *Instagram* та *Facebook* активно використовувалися російською аудиторією. В той же час, *Telegram* мав значно повільніше, проте стабільне зростання, і став першим вибором для особистого спілкування, але не використовувався як повноцінна соціальна мережа. Тому російські лідери думок почали створювати канали виключно як резервні, або ж альтернативні способи розповсюдження інформації для молодших аудиторій.

З самого початку існування *Telegram* створювався як платформа з мінімальною кількістю модерації в порівнянні з іншими внутрішніми соціальними мережами, такими як *Vkontakte* чи *odnoklassniki*. Тому російський уряд неодноразово намагався заблокувати доступ до нього на території країни [63]. Проте, після 24 лютого 2022 року, після запровадження західних санкцій, уряд Росії усвідомив потребу в соціальній мережі, що толеруватиме пропаганду, тому *Telegram* став для них максимально корисним. Всього за 2 тижні російська аудиторія повністю мігрувала в *Telegram*, див Рис. 3.6.

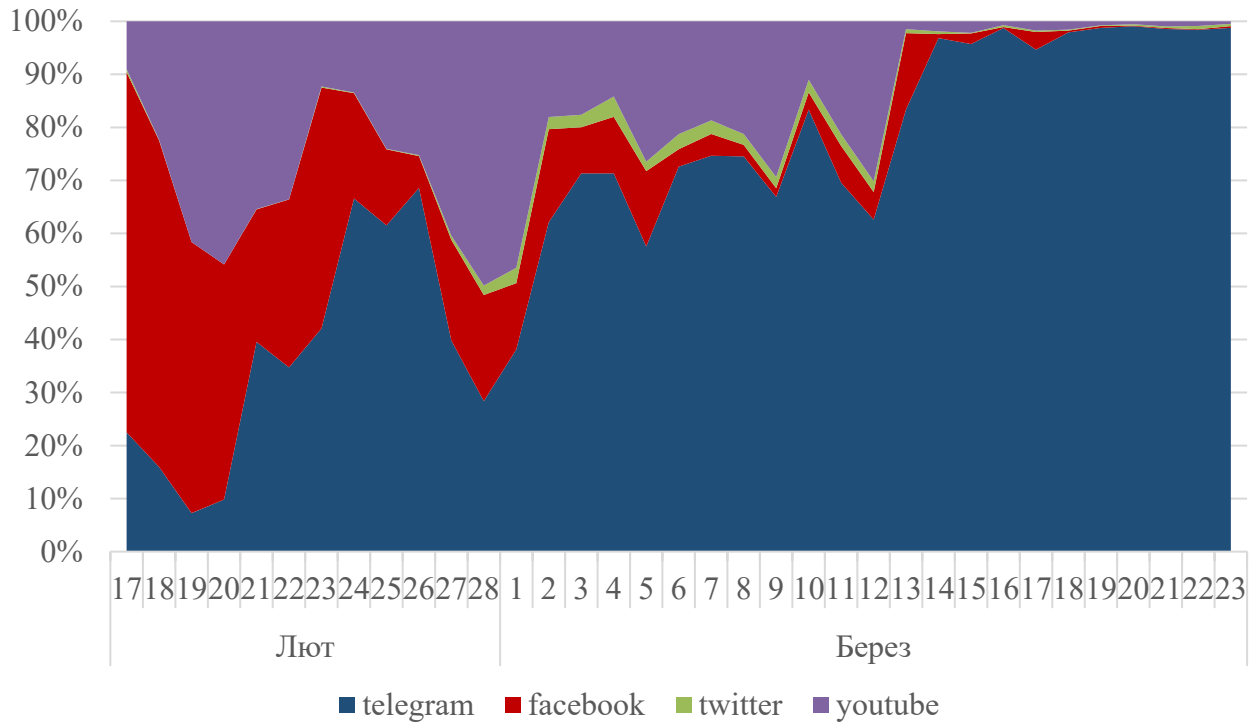


Рис. 3.6. Відсоток щоденних публікацій, присвячених подіям в Україні та війні у російських соціальних мережах.

3.2.1 Набір даних для побудови мережі перетину аудиторій

Для проведення дослідження, по-перше, було побудовано набір даних соціальної мережі *Telegram*, що складається зі всіх публікацій каналу *solovievlive* (36911 публікацій починаючи з 1 лютого 2022 року, включаючи 23548 публікації інших каналів що були ним переслані), див Рис. 3.7. По друге, зібрано всі публікації *Telegram* каналів які були ним рекламовані (3678480 публікацій від 1048 каналів) включаючи дату публікації, кількість переглядів, коментарів та реакцій. Зрештою, було зібрано російськомовні публікації в такому ж самому часовому проміжку з каналів що характеризуються інтересами в економіці, війні та політиці, які не були переслані каналом *solovievlive* (8346057 публікацій з 119916 каналів).

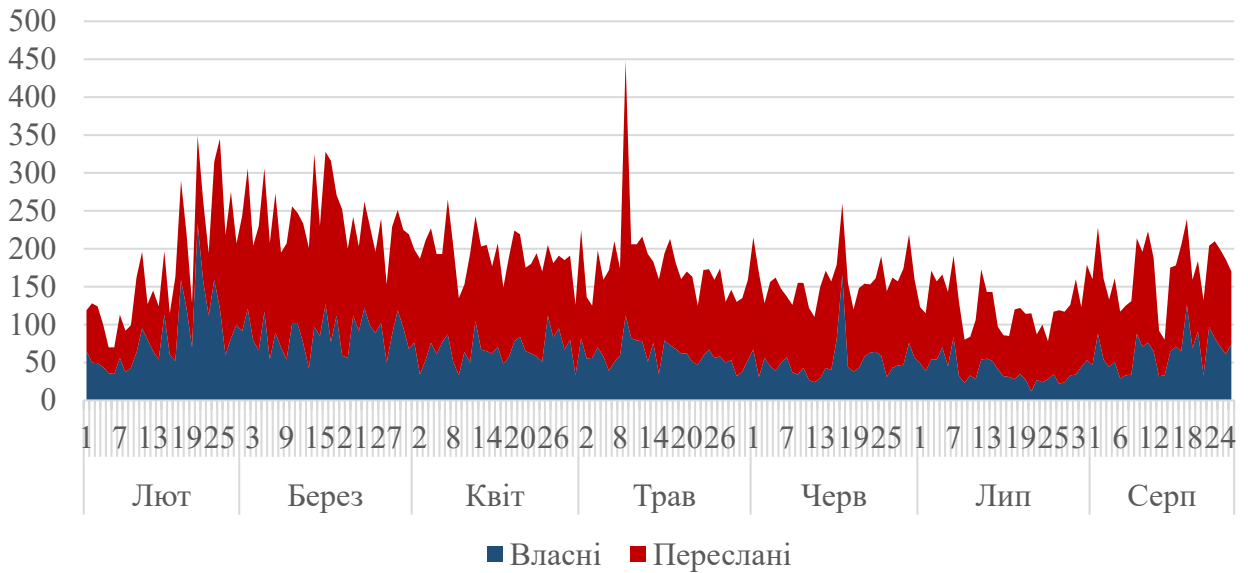


Рис. 3.7. Щоденна кількість власних та пересланих публікацій на *Telegram* каналі *solovievlive*.

Канал *solovievlive* оперує за наперед визначеним графіком. Він добре розуміє множину “перевіраних”, з якими він координує діяльність, систематично пересилаючи їх публікації у випадку появи резонансних новин. В інших випадках, канал фокусується на пересиланні повідомлень від нових каналів, див Рис. 3.8.

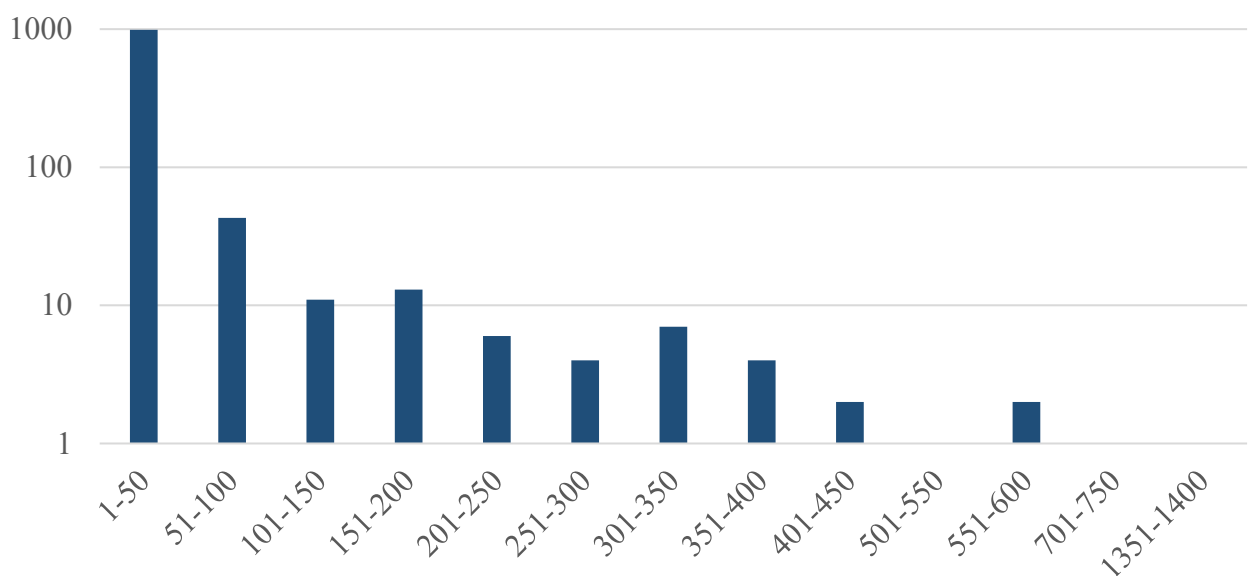


Рис. 3.8. Кількість *Telegram* каналів пересланих *solovievlive* за кількістю публікацій.

В Табл. 3.4 наведено список каналів, публікації яких *solovievlive* були найчастіше переслані. Також, ці данні перетворено та візуалізовано у вигляді мережі пересилань, див. Рис. 3.11.

Табл. 3.4. Канали що найчастіше пересилалися каналом *solovievlive*.

Канал	Місяць							
	Лют	Берез	Квіт	Трав	Черв	Лип	Серп	Заг
dimsmirnov175	67	102	51	74	145	86	96	621
imerkouri	58	99	53	48	49	28	20	355
mariavladimirovnazakharova	53	49	38	30	35	56	42	303
sergeykolyasnikov	12	87	58	42	33	14	21	267
rybar	11	21	26	53	54	42	49	256
rkadyrov_95	2	105	40	33	23	26	18	247
romagolovanov	40	89	50	18	11	19	10	237

3.2.2 Дослідження динаміки еволюції мережі

Для дослідження ефективності дій каналу *solovievlive*, для кожної пересланої публікації, було розраховано середню кількість переглядів для публікацій каналу-автора за 5 днів до та за 5 днів після того як *solovievlive* її переслав. Після цього було розраховано відношення між двома середніми значеннями як функцію від місяця та середнього значення переглядів, до пересилання, див Табл. 3.5.

Варто зазначити, що для каналів які були переслані виконується:

- Темпи зростання у лютому та березні були універсально високими, досягаючи 70%, проте згодом почали знижуватися.
- Малі канали продовжували збільшувати свої аудиторії та перегляди, щонайменше на 25% кожні 10 днів.

- Темпи зростання середніх та великих каналів припинився, та навіть став негативним.

Табл. 3.5. Вплив каналу *solovivlive* на середню кількість переглядів у відсотках.

Сер. перегляди	Month						
	Лют	Берез	Квіт	Трав	Черв	Лип	Серп
< 20000	171%	138%	154%	128%	133%	125%	165%
20000-60000	161%	119%	109%	105%	110%	106%	112%
60000-180000	155%	111%	105%	95%	104%	99%	100%
> 180000	174%	107%	107%	98%	95%	101%	89%

Дані з Табл. 3.5 не можуть бути використані для підтвердження залежності динаміки зростання від підтримки каналу *solovivlive*, оскільки вони можуть бути спричинені іншими внутрішніми процесами в *Telegram*. Для підтвердження залежності, було використано випадкову вибірку публікацій, що ніколи не підтримувалися *solovivlive*, та розрахував 10-ти денні темпи зростання використовуючи схожу процедуру. Темпи зростання для випадкової вибірки наведені в Табл. 3.6, а різниця між темпами зростання випадкової вибірки та публікацій що були підтримані *solovivlive* та іншими каналами наведена в Табл. 3.7.

Табл. 3.6. Темпи зростання переглядів для каналів випадкової вибірки у відсотках.

Сер. перегляди	Місяць						
	Лют	Берез	Квіт	Трав	Черв	Лип	Серп
< 20000	164%	107%	111%	104%	111%	113%	110%
20000-60000	143%	101%	100%	97%	100%	101%	90%
60000-180000	137%	99%	97%	97%	96%	98%	89%
> 180000	127%	98%	98%	97%	97%	98%	87%

Табл. 3.7. Різниця між темпами зростання переглядів для каналів що підтримувалися *solovievlive* та випадкової вибірки каналів.

Сер. перегляди	Місяць						
	Лют	Берез	Квіт	Трав	Черв	Лип	Серп
< 20000	6%	31%	43%	23%	21%	13%	56%
20000-60000	17%	18%	10%	8%	11%	5%	21%
60000-180000	18%	12%	8%	-2%	8%	1%	11%
> 180000	48%	9%	10%	1%	-1%	3%	3%

Виходячи з даних Табл. 3.7, канал *solovievlive* значно прискорив зростання пересланих каналів аж до травня 2022 року, в порівнянні з середньостатистичними темпами зростання каналів в російському Telegram. В травні та червні, він продовжував прискорювати темпи росту малих та середніх каналів, проте його вплив на великі канали був незначним, потенційно негативним.

Одним з пояснень відсутності значного впливу на зростання великих каналів є поведінковий аспект. Можна припустити, що вже у травні більшість користувачів підписали на великі канали, та вимкнули сповіщення на їхні публікації. В момент, коли *solovievlive* пересилає допис з великого каналу для якого вимкнуті сповіщення, користувач не знатиме про таку публікацію. Зазвичай, *solovievlive* пересилає відносно нові публікації, тому навіть якщо користувач побачить таку публікацію, під час її відкриття, він буде переправлений до її контексту в каналі автора, яка буде найновішою. Тому тільки у випадку коли після відкриття такої публікації користувач вирішить переглянути інші публікації каналу, він отримає перегляди. Цей процес призводить до появи негативних темпів зростання.

Повернення до збільшення темпів зростання в липні та серпні може бути показником наступних факторів:

- Користувачі відновлюють інтерес до пропагандистських каналів для яких були вимкненні сповіщення.

- Існує хвиля новостворених авторських та каналів середнього розміру, які мають інший погляд на поточну ситуацію, в яких зацікавлений *solovievlive* див, Табл. 3.8.

Табл. 3.8. Нові канали, що пересилаються *solovievlive*

Канал	Місяць							
	Лют	Берез	Квіт	Трав	Черв	Лип	Серп	Заг
Mediastanciacom	0	0	0	0	15	15	5	35
Mkhusnullin	0	0	0	6	9	5	8	28
turchak_andrey	0	0	0	3	7	6	3	19
Bankingtruth	0	0	0	0	2	9	7	18
atas_atas	0	0	0	0	1	11	6	18

3.2.3 Побудова мережі перетину аудиторій та пересилань

Мережа перетину аудиторій

Для дослідження еволюції мережі, розповсюдження впливу та зміну ставлення до поточних новин, ефективно застосувати мережу перетину аудиторій. Зважаючи на те, що списки послідовників в *Telegram* є прихованими, було використано комбінований підхід до побудови мережі перетину аудиторій. Комбінуючи мережу подібності публікацій, що була побудована на основі даних про публікації каналів, а також мережу взаємодій, побудовану на інформації про користувачів що залишили коментарі, було отримано мережу перетину аудиторій.

5 травня 2022 року, 85% користувачів *Telegram* що є послідовниками російських каналів максимально підтримували про-російську позицію стосовно поточної ситуації. Серед тих, хто підтримував військові дії, було висловлено

позицію про необхідність ескалації до максимальної інтенсивності з повною мобілізацією, див Рис. 3.9.

Проте вже 15 травня 2022, публікації, що закликали до повної мобілізації повністю зникли, зменшуючи загальну кількість тих, хто підтримує поточний підхід до ведення операції. Важливо зазначити, що група провокативних користувачів та каналів, об'єдналася для висміювання різноманітних анти-воєнних ліберальних каналів таких як *Nexta*, див Рис. 3.10.

Таким чином, за допомогою мережі перетину аудиторій можна спостерігати за динамікою процесів в соціальних мережах та ідентифікувати лідерів думок як основного джерела розповсюдження впливу.

Табл. 3.9. Класифікації каналів на основі ставленням до поточної ситуації.





Колір	Ставлення	5 травня	15 травня
	“Зупиніть війну”	15%	19%
	“Ми перемагаємо.”	15%	77%
	“Час повної мобілізації.”	69%	<1%
	Провокатори	< 1%	4%



Рис. 3.9. Мережа перетину аудиторій *Telegram* каналів з 25 квітня по 5 травня.

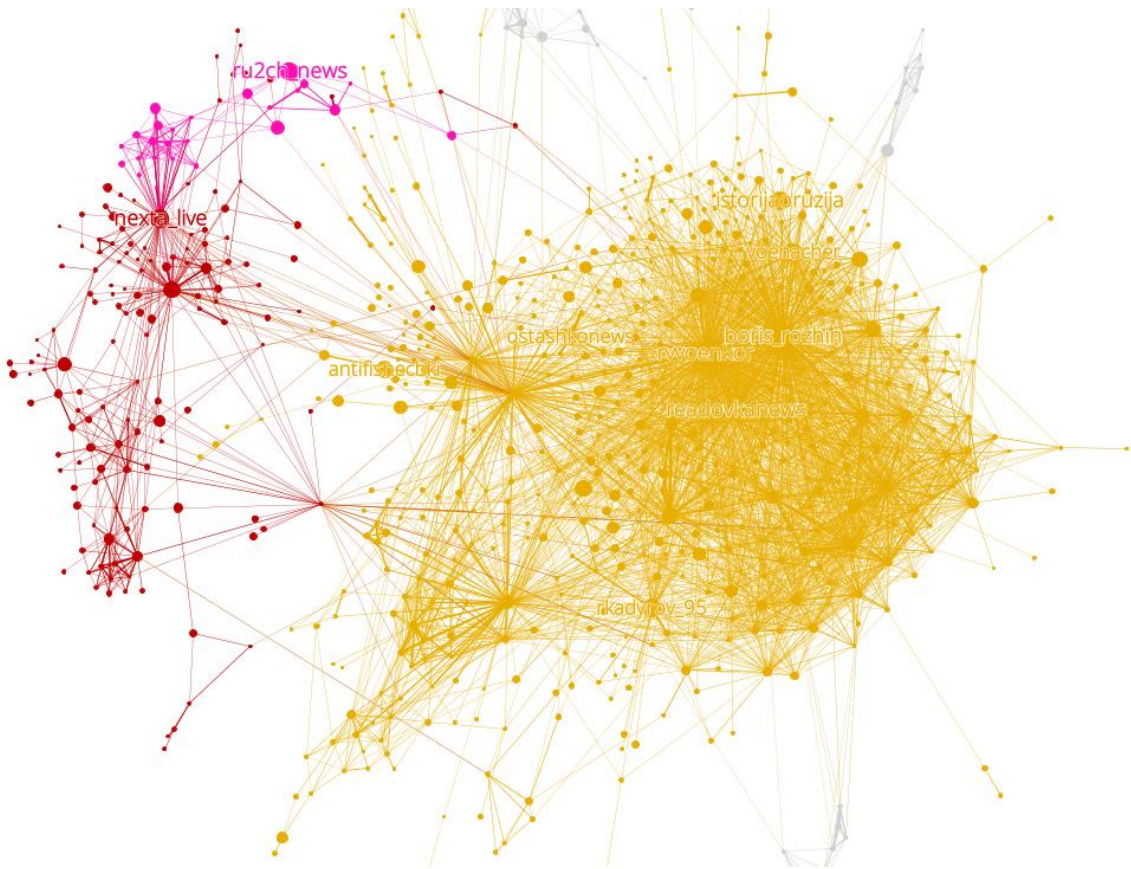










Рис. 3.10. Мережа перетину аудиторій *Telegram* каналів з 5 до 15 травня.

Мережа пересилань публікацій

Аналізуючи важливість ролі *solovivlive* в контексті розповсюдження пропаганди, доцільно проаналізувати структуру пересилань всіх каналів у вигляді мережі пересилань публікацій, що характеризуються інтересами в економіці, війні та політиці, див Рис. 3.11. Для розуміння структури мережі було застосовано алгоритм ідентифікації спільнот [64].

Відповідно до побудованої мережі, російська пропаганда оркеструється через канал *solovivlive* як основу до розповсюдження інформації. Проте, після ідентифікації спільнот, можна легко помітити наявність в кожній з них каналів які відповідають за зв'язок всередині спільноти та комунікацію з іншими спільнотами. В таких умовах, достатнього всього 6 каналів для того щоб з'єднати канал *solovivlive* та офіційний канал президента України.

Табл. 3.10. Характеристика спільнот в *Telegram*.

Колір	Спільнота
	Основна внутрішня російська пропаганда
	Російські ліберальні канали
	Російські пропагандистські канали-посередники
	Російські канали, які намагаються виглядати незалежними
	Канали ДНР та ЛНР
	Офіційні російські урядові канали
	Білоруські пропагандистські канали.
	Офіційні українські канали

поведінкою користувачів протидіють створенню централізованого механізму впливу який базується на активності кількох великих каналів.

Користувачі природньо схиляються до невеликих або середніх за розміром авторських каналів з помірною частотою публікацій.

Ймовірно, російський пропагандистський апарат усвідомлює ці динаміки та підтримує активність каналу *solovievlive* таким чином, що канали, які стали занадто великими або почали втрачати інтерес користувачів через вимкнені сповіщення або не змогли підтримувати якісь публікацій, більше не пересилалися.

3.3 ЗАСТОСУВАННЯ ПЕРЕТИНУ АУДИТОРІЙ ДЛЯ РОЗПОВСЮДЖЕННЯ ВПЛИВУ

Основна мета впливу в соціальних мережах полягає в максимізації розповсюдження інформації для формування громадської думки. Операції з розповсюдження впливу спрямовані в першу чергу на ідентифікацію цільових аудиторій, взаємодію з ними шляхом інтеграції специфічних наративів що відповідають їх поглядам в спілкуванні з ними. Це призводить до зміни сприйняття та інтерпретації інформації окремими особами з цільової аудиторії. Цільові аудиторії можуть бути описаним як набір сутностей в соціальних мережах, що мають схожі інтереси, погляди або поведінку.

Саме тому, мережі перетину аудиторій надають глибоке розуміння цих цільових аудиторій. Розглянемо Рис. 3.9 та спільноту що складається з російських пропагандистів. Розуміючи, що значна частина населення слідкує за одним і тим же набором каналів, кожен з них повинен мати свою власну унікальну інтерпретацію подій та висловлювання, дотримуючись офіційної державної позиції.

Для того, щоб порушити цю систему впливу, потрібно було б створити новий набір *Telegram* каналів, спрямованих на ті ж цільові аудиторії. Такий підхід був би успішним лише за умови використання на ранніх етапах формування мережі. Однак, в умовах коли мережа вже сформована, це вимагало б наявності також нових користувачів *Telegram*, які виступили б в ролі початкової аудиторії для нових каналів. Проте така ситуація практично неможлива, тому може розглядатися в теорії.

3.3.1 Концепція впливу в контексті користувача

Альтернативний підхід базується не на розповсюдженні впливу на рівні каналів, а на рівні вже існуючих користувачів *Telegram*. В загальному в *Telegram* існує два типи каналів: ті які дозволяють користувачам коментувати свої публікації, та ті які повністю обмежують можливість коментування. Канали, з вимкненими коментарями функціонують як новинні ресурси, які лише створюють новини, та передають їх до каналів з коментарями для перевірки ставлення користувачів. Тому канали, з коментарями навпаки заохочують обговорення та взаємодію між користувачами. У таких каналах, коментарі стають такими ж важливими як і публікації.

Користувачі в *Telegram* можуть реагувати та відповідати на коментарі інших користувачів, роблячи найбільш актуальні та запеклі коментарі легко доступними в секції коментарів. Це породжує динамічне середовище, в якому коментарі користувачів та їх взаємодія формують загальну думку обговорення.

В цьому розділі, було продемонстровано, як використання штучних агентів може вплинути на стан мережі *Telegram*, зосереджуючись на денному розповсюдженні інформації (перегляди - кількість унікальних користувачів *Telegram*, які бачили хоча б один коментар, написаний штучним агентом), та реакціях досягнутих штучними агентами розміщених по всій мережі *Telegram*.

3.3.2 Набір даних для моделювання розповсюдження впливу

Для дослідження динаміки коментарів, було зібрано набір даних, який складається з 12200 російських *Telegram* каналів, які робили публікації про “спеціальну воєнну операцію” (СВО), російську політику чи російську економіку. В Табл. 3.11 наведено денну статистику цих каналів за 12 травня 2022. Ця статистика є отримана напряму з *Telegram* та включає інформацію про кількість переглядів публікацій та кількість отриманих ними коментарів.

Важливість вибору соціальної мережі *Telegram* в період травня 2022 року пояснюється активністю користувачів соціальних мереж, які зосередили спілкування на тему СВО саме в *Telegram*. В той самий час інші соціальні мережі в міру застосування санкцій та модерації публікацій не заохочували активність користувачів, див Рис. 3.12. Запропонована в цьому розділі моделювання розповсюдження впливу зосереджено на *Telegram*, проте може бути застосоване для інших соціальних мереж.

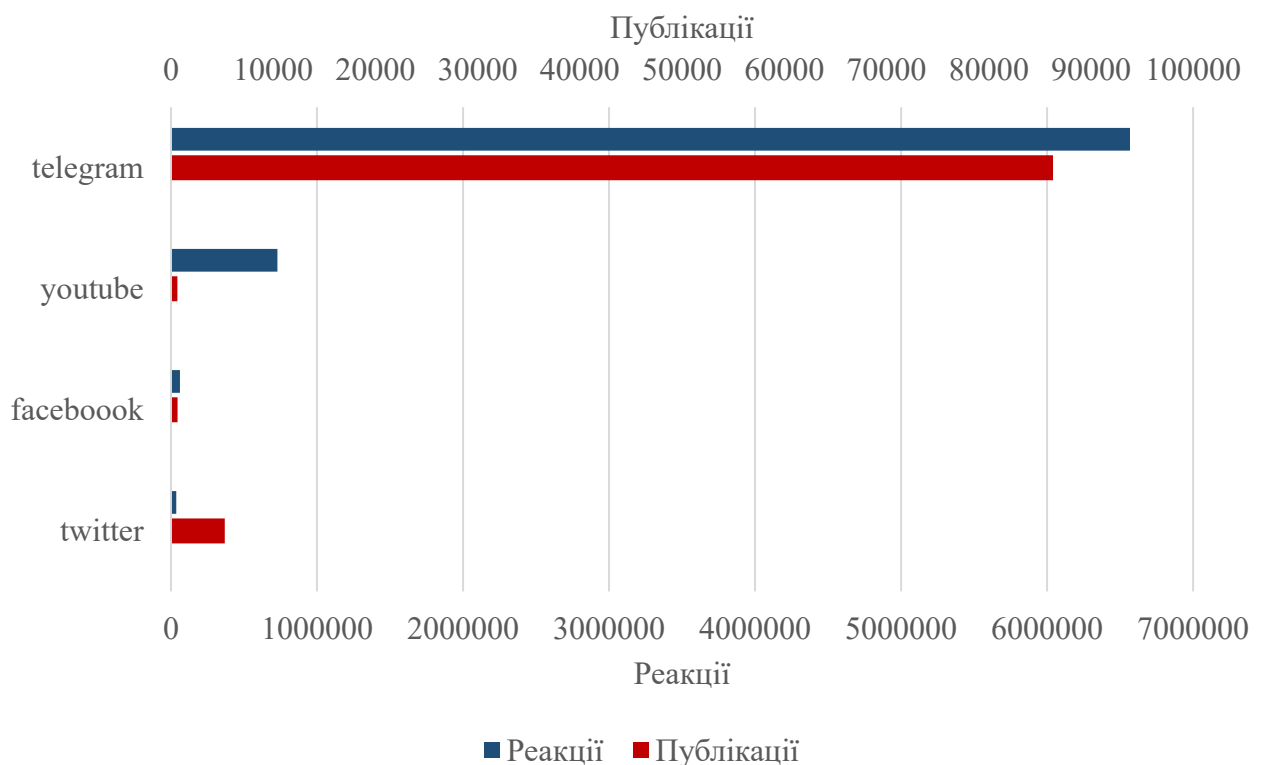


Рис. 3.12. Активність користувачів різних соціальних мереж на тему СВО 12 травня 2022 року.

Табл. 3.11. Статистика російських *Telegram* каналів.

Підписники	Канали	Наявність коментарі	Денні показники		
			Публікацій	Переглядів публікації	Коментарі публікації
< 20000	9050	59%	6	2750	14
20000 – 50000	1750	50%	8	13100	42
50000 – 100000	670	42%	10	22400	87
100000–200000	450	39%	12	43900	170
200000 – 500000	250	38%	15	108400	320
> 500000	80	33%	30	301800	1,520

Перш за все, розмір аудиторії каналу визначає властивості динаміки коментарів:

- Малі канали, як правило, але кількість користувачів, що беруть участь в обговоренні є незначущою або нульовою.
- Великі канали, як правило, спамлять, публікуючи більше ніж 30 публікацій за день.
- Великі канали також будуть обережними, та не будуть мати коментарів доступних під кожною публікацією, відповідно дозволяючи поверхневе обговорення, з максимальною його модерацією.

Тому оптимальним розміщенням агентів, будуть коментарі публікацій *Telegram* каналів, аудиторія яких знаходиться в межах від 25000 до 200000 підписників, див Табл. 3.12.

Табл. 3.12. Цільові особливості *Telegram* каналів.

Особливості каналу	Денна доступність		
	Канали	Публікації	Перегляди
Відповідна тематика: СВО, російська політика чи економіка	12200	87200	1630000000
<ul style="list-style-type: none"> Має відкриті коментарі: доступний для втручання агента 	6600	45200	570200000
<ul style="list-style-type: none"> Не публікує спам: втручання агента є ефективним 	6100	24100	192050000
<ul style="list-style-type: none"> Має активну аудиторію: головна ціль агентів 	6000	23800	161000000

12 травня 2022 року, російські користувачі переглянули 87200 публікацій про “СВО”, політику або економіку 1.63 млрд. разів.

Коментарі 45200 публікацій отримали 570 млн. денних переглядів, там можуть бути використаними як ціль для агентів. Атаки на 24000 таких публікацій починають бути ефективними, оскільки це публікації якісних каналів, в яких обговорення користувачами тривають годинами а не хвилинами.

Серед 23800 публікацій з 160 млн. денних переглядів, агентам оптимально залишити коментарі, оскільки з високою імовірністю інші користувачі зможуть побачити і відреагувати на ці коментарі.

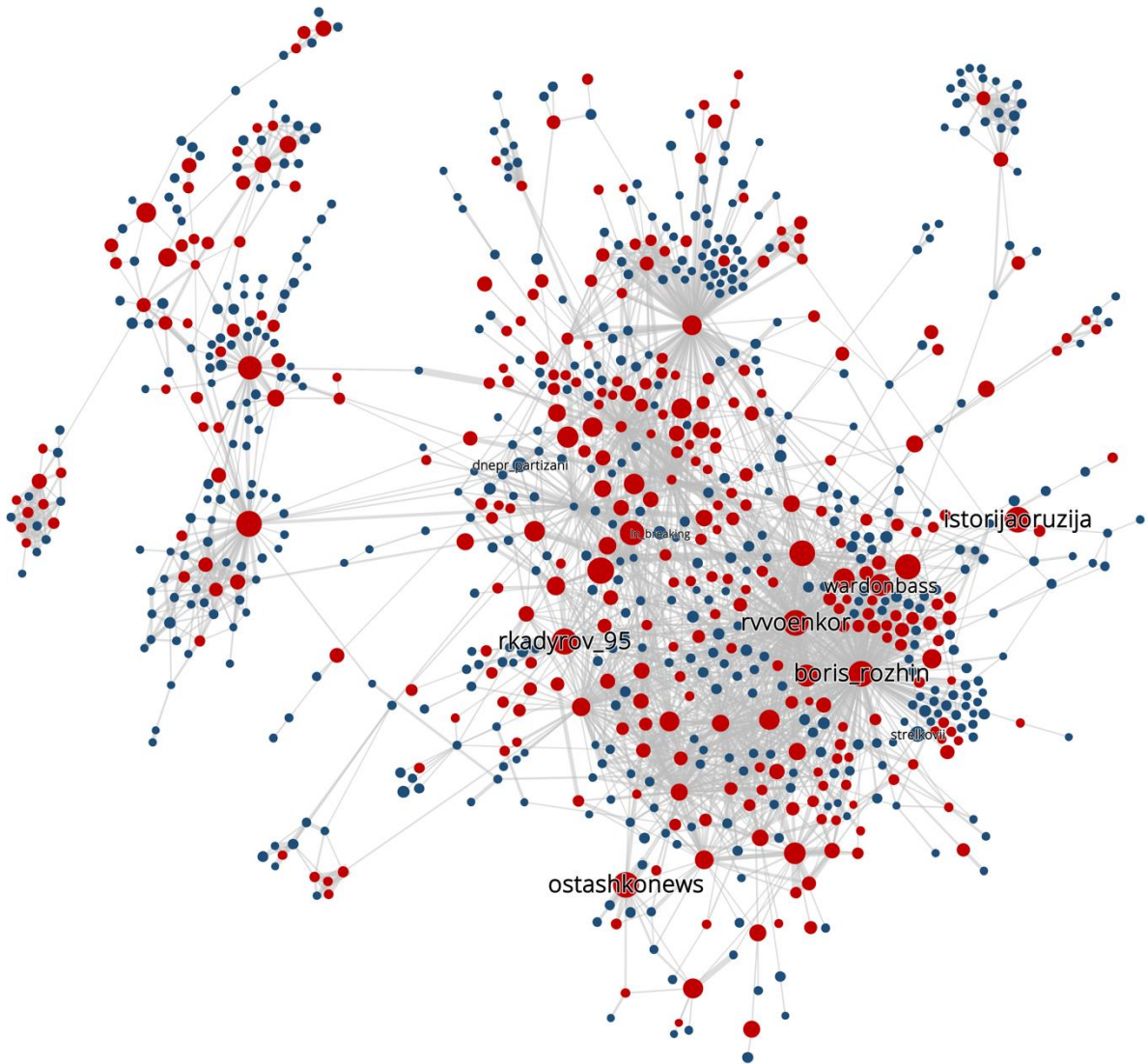


Рис. 3.13. Мережа перетину аудиторій каналів Telegram станом на 12 травня 2022 року. Колір каналу вказує на потенційну ціль для агентів, див Табл. 3.13.

Табл. 3.13. Класифікації каналів на основі ідентифікації цілей агентів.

Колір	Категорія
	Не може бути ціллю агентів
	Є ціллю агентів

Існує важливий аспект використання *Telegram*, оскільки він не надає точної статистики переглядів коментарів. Проте, доцільно використати значення реакцій на публікації (вподобання, тощо) для того, щоб оцінити наступну статистику для 6000 цільових каналів:

- Перший коментар отримує 12% від реакцій на публікацію і може бути переглянутий до 54% користувачів які переглянули публікацію.
- В середньому десятий коментар отримує всього 2% від реакцій на публікацію, проте може бути переглянутий 35% користувачів, які переглянули публікацію.

Відповідно, для максимізації переглядів, агенти повинні діяти продовж кількох хвилин після публікації, для забезпечення вдалого розміщення свої коментарів, див Рис. 3.14 та Рис. 3.15.

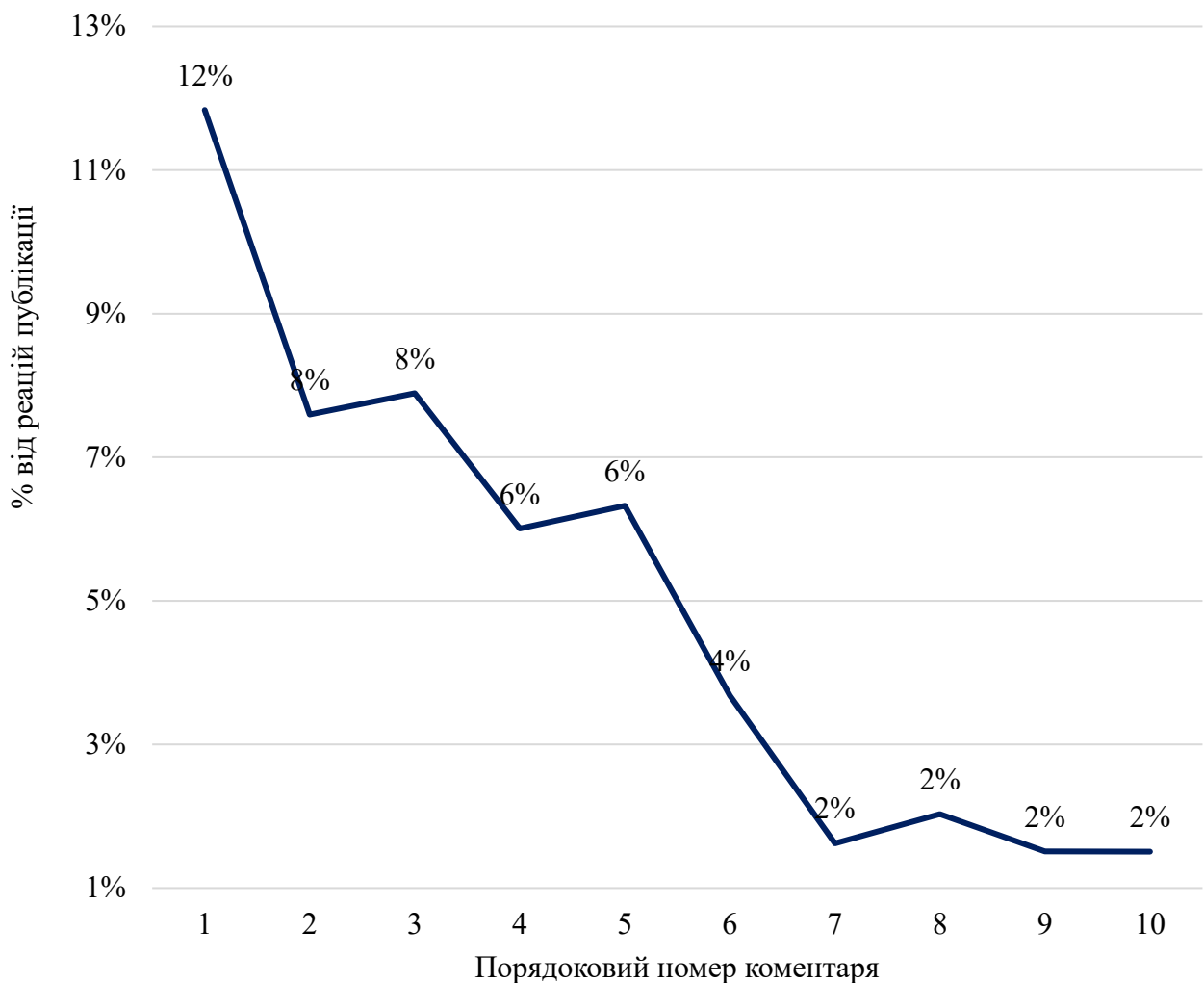


Рис. 3.14. Кількість реакцій отриманих коментарем від загальної кількості реакцій публікації в залежності від його порядкового номера.

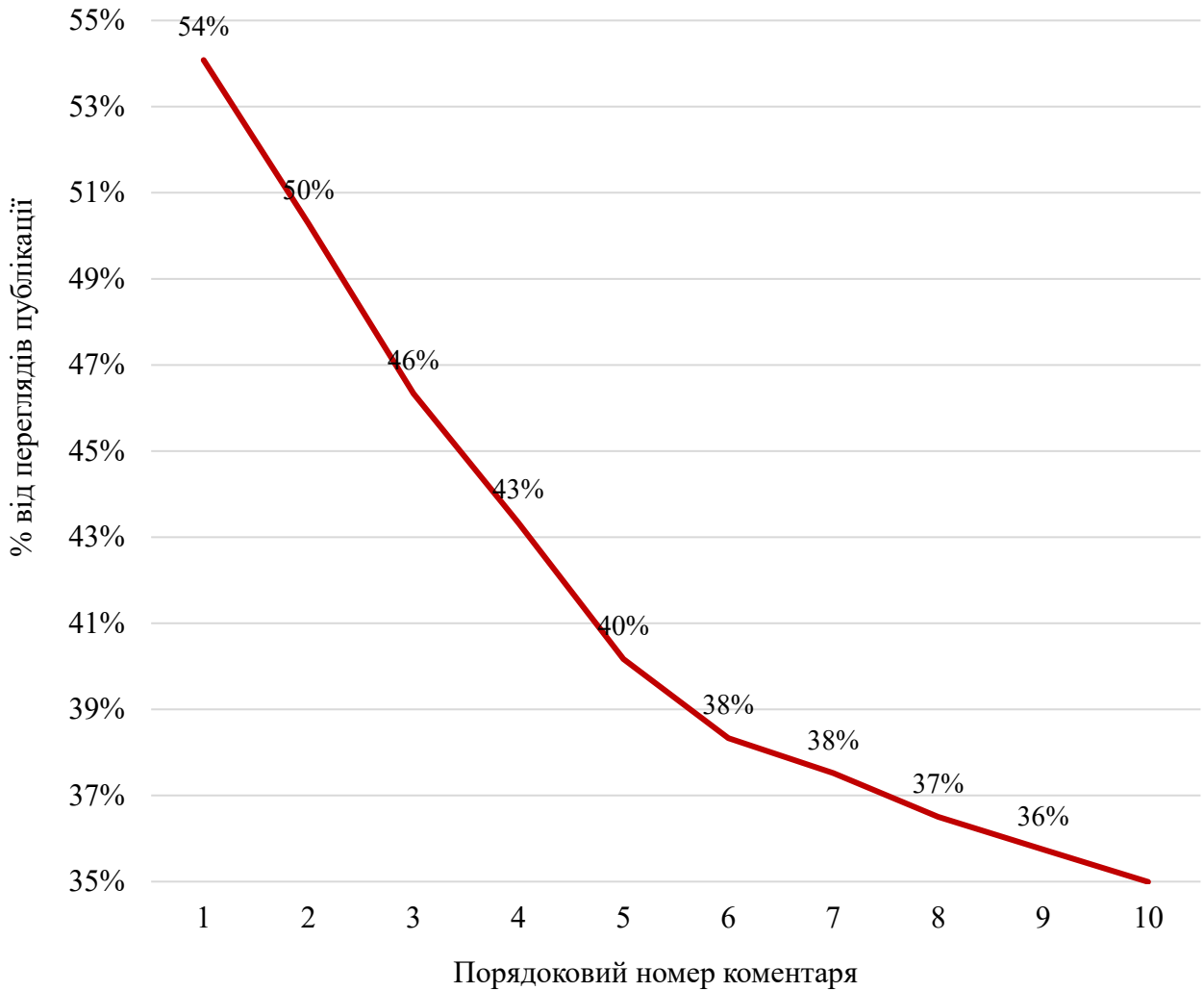


Рис. 3.15. Кількість переглядів отриманих коментарем від загальної кількості переглядів публікації в залежності від його порядкового номера.

3.3.3 Моделювання розповсюдження впливу

Розглянемо симуляцію в якій поведінка кожного агента базується на виконанні двох припущень:

Припущення 1

Нехай структура мережі залишається статичною під час аналізу.

Припущення 2

Для того, щоб агенти сприймалися користувачами як реальні люди, та не блокувалися ними, вони повинні дотримуватися людської поведінки:

- Агенти не можуть публікувати більше 20 коментарів за день.
- Агенти не можуть публікувати коментарі занадто часто: не більше одного коментаря за 10 хвилин.
- Агенти повинні мати реалістичний час реакції на нові публікації.

Вхідні дані

1. Мережа перетину аудиторій *Telegram* каналів.
2. Вектор публікацій за день \bar{P} , з наступними атрибутами:
 - a. Канал $p.channel$
 - b. Час публікації $p.timestamp$
 - c. Вектор коментарів $\overline{p.comments}$ з часом їх публікації та кількістю переглядів.
3. Вектор цільових призначень агентів до *Telegram* каналів. Кожен агент може бути призначеним до декількох каналів.

Алгоритм

Для кожної публікації $p \in \bar{P}$ виконуємо наступні кроки:

1. Зчитуємо вхідні дані.
2. Обираємо агентів, цільове призначення яких включає канал публікації p .
3. Пропускаємо всіх агентів, які використали свій денний ліміт коментарів або опублікували коментар менше ніж 10 хвилин.
4. Для кожного агента який задовольняє необхідні умови, обираємо випадкове значення з інтервалу від 1 до $\|\overline{p.comments}\|$, відображаючи

можливий порядок коментаря, який агент може опублікувати під публікацією.

5. У випадку доступності двох і більше агентів, сортуємо потенційні коментарі за їх порядком, ігноруємо всі коментарі, крім двох з найнижчим порядком.
6. Оновлюємо атрибути агентів, які опублікували коментарі, позначаючи коли відбулось коментування.
7. Записуємо кількість переглядів кожного опублікованого коментаря.

Після того, як всі публікації були опрацьовані, загальна кількість переглядів розраховується як максимальна кількість переглядів коментарів агентів на кожному з каналів.

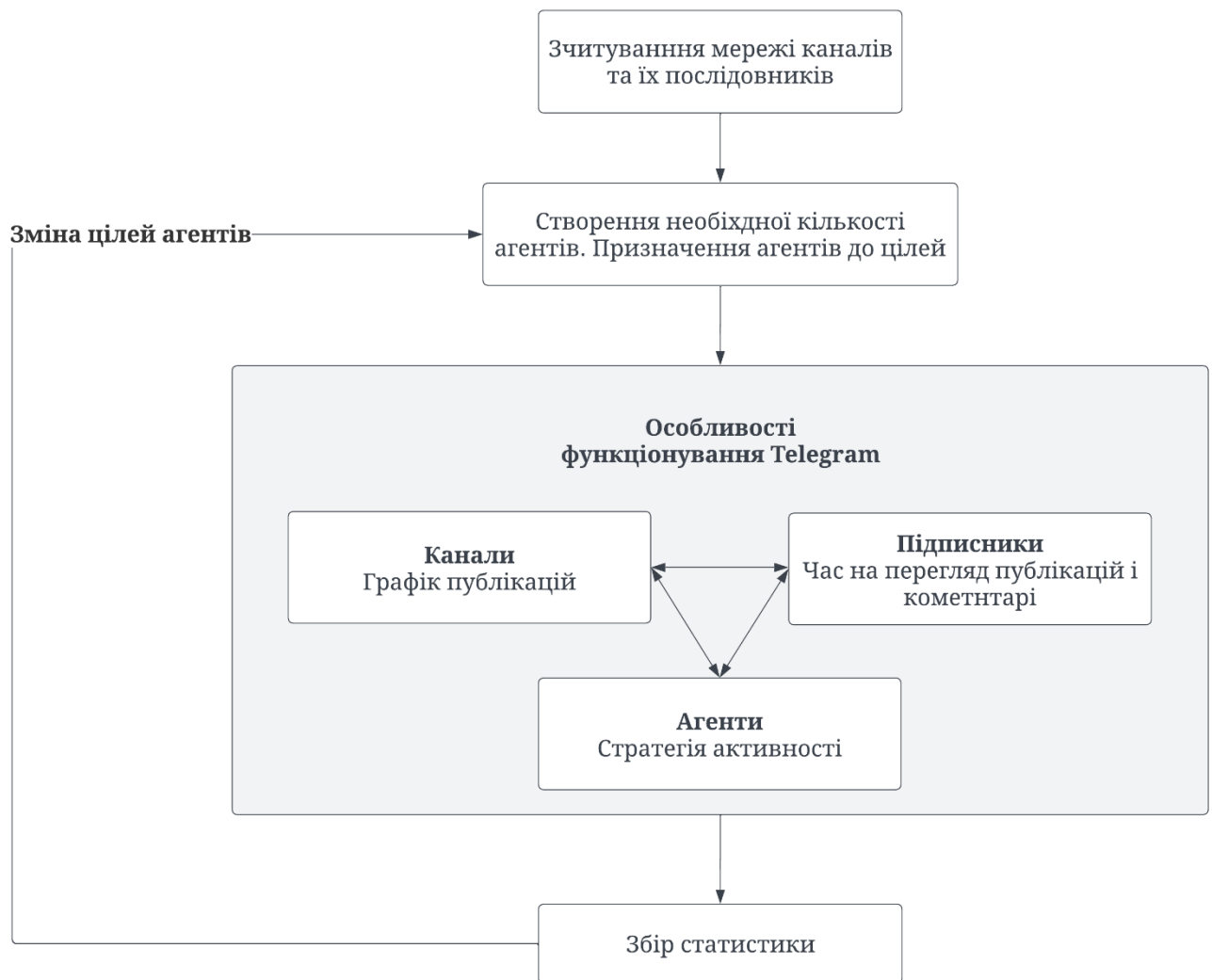


Рис. 3.16. Схема процесу моделювання коментування в *Telegram*.

3.3.4 Аналіз отриманих результатів

Результати моделювання були отримані шляхом симуляції 87222 архівних публікацій та 664309 коментарів опублікованих 12 травня 2022 року.

Кожен з агентів був призначений до 3 каналів, обраних випадковим чином, на основі імовірності пропорційної загальній кількості переглядів коментарів кожного каналу. В той же час, до кожного каналу не могло бути призначено більше 5 агентів.

Ця процедура була виконана 100 разів для поточної конфігурації агентів.

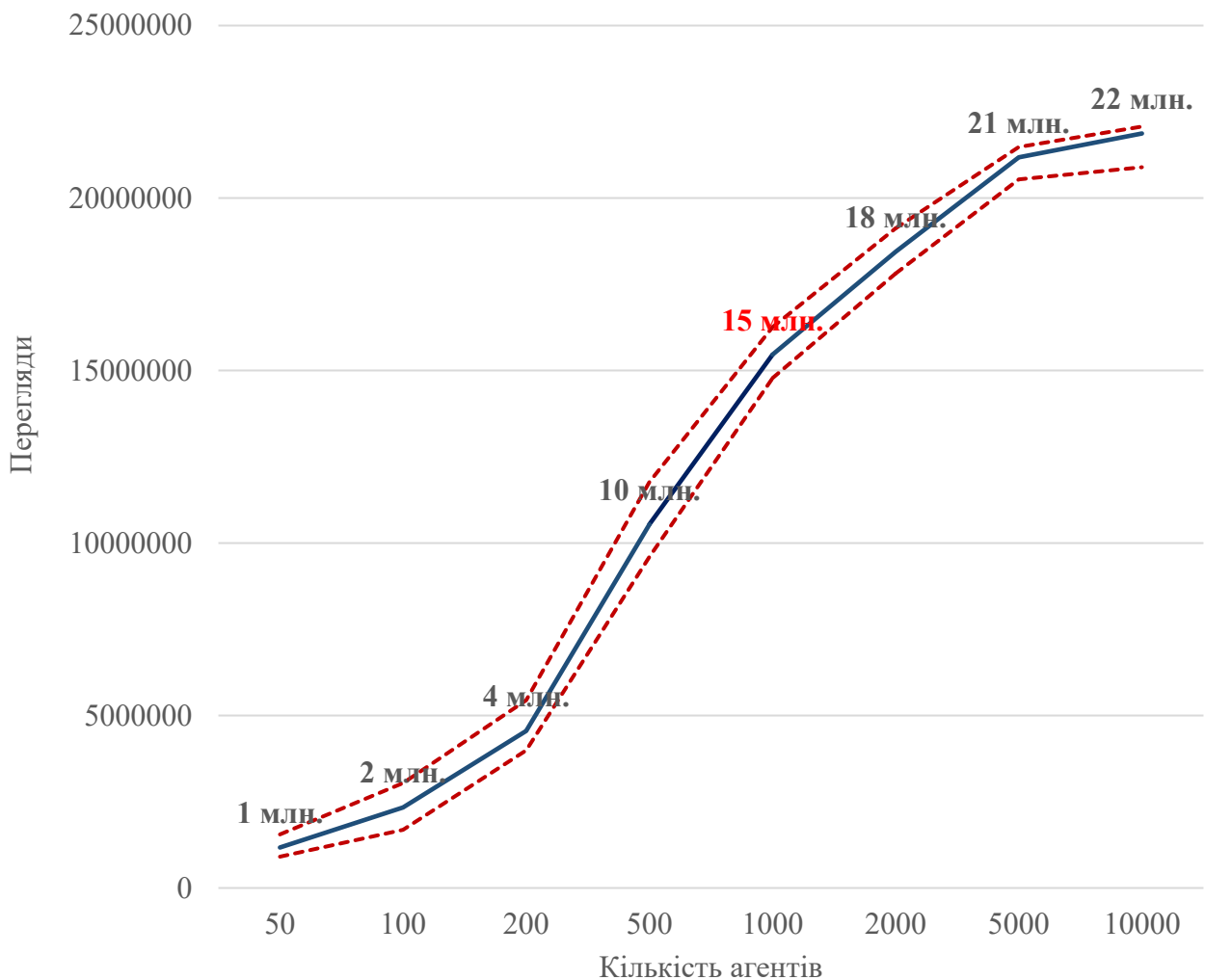


Рис. 3.17. Залежність кількості можливих від кількості агентів.

Основні результати моделювання:

- В середньому, 1000 агентів що координуються спільної метою може генерувати 15 млн. переглядів на день, в той час як 2000 агентів збільшує кількість переглядів до 18 млн.
- Використання більше 2000 агентів не гарантує збільшення кількості переглядів, через відсутність достатньої кількості публікацій, як підпадають під критерії цілей.

Для довготривалої стабільності функціонування групи агентів, їх кількість варто подвоїти, для запобігання втрати ключових коментарів шляхом блокування агентів модераторами каналів.

Для максимальної ефективності роботи, запропонований алгоритм доцільно регулярно тестувати, з використанням різних алгоритмів вибору каналів для агентів.

3.4 ВИСНОВКИ ПО ТРЕТЬОМУ РОЗДІЛІ

У третьому розділі дисертаційної роботи було проведено дослідження особливостей різних соціальних мереж, формалізовано їх основні сутності та способи взаємодії. Для ефективного моделювання процесу розповсюдження інформації, було запропоновано модель мережі перетину аудиторій на основі публічних даних. В умовах обмеження соціальними мережами доступу до даних, було показано як модель мережі подібності публікацій та мережі взаємодій дозволяють оцінювати перетин аудиторій. З використанням мережі перетину аудиторій, було досліджено динаміку еволюції спільнот соціальної мережі *Telegram* та процес розповсюдження впливу.

Для розв'язання задачі оптимального поширення впливу було запропоновано модель з використанням агентів, а мережу перетину аудиторій використано для вибору оптимальних цільових аудиторій для агентів. Для ефективного застосування агентів, було проведено моделювання поширення інформації для випадків з різною кількістю агентів.

РОЗДІЛ 4: АНАЛІЗ СОЦЬЛЬНИХ МЕРЕЖ ЕКОНОМІЧНИХ ВІДНОСИН

Впродовж останніх 50 років світ зіткнувся з численним, єдиними у своєму роді пандеміями вірусних захворювань. Серед яких як пандемія *HIV/AIDS* (з 1981 року по сьогодні), пандемія грипу *H1N1* (з грудня 2009 року до серпня 2010 року), окремі випадки епідемії гарячки Ебола (з 2013 до 2016 року) та *SARS* епідемія тяжкого гострого респіраторного синдрому (з 2002 до 2004 року). Кожна з цих епідемій та пандемій, незважаючи на присутність в усіх куточках світу, залишається сфокусованою на окремих характеристиках способів передачі вірусу. Проте, пандемія *COVID-19* виділяється серед усіх інших, призводячи до глибоких наслідків для усіх країн. Пандемія призвела до 6.5 млн. смертей, породжуючи небачені до цього кризи та ставлячи виклик системам охорони здоров'я.

Початкові етапи пандемії були неймовірно хаотичними, уряди країн з найбільшою кількістю нових захворювань, усіма силами намагалися зупинити розповсюдження вірусу приймаючи рішення на основі експертних оцінок різноманітних науковців. Проте, такий підхід виявився неефективним. Багато з рішень було прийнято на основі експертних оцінок, що в свою чергу базували на використанні статистичних моделях в яких повністю відсутній поведінковий аспект.

В комбінації з іншими факторами, уряди не змогли запобігти розповсюдженню вірусу, а системи охорони здоров'я стали перевантаженим по усьому світу. Слідом, за введенням карантинів, люди стали ізольованими у власних помешканнях, що стало одним з перших кроків наближення економічної кризи. Бізнеси та приватні підприємства були змушені припиняти свою діяльність, звільняти працівників, тим самим створюючи кризу ринку працевлаштування. В цей момент, урядам довелося створювати програми для протидії кризі. На жаль, багато з них виявилися малоефективними, в першу чергу відсутність розуміння специфіки основних процесів, що спричиняють кризи.

Для систематичного вирішення таких проблем, в цьому розділі було досліджено реальні наслідки пандемії *COVID-19* для населення та економіки штату Гаваї. Розвиток економічне зростання штату Гаваї, що складається з 8 основних островів, в цілому залежав від доходів з туристичної галузі. Проте, епідемія цілком зупинила туризм, змусивши уряд штату активно втручатися в економічну ситуацію за допомогою численних програм підтримки населення та його зайнятості. Усвідомлюючи невизначеність потенційних наслідків таких програм, та імовірність появи нових хвиль пандемії, уряд штату сфокусувався на дослідженні основних показників економічної ситуації:

1. На скільки значними є поточні економічні збитки від пандемії *COVID-19*?
2. На скільки урядові програми змогли зменшити ці збитки?
3. Якою повинна бути швидкість відновлення кількості туристів та відвідувачів островів для стабілізації економіки штату?

Для надання уряду та іншим впливовим організаціям точні та швидкі відповіді, було розроблено комплексну систему, що складається з трьох основних компонент: синтез даних, агентне моделювання та дослідження сценаріїв.

Синтез даних

Під час процедури синтезу було використано виключно публічно доступні дані, такі як (a) реєстраційні дані бізнесів та працевлаштувань, (b) щорічні опитування діяльності бізнесів, дані оподаткування підприємств між галузями економіки, (c) анонімізована 5-відсоткова вибірка даних перепису населення США *IPUMS* [65], (d) статистичні оцінки ключових характеристик населення, (e) дані туризму. Під час проведення дослідження, ці дані були актуалізовані для максимально точного відображення популяції та економіки штату на кінець 2019. Синтезовані дані відображають штат в вигляді синтетичної популяції - точної копії населення, підприємств та економічних зв'язків.

Агентне моделювання

Агентна модель відображає економічні дії домогосподарств та підприємств на тижневому інтервалі. Домогосподарства приймають рішення про тижневі витрати, а підприємства – про бажаний рівень зайнятості. Ці рішення спричинені кількістю туристів, та залежними від сценарію урядовими програмами, карантинном, обмеженнями мобільності населення та захистом їх доходів.

Дослідження сценаріїв

Моделювання з використанням агентної моделі проводиться на всій синтетичній популяції, в якій всі домогосподарства та підприємства діють виходячи з власних корисних інтересів, впродовж року. Агенти взаємодіють між собою таким чином впливаючи на аргументацію дії інших агентів. Враховуючи внутрішню стохастичність під час прийняття рішень, невизначеність початкових умов в синтетичній популяції та динамічну природу прогнозів і сценаріїв, було проведено сотні незалежних досліджень для кожного сценарію, а також сотні тисяч досліджень присвячених тестуванню та калібрації моделі.

4.1 ПОБУДОВА СИНТЕТИЧНОЇ ПОПУЛЯЦІЇ

В цьому розділі, було запропоновано комбінований підхід до побудови (синтезу) комплексної синтетичної популяції осіб, господарств, підприємств, та мережі економічної активності між ними. На Рис. 4.1 представлено детальну діаграму етапів створення синтетичної популяції. Розглянемо детально кожен з етапів створення.

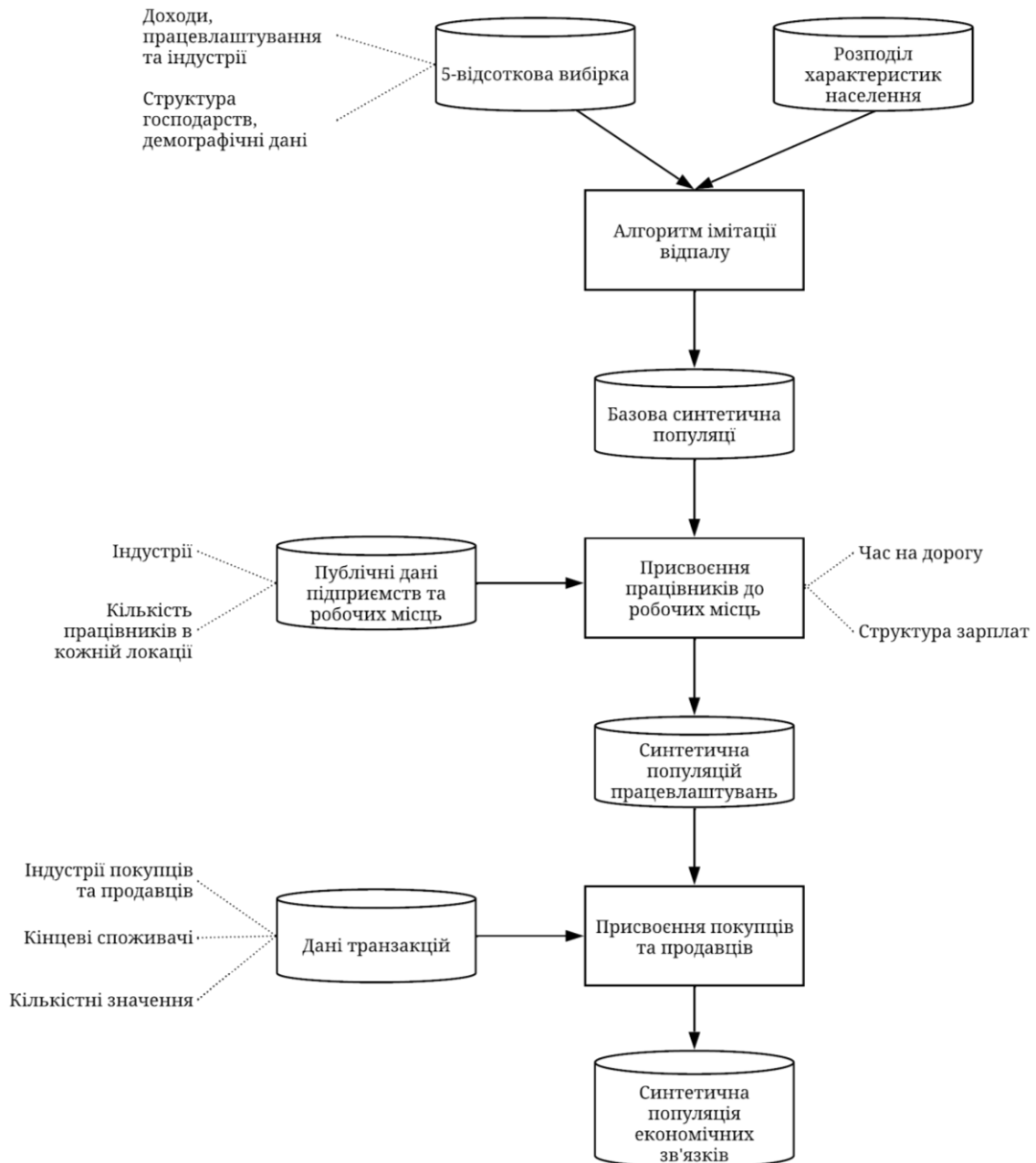


Рис. 4.1. Діаграма етапів створення синтетичної популяції осіб, господарств та підприємств.

4.1.1 Набір даних для побудови синтетичної популяції

Основним призначенням бюро перепису населення США є проведення систематичного перепису населення США для визначення кількості представників від кожного штату в Конгресі. Проте, бюро перепису також

проводить багато програм дослідження різноманітних соціоекономічних характеристик населення. Не зважаючи на проведення перепису раз на 10 років, бюро займається аналізом щорічних змін в структурі населення, та надає актуальні і історичні анонімізовані дані для дослідників у вигляді 5-відсоткової вибірки.

Вибірка в свою чергу складається з двох груп даних: осіб та господарств до яких належать ці особи. Кожен анонімізований запис у вибірці вказує на приблизне місцезнаходження господарства, його соціоекономічні, демографічні та поведінкові характеристики та осіб що до нього належать. Кожне господарство описане за допомогою 96 атрибутів, а особи, що до нього належать описані за допомогою 210 атрибутів. Для прикладу, домогосподарство що складається з 5 осіб, двоє з яких одружені, з сумарним доходом в \$80000, що орендують житло, та отримують матеріальну допомогу від уряду. В цьому господарстві чоловік витрачає 75 хв на дорогу до місця працевлаштування. В той же час в цьому господарстві також проживає студент, та особа на пенсії. Дані вибірки створенні таким чином, щоб максимально описати унікальну структуру населення, тому для штату Гаваї вона складається з 28337 господарств та 72078 осіб. Приблизні місцезнаходження записів з вибірки обмежені на локації публічного використання *PUMAs*. Штат Гаваї в загальному складається з 137 островів, серед яких виділяють 8 основних, та описується 10 локаціями *PUMA*, див Рис. 4.2.

Проте, даних вибірки не достатньо для побудови синтетичної популяції. В першу чергу це пов'язано з відсутністю даних на місцевому рівні, оскільки одна локації *PUMA* описує територію з населенням в сотні тисяч мешканців. Для розв'язання цієї проблеми, було використано статистичні дані бюро перепису США. Ці дані надають оцінки розподілу характеристик на рівні локацій перепису *Census tract* (351 локація), див Рис. 4.3. Для прикладу, в локації перепису проживає 91400 осіб віком від 40 до 44 років. Характеристики для яких доступні розподіли, частково присутні у вибірці перепису населення.

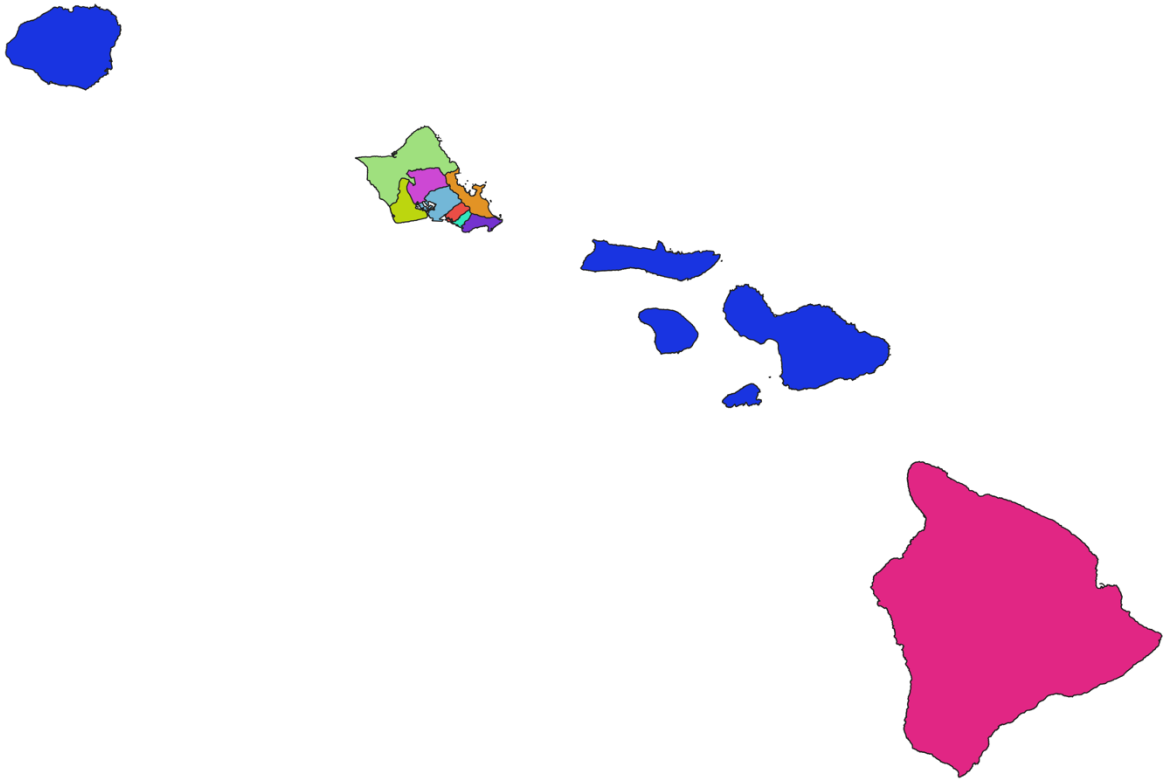


Рис. 4.2. Структура локацій *PUMAs* штату Гаваї.

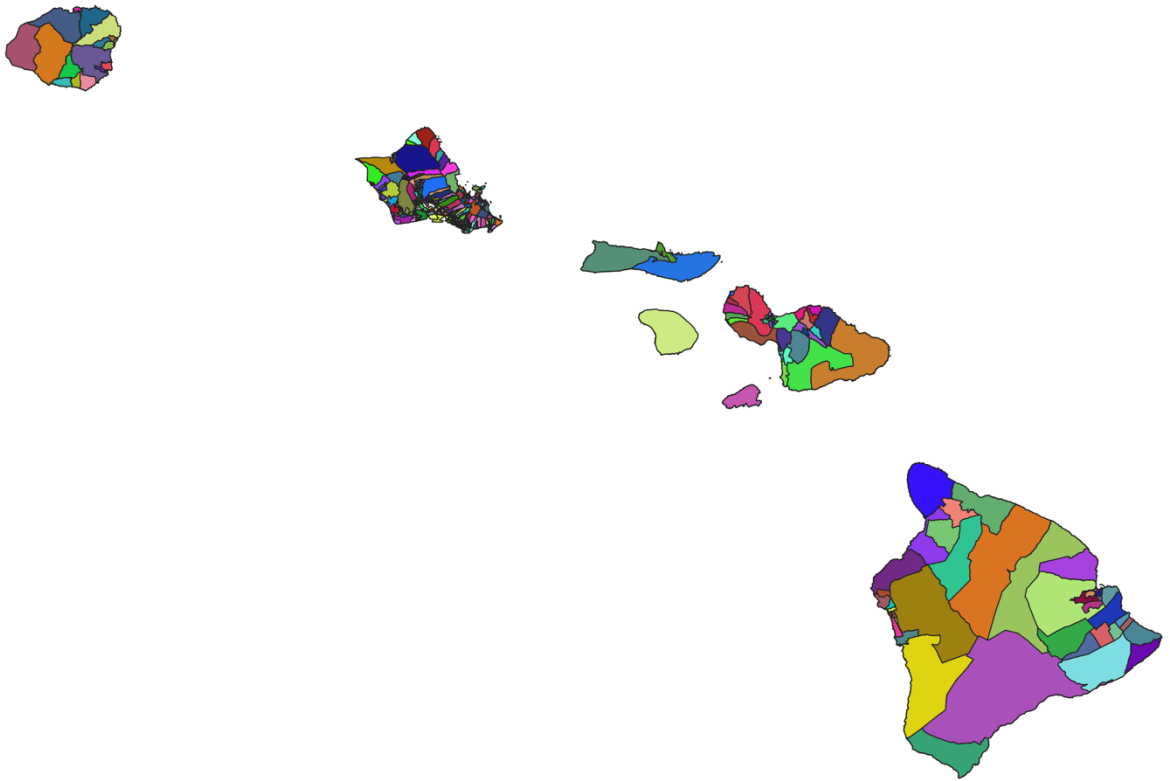


Рис. 4.3. Структура локацій перепису населення *Census tracts* штату Гаваї.

Таким чином створення синтетичної популяції повинно відбуватися з використанням вибірки та характеристик населення.

4.1.2 Синтез господарств та осіб

Для створення синтетичної популяції необхідно провести оптимізацію підбору необхідних осіб та господарств для досягнення бажаних розподілів.

Розглянемо алгоритм підбору на основі алгоритму імітації відпалу [66], для кожної локації перепису населення.

Вхідні дані

1. 5-відсоткова вибірка динних перепису населення, 96 атрибутів господарств та 210 атрибутів осіб.
2. Розподіл найбільш важливих атрибутів для кожної локації перепису:
 - a. Розподіл осіб за віком (інтервали в 5 років) та статтю.
 - b. Розподіл домогосподарств за розміром (точні значення).
 - c. Розподіл домогосподарств за загальним доходом (інтервали в \$10000) з обмеженням максимального доходу в \$250000.
 - d. Розподіл домогосподарств за власністю житла (оренда або володіння), або проживання в спеціальних умовах (гуртожитки, госпіси, притулки).
 - e. Розподіл кількості працевлаштованих осіб за індустрією.
 - f. Розподіл осіб за роком закінчення навчання.

Алгоритм

Нехай, для вибраної локації потрібно побудувати N господарств.

1. Ініціалізація:

- a. Зчитуємо вхідні дані.
 - b. Жадібним чином обираємо N господарств для вибраної локації які відображають початковий стан синтетичної популяції.
 - c. Розраховуємо початкове значення функції втрат L , як суму відстаней між усіма гістограмами бажаного розподілу та поточного розподілу атрибутів.
2. Алгоритм імітації відпалу, виконуємо 2 млн. ітерацій:
- a. Випадковим чином обираємо господарство з синтетичної популяції $L_{synthetic}$.
 - b. Випадковим чином обираємо господарство з вибірки L_{sample} .
 - c. Замінюємо $L_{synthetic}$ на L_{sample} в синтетичній популяції.
 - d. Розраховуємо нове значення функції втрат L' з врахуванням заміни.
 - e. Розраховуємо приріст функції втрат $\Delta L = L' - L$.
 - f. Якщо значення $\Delta L < 0$, приймаємо заміну.
 - g. Якщо значення $\Delta L \geq 0$, приймаємо заміну з ймовірністю $P = e^{-\Delta L/T}$, де T – параметр температури, що зменшується з кожною ітерацією алгоритму.
 - h. Оновлюємо значення $L = L'$.
3. Збереження синтетичної популяції осіб та господарств.

Варто зазначити, що початкове значення параметру температури T є великим, проте зменшується на кожній ітерації. Це гарантує, що на початкових ітераціях алгоритму зміни приймаються з вищою імовірністю, навіть якщо вони спричиняють збільшення значення функції втрат. По мірі виконання алгоритму значення T зменшується, дозволяючи приймати лише заміни, що мінімізують функцію втрат.

З використанням ефективної параметризації, алгоритм імітації відпалу дозволяє будувати синтетичні популяції, що ідеально відповідають бажаними характеристикам населення. На Рис. 4.4 зображено перевірку ефективності роботи алгоритму з використанням критерію узгодженості Колмогорова [67], на

розподілі доходів господарств. В найгіршому випадку, на островах, де проживають одинці, похибка становить не більше 3%, а для густо населених районів адміністративного центру штату міста Гонолулу, похибка не перевищує 1%. Побудована синтетична популяція налічує 1.2 млн. господарств та 2.9 млн. осіб.

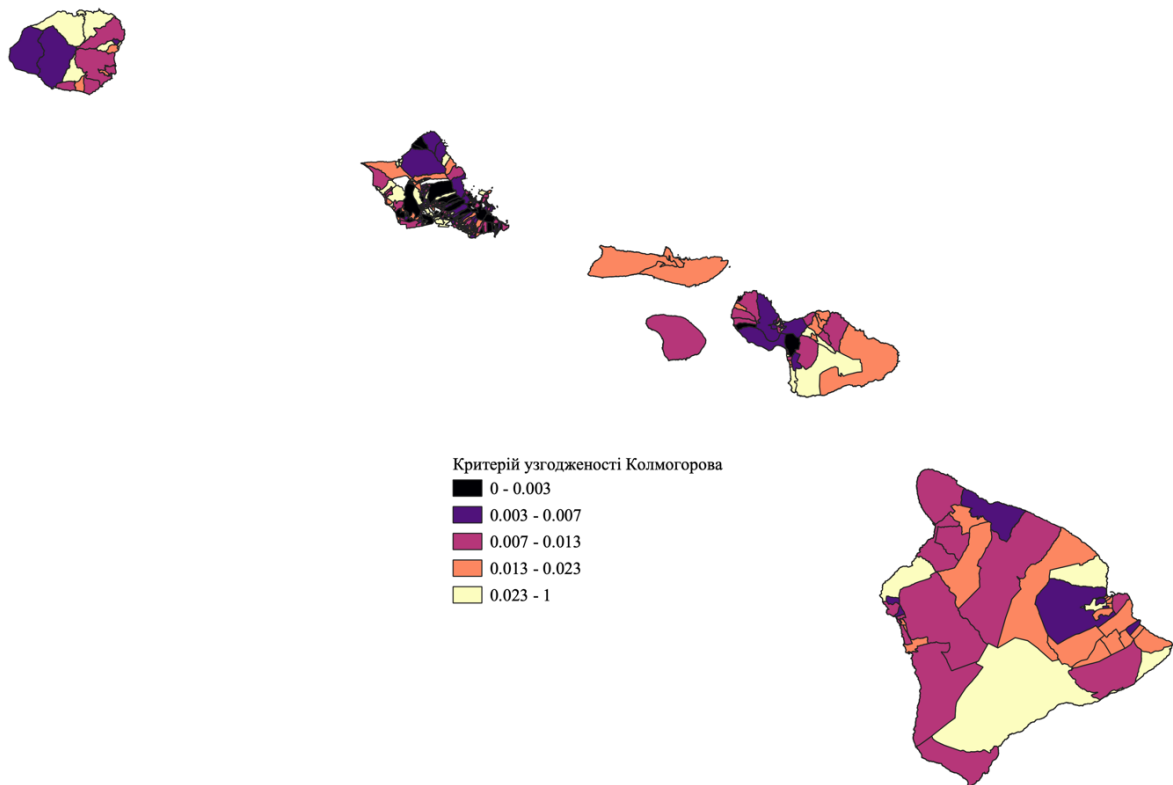


Рис. 4.4. Валідація синтетичної популяції з використання критерію узгодженості Колмогорова.

4.1.3 Синтез підприємств та робочих місць

Для синтезу підприємств та робочих місць необхідно провести схожу оптимізацію підбору осіб, активних на ринку праці, для конкретних підприємств.

Розглянемо алгоритм підбору на основі алгоритму переважного приєднання, при якому виконується принцип, що чим більше працівників працюють на підприємстві тим більша імовірність появи нового працівника.

Використання алгоритму дозволило прискорити складний процес побудови вибірки, та зберегти емпіричний принцип зайнятості, при якому 1% підприємств наймають 50% співробітників.

Вхідні дані

1. Синтетична популяція осіб. Їх точне розташування, час на дорогу до місця працевлаштування в хвилинах, та галузь в якій вони працюють.
2. База даних подорожей. Час подорожі між двома, будь-якими, локаціями, отриманий на основі даних громадського транспорту та автомобільної інфраструктури.
3. Реєстраційні дані підприємств та робочих місць. Їх точне розташування, кількість працівників на конкретному робочому місці, а також основний галузевий ідентифікатор.

Алгоритм

1. Ініціалізація:
 - a. Зчитуємо вхідні дані.
 - b. Жадібним чином обираємо робоче місце для кожної особи, на основі переважного приєднання, переконуючись що час на дорогу знаходиться в діапазоні (+100%, -80%).
2. Обчислення функції втрат L як суму помилок для всіх працівників та робочих місць у випадку якщо:
 - a. Галузь працівника відрізняється від галузі підприємства.
 - b. Працівник змушений витратити на дорогу більше або менше очікуваного часу (+30%, -30%).
 - c. До робочого місця приєднано забагато або замало працівників.
3. Алгоритм підбору:

- a. Випадковим чином обираємо вже приєднаного працівника з робочим місцем $B_{synthetic}$, та підбираємо для нього альтернативне робоче місце B_{sample} .
 - a. Замінюємо робоче місце $B_{synthetic}$ на B_{sample} в синтетичній популяції.
 - b. Розраховуємо нове значення функції втрат L' з врахуванням заміни.
 - c. Розраховуємо приріст функції втрат $\Delta L = L' - L$.
 - d. Якщо значення $\Delta L < 0$, приймаємо заміну.
 - e. Якщо значення $\Delta L \geq 0$, приймаємо заміну з ймовірністю $P = e^{-\Delta L/T}$, де T – параметр температури, що зменшується з кожною ітерацією алгоритму.
 - f. Оновлюємо значення $L = L'$.
4. Збереження синтетичної популяції осіб та господарств.

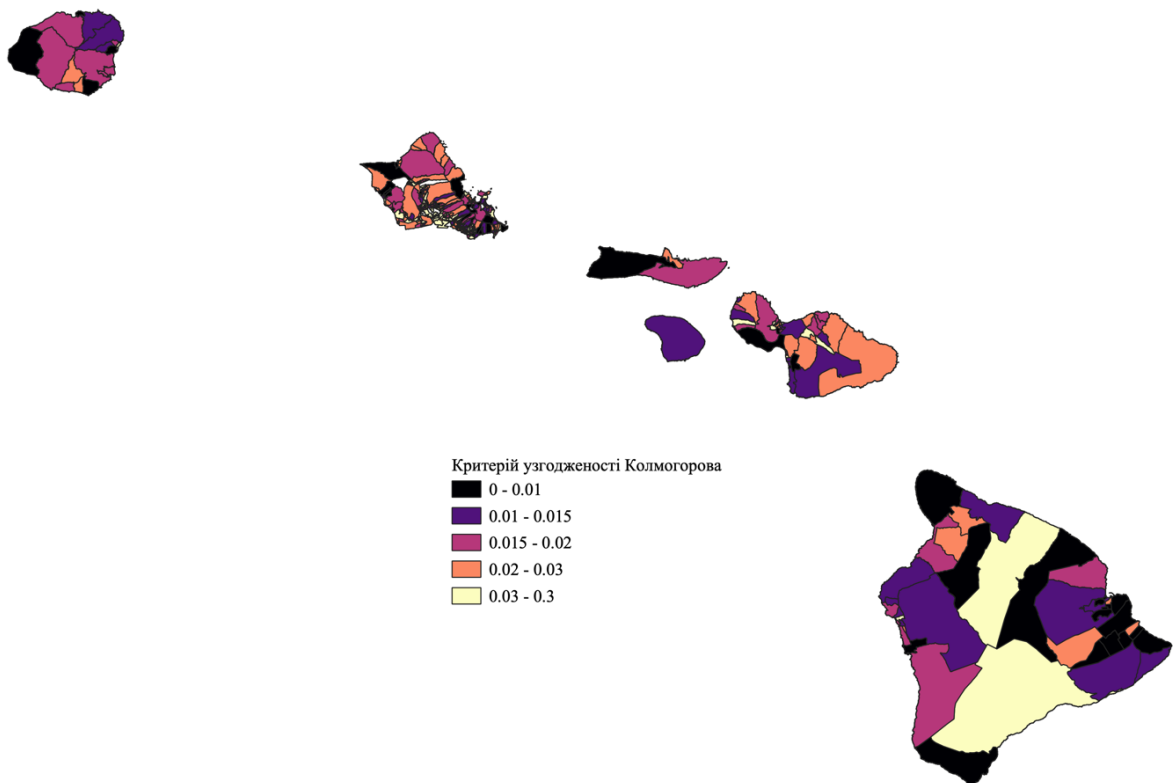


Рис. 4.5. Валідація синтетичної популяції з використання критерію узгодженості Колмогорова.

На Рис. 4.4 зображено перевірку ефективності роботи алгоритму з використанням критерію узгодженості Колмогорова [67], на відповідність даним адміністрації малого бізнесу. До попередньої синтетичної популяції було додано 75000 підприємств та 1.3 млн. їх працівників.

4.1.4 Синтез мереж постачання продуктів та послуг

Останнім кроком побудови синтетичної популяції, є побудова зв'язків туристами, господарствами та підприємствами в мережі економічних транзакцій. Для цього було застосовано модифікований алгоритм переважного приєднання, де кожен турист, господарство та підприємство були з'єднані з принаймні одним іншим підприємством для кожної з 20 галузей. Також туристам та господарствам було дозволено з більше ніж одним підприємством з галузей роздрібною торгівлі та громадського харчування. Вага таких з'єднань відображає загальний щорічний попит на товари та послуги яке надає конкретне підприємство. Побудова зв'язків базувалася на даних транзакцій оприлюднених бюро економічного аналізу США, див. Додаток 1. Цей крок виявився складним з точки зору складності обчислювальних операцій, тому його реалізація була виконана на мові Julia [68].

До попередньої синтетичної популяції було додано 106 млн. зв'язків економічної активності.

Розглянемо окремі випадки синтезу підприємств, аналізуючи структуру поставок та замовлень. Випадковим чином було обрано два підприємства, одне з галузі національної безпеки, інше з галузі нерухомості.

Виходячи зі структури мережі, див Рис. 4.6, підприємство з галузі національної безпеки є одним з підприємств які тісно працюють з Міністерством оборони США. Воно спеціалізується на організації та проведенні заходів навчання військових з інших штатів їх реабілітації, та побудови інфраструктури їхнього проживання. Природньо, це підприємство отримує кошти на проведення

діяльності від уряду США, і згодом ефективно розпоряджається ними, замовляючи послуги в приватних компаній штату. Тому, в основному це підприємство буде організовувати авіаперельоти, згодом поселяти військових в готелі, забезпечувати їх харчування та прибирання приміщень, в яких відбуваються навчання, орендувати необхідне спорядження та готувати тренувальні майданчики, навчальні центри. В той же час, підприємству потрібно забезпечити постійне проживання інструкторів та співробітників.

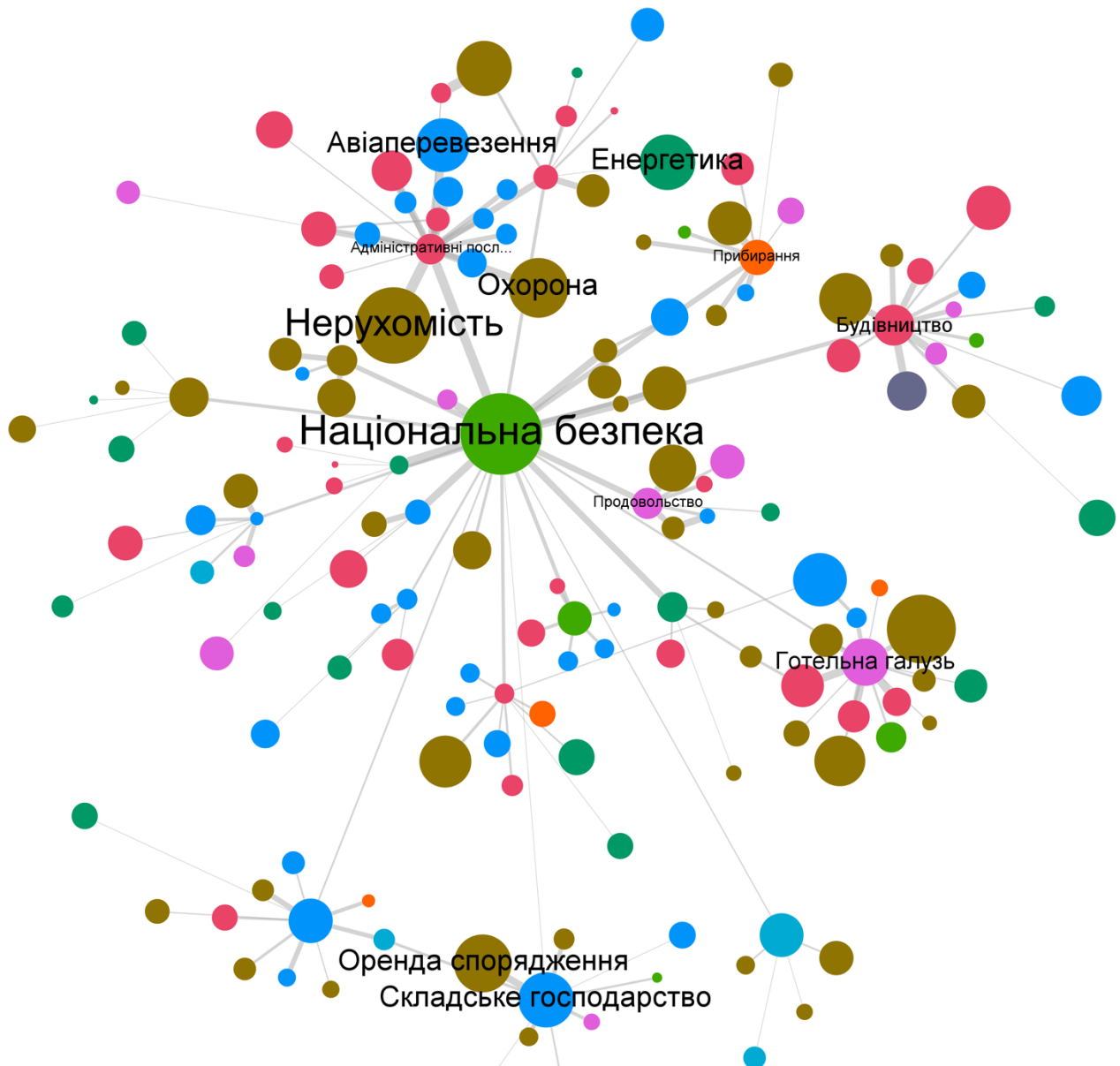


Рис. 4.6. Мережа постачання між підприємствами для вибраного підприємства галузі національної безпеки. Кольори вершин відповідають галузям підприємств, див Табл. 4.1.

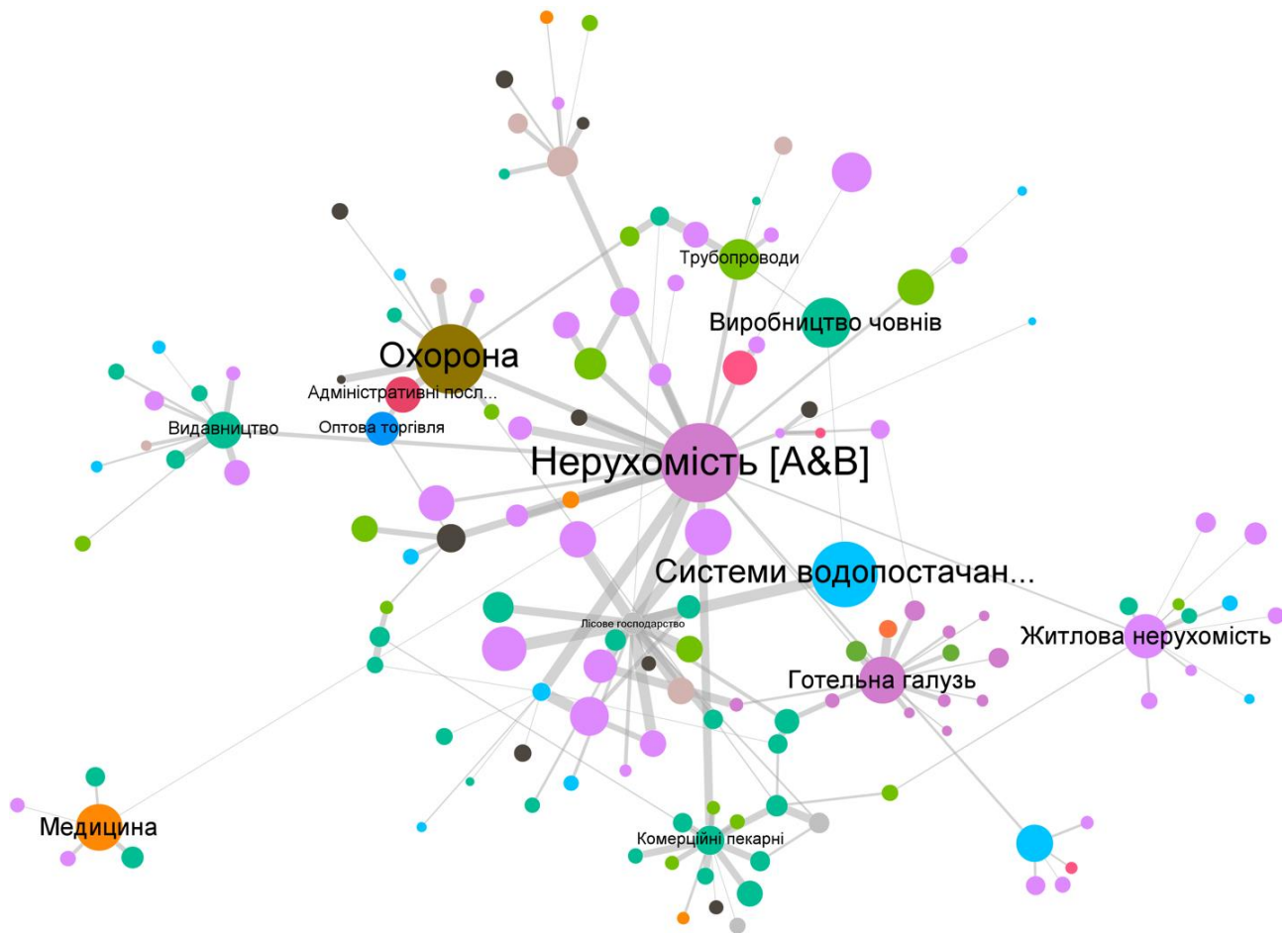


Рис. 4.7. Мережа постачання між підприємствами для вибраного підприємства галузі нерухомості. Кольори вершин відповідають галузям підприємств, див Табл. 4.1.

Економіка штату Гаваї є доволі різноманітною, та фокусується на багатьох галузях, проте в ній домінує присутність “Великої п’ятірки” підприємств. Аналізуючи структуру мережі постачань випадкового обраного підприємства зі сфери нерухомості, див Рис. 4.7, було ідентифіковано відоме підприємство *Alexander & Baldwin*. Це підприємство володіє значною кількістю сільськогосподарських земель на різних островах штату Гаваї, в той же час воно інвестує у всі можливі індустрії, диверсифікуючи свої активи. Тому в мережі постачань можна побачити сильні зв’язки з підприємствами водопостачання та зрошень земель, підприємства, що виготовляють трубопроводи, підприємства оптової торгівлі сільськогосподарською продукцією. В той же час, це

підприємство активно інвестує в медицину, будуючи власні лікарні, в житлову нерухомість, а також в засоби масової інформації, монополізуючи економіку.

Табл. 4.1. Характеристика індустрій підприємств.

Колір	Загальна індустрія
	Нерухомість, фінанси, страховий сектор
	Промисловість
	Оптова торгівля
	Комунальні послуги
	Проживання та харчування
	Державне управління
	Освіта та охорона здоров'я
	Інші послуги
	Сільське господарство, лісове господарство, рибальство, мисливство

4.1.5 Аналіз отриманих результатів

Вибір цього штату Гаваї як основного об'єкту дослідження був спричинений потребою місцевих органів влади, в розумінні соціальних та демографічних характеристик населення в умовах економічної кризи. Агентне моделювання на побудованій синтетичній популяції дозволило надати критичну інформацію для підтримки прийняття рішень. Окремо, синтетична популяція була ефективно застосованою при розробці та моделюванні бюджетів благодійних організацій штату не пов'язаних з пандемією.

Вибір такого об'єкту дослідження мав свої переваги. США мають розвинену та добре фінансовану мережу федеральних організацій, завдання яких полягає в детальному аналізі суспільних процесів. Незважаючи на те, що метою цих організацій, є в першу чергу оптимізація збору податків населення, проте

дані які вони акумулюють та в подальшому публікують для дослідників є надзвичайно якісними.

В умовах необхідності проведення схожого дослідження в інших штатах або країнах, можуть виникнути додаткові проблеми пов'язані з якістю даних. Країни Європи та Північної Америки систематично проводять переписи населення, в той час як інші, не завжди мають фінансові можливості його провести. Наприклад, останній перепис населення в Україні відбувся в 2001 році, унеможливаючи проведення будь-яких досліджень.

Для вирішення цієї проблеми, можна використовувати альтернативні набори даних [69], в яких досліджується динаміка зміни густини населення різних країн. Ці дані, в комбінації з алгоритмом імітації відпалу та алгоритмом переважного приєднання, дозволять побудувати синтетичну популяцію схожої якості.

4.2 ПОБУДОВА АГЕНТНОЇ МОДЕЛІ СОЦІАЛЬНИХ ВЗАЄМОДІЙ

Основний принцип агентного моделювання полягає у створенні симуляції взаємопов'язаних агентів, кожен з яких описаний набором атрибутів та поведінкою. Механізм роботи симуляції в загальному доволі простий, на кожному кроці (часовому інтервалі) відбувається активація агентів (кожен агент має власний графік активації), які взаємодіють між собою, змінюючи значення атрибутів.

Для початку проаналізуємо властивості тривіального прикладу агентного моделювання. Розглянемо мережу технологічних інновацій, в якій агентами виступають компанії, що співпрацюють між собою, діляться, розвивають та іноді крадуть інновації. Мережа постійно зазнає перебої через створення нових принципових технологій.

Атрибути агентів

1. База знань: набір технологій про які агент знає на поточному кроці симуляції.
2. Критерій співпраці: міра готовності агента співпрацювати з іншими.
3. Критерій інновацій: наскільки швидко агент може створювати або вдосконалювати технології.
4. Критерій безпеки: наскільки агент захищає свої технології від крадіжки.

Поведінка агентів

1. Інновації: на основі поточної бази знань агенти періодично пропонують вдосконалення технологій або створюють нові.
2. Співпраця: агенти з високим критерієм співпраці можуть об'єднуватися, ділитися своєю базою знань, і консолідувати зусилля для прискорення інновацій.
3. Конкуренція: агенти прагнуть стати першими в реалізації нової технології.
4. Захист та крадіжки: для пришвидшення власних інновацій, агенти можуть копіювати прогрес агентів зі слабким критерієм безпеки, тоді як інші інвестують у збільшенні рівня безпеки для захисту від копіювання.

Динамічна поведінка агентів

1. Кластери інновацій: з часом агенти об'єднуються у високоефективні групи, що сприяють швидкому виникненню інновацій.
2. Лідери та аутсайтери: з часом агенти з високим рівнем інновацій визначають динаміку розвитку інновацій, на них рівняються інші, менш успішні агенти.

3. Розповсюдження інновацій: виходячи з співпраці між агентами, відбувається колективне розповсюдження інновацій, створюючи множину спільних ідей та технологій.

Така агента модель симуляція дозволяє досліджувати вплив регуляторних сценарії на процес створення інновацій, допомагаючи створювати ефективну мережу інновацій.

4.2.1 Формалізація поведінки агентів

Розглянемо поведінки агентів симуляції економічних мереж штату Гаваї в умовах пандемії *COVID-19*.

Дії уряду

Дія федерального та місцевого уряду для допомогти економіці штату наведені в Табл. 4.2 та відповідають реальним програмам, включаючи їх критерії спрямування, терміни реалізації та бюджети. У випадку, коли господарства отримують доходи вищі за номінальні доходи до пандемії *COVID-19*, або заощадили кошти які не могли витратити під час карантину, було введено функцію, що визначає скільки з цих коштів буде витрачено на негайне споживання, скільки для збільшення заощаджень, а скільки для зниження боргового навантаження.

Табл. 4.2. Урядові програми та їх плив на агентів.

Програма	Спрямування	Характеристика	Недоліки
Карантин	Господарства та підприємства	Деякі підприємства, в залежності від галузі, отримують наказ припинити свою діяльність або працювати в обмеженому режимі. Особи обмежують своє пересування та залишаються вдома.	Недотримання урядових наказів було обмеженим.
Федеральний закон <i>CARES</i>	Особи з низьким рівнем доходів	Одноразова виплата в розмірі \$1200 на дорослого, \$700 для осіб на утриманні. Обмежується для осіб з річним доходом в \$75000 і більше.	Особи та домогосподарства з низьким рівнем, часто не подають податкові декларації, тому уряд не може надати їм виплат.
Федеральна програма захисту заробітних плат <i>PPP</i>	Вибрані підприємства	Повне відшкодування заробітної плати працівникам за 8 тижнів, за умови збереження повного штату працівників.	Складна процедура оформлення, що потребує наявності кредитної історії позбавила можливості отримати програму малими

			підприємствами.
Програма допомоги працівникам <i>SUI</i>	Обрані особи, які втратили роботу	Щотижнева виплата в розмірі до \$625, продовж 25 тижнів.	Обмежена здатність обробляти нові заявки про втрату роботи (всього 10000 оброблених заявок на тиждень).
Підтримка працівників в умовах пандемії <i>PUA</i>	Особи з програми <i>SUI</i>	Додаткова тижнева виплата в розмірі \$600, до 31 липня.	Лише особи з опрацьованими заявками <i>SUI</i> мають доступ до програми.
Федеральна програма страхування працівників <i>UIR</i>	Особи які втратили роботу, для яких закінчилася програма <i>SUI</i>	Додаткові 13 тижнів виплат програми <i>SUI</i> , після закінчення періоду програми <i>SUI</i> .	Лише особи з опрацьованими заявками <i>SUI</i> мають доступ до програми.

Туристи

1. Дотримання наперед визначених сценаріїв:
 - a. Туристи дотримуються вхідних сценаріїв візитів.
 - b. До початку пандемії, візити базуються на тижневих даних, наданих адміністрацію туризму США.
2. Дотримання динамічних моделей поведінки:

- a. Після початку пандемії, сценарії візитів базуються на динамічному прогнозі відновлення економіки штату.
3. Дотримання моделей поведінки на основі історичних даних:
 - a. Характеристики поведінки туристів, такі як тривалість візиту, загальні витрати та відвідувані місця, відповідають історичним даним адміністрації туризму США.

Господарства

1. Керування джерелами доходу:
 - a. Домогосподарства мають три основні джерела доходів: зарплата, поза зарплатні доходи (інвестиції, дивіденди, оренда, успадкування) та субсидії.
 - b. Більшість поза зарплатних доходів залишилися стабільними, розмір субсидії визначається в залежності від урядових програм.
2. Реагування на звільнення:
 - a. Зарплатний дохід господарства складається з сукупних зарплат всіх працевлаштованих осіб, що належать до нього.
 - b. Якщо всі працевлаштовані особи були звільнені, зарплатний дохід може наближуватися до нуля.
3. Використання урядової підтримки:
 - a. Приріст доходів господарств в першу чергу спричинений надходженням від різних урядових програм страхування від безробіття та пов'язаних з ними виплат.
 - b. Такі доходи зазвичай нівелюються зменшенням зарплатних доходів, тому не покращують загальну економічну ситуацію господарства.
 - c. Такі доходи в основному використовуються господарствами для задоволення термінових потреб.
4. Зміна поведінки на основі поточних доходів:

- a. Після визначення власного тижневого доходу господарство змінює свої потреба на товари і послуги відповідно до доходів перед *COVID-19*.
- b. Карантинні обмеження додатково знижують попит на певну групу товарів та послуг.

Підприємства

На Рис. 4.8 зображено детальну схему поведінки підприємств.

1. Оцінювання попиту:
 - a. Впродовж кожного тижня, підприємства отримують замовлення від клієнтів, які можуть бути господарствами, туристами або іншими підприємствами.
2. Аналіз замовлень:
 - a. Підприємства порівнюють фактичну кількість замовлень з кількістю замовлень в нормальних умовах (використовуючи той самий тиждень минулого року).
3. Керування працівниками:
 - a. В залежності від кількості поточних замовлень, підприємства роблять замовлення у свої постачальників.
 - b. Підприємства визначають, скільки працівників їм потрібно для виконання замовлень наступного тижня. Якщо це значення відрізняється від фактичної кількості працівників, підприємства адаптують свій персонал, звільняючи або наймаючи нових працівників.

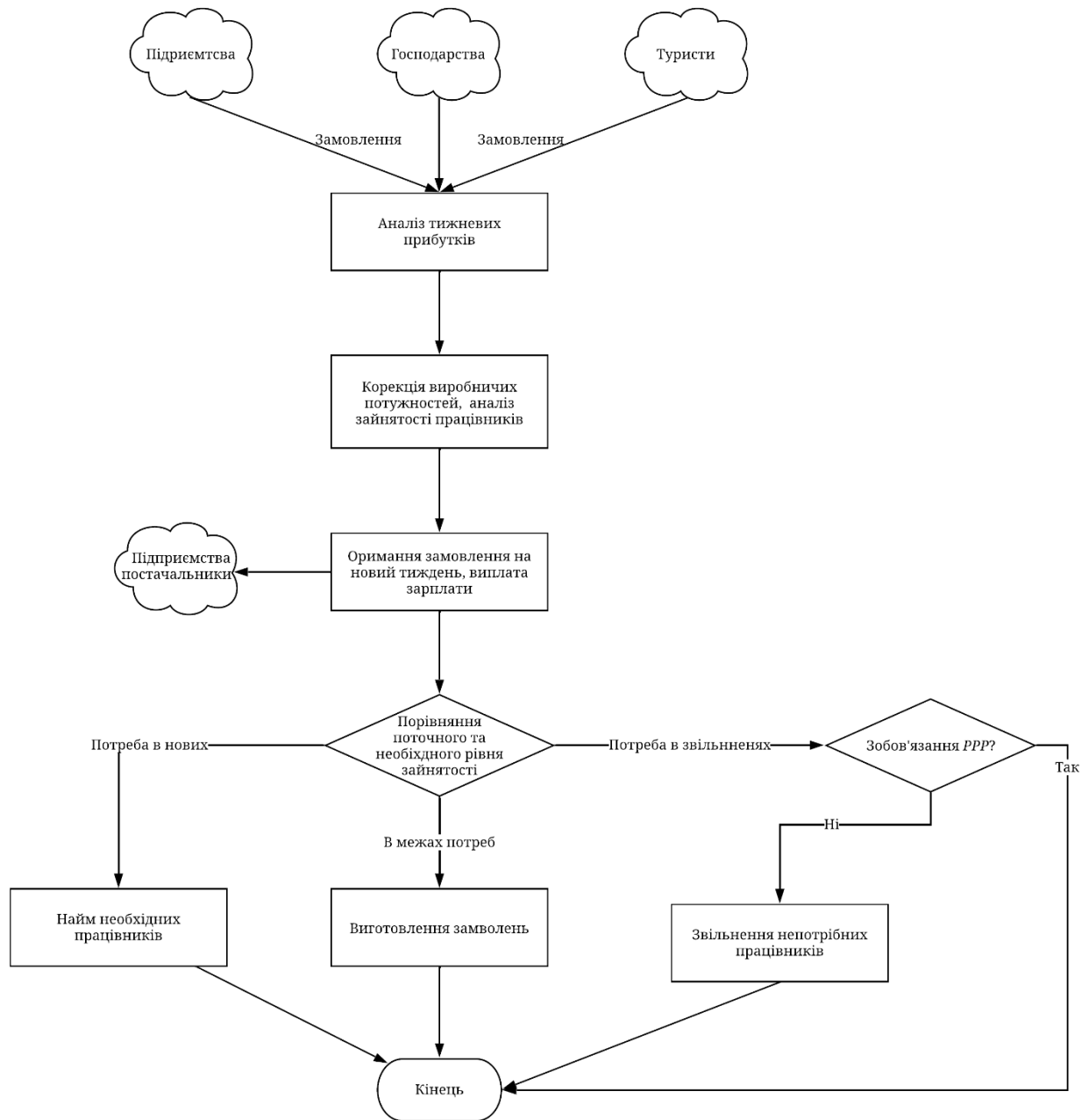


Рис. 4.8. Схема поведінки підприємств: формування очікуваних замовлень, виконання замовлень і корегування рівнів зайнятості.

Під час найму нових працівників, поведінка підприємств включає надання переваги поверненню на роботу нещодавно звільнених працівників. Крім того, в моделі відсутнє мінімальне фіксоване значення кількості працівників. Доти, поки в підприємстві є принаймні один працівник, воно дія в межах економічної мережі, приймаючи та виконуючи замовлення. Агента модель була реалізована з використанням фреймворку *MASON Simulation Framework* [70].

4.2.2 Особливості агентів

Важливо зазначити, що поведінка всіх агентів базується на перших принципах ринкових відносин, підприємства намагаються мінімізувати свої витрати, та максимізувати прибутки, не зважаючи на матеріальний стан працівників. З іншого боку, особи та господарства, не зважаючи на планування тижневих бюджетів можуть їх перевищувати, потребуючи допомоги від уряду.

4.3 МОДЕЛЮВАННЯ СЦЕНАРІЇВ ЗОВНІШНЬОГО ВПЛИВУ

Побудована ангента симуляція дозволяє варіювати вхідні дані у вигляді сценаріїв, та на їх основі прогнозувати майбутній стан агентів та системи в цілому. В цьому розділі було досліджується набір сценаріїв відновлення економіки штату, що базуються на урядових програмах та прогнозу відновлення туризму. Для узгодження результатів симуляції з реальними даними було розроблено процес калібрування.

4.3.1 Формалізація сценаріїв зовнішнього впливу

В Табл. 4.3, наведено характеристику сценаріїв, розроблених для інформування федерального уряду, та сприянню прийняття рішень урядом штату на місцевому рівні. Кожен сценарій не враховує нових потенційних карантинних обмежень, незалежно від траєкторії пандемії. З огляду на те, що урядові прогами є законодавчо затвердженими, вони присутні у всіх сценаріях.

Табл. 4.3. Характеристика досліджуваних сценаріїв.

Сценарій	Характеристика
<i>S4</i>	Відсутність <i>COVID-19</i>
<i>S5</i>	Відсутнє економічне відновлення: відсутні урядові програми підтримки окрім <i>SUI</i>
<i>S3</i>	Швидке економічне відновлення: на кінець рудня 2020 року туризм відновився на 50%
<i>S2</i>	Повільне економічне відновлення: на кінець рудня 2020 року туризм відновився на 20%
<i>S1</i>	Відсутнє економічне відновлення: туризм не відновився

Кожен сценарій починається з 1 січня 2020 року, та відповідає історичним даним до 31 липня 2020 року. Після цього, динаміка притоку туристів визначається на рівні сценарію, див Рис. 4.9.

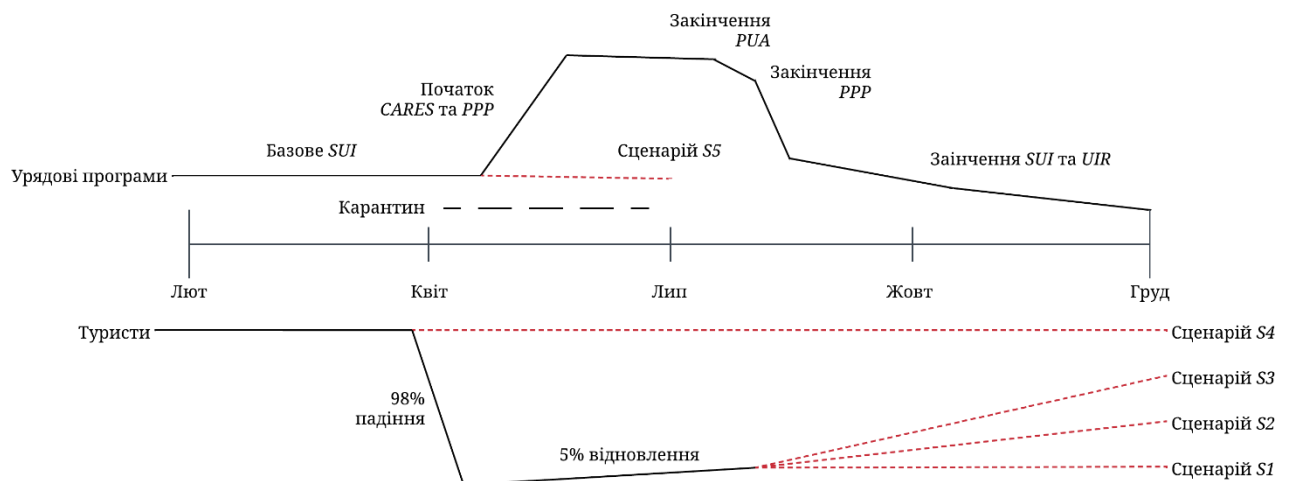


Рис. 4.9. Часова структура досліджуваних сценаріїв.

4.3.2 Калібрування агентної моделі

Рівень безробіття є одним з найважливіших показників для урядів та осіб, що приймають рішення. Для цього, міністерство праці США щотижня збирає дані, про кількість осіб, які втратили роботу та вперше подали заявку на отримання виплат страхування від безробіття *SUI*, а також загальну кількість здійснених страхових виплат. Особі, для того, щоб отримати виплату спочатку необхідно подати заявку, та дочекатися на її обробку та прийняття позитивного або негативного рішення. Крім цього, особа повинна підтвердити статус безробітного та приготувати реквізити для отримання виплати.

Це призводить до частково викривленої статистики, оскільки багато осіб не подають заявку через її складність. Також, особи зазвичай мають сумніви стосовно свого права на отримання такої виплати, або ж працюють на тимчасових роботах, отримуючи одночасно виплати та зарплату. В той же час заявки можуть потребувати складної процедури апеляції.

Це в загальному призводить до запізнених оновлень даних, які можуть включати нові значення для попередніх тижнів, що в свою чергу не дозволяє уряду приймати вчасні рішення.

Саме тому рівень безробіття відіграє ключову роль і в симуляції, для якої архівні дані *SUI* є вхідними даними для калібрування, і водночас одним з значень які моделюються.

Для забезпечення відповідності результатів роботи симуляції реальним даним, було розроблено механізм калібрування. В основі цього процесу лежить алгоритм еволюційних стратегій [71], що проводить пошук на множинні параметрів використаних в симуляції, для наближення результатів до реальних даних. Для процесу калібрування було використано дані про кількість нових заявок та виплачених страхових платежів *SUI*. На Рис. 4.11 та Рис. 4.12 представлено порівняння результатів симуляції та реальних даних після проведення калібрування.

Дані Рис. 4.11 та Рис. 4.12 вказують на дві основні відмінності між реальними та симульованим даними:

- Реальні дані про нові заявки не знизилися у липні, як це відбулося в симуляції. Можна припустити, що це пов'язано з програмою *PPP*. А саме, підприємства могли вичерпати кошти програми *PPP* набагато швидше, ніж було заплановано, оскільки вони використовували ці кошти не тільки для зарплат, але і для оренди приміщень чи комунальних послуг.
- Реальні дані про виплачені страхові платежі в квітні були значено вищі ніж в симуляції. Можна припустити, що це пов'язано з адміністративними затримками, оскільки в кінцевому випадку ці виплати були здійснені для тижнів в яких заявки були поданими але ще не опрацьованими. Варто зазначити, що в симуляції затримані платежі зараховуються в той момент, коли вони були власне виплачені.

Жодна з цих відмінностей не є суттєвою на практиці, враховуючи характеристики досліджуваних сценаріїв.

Значення похибки процесу калібрування розраховується як середнє відносне відхилення між симульованої та реальною статисткою безробіття, див.

Error! Reference source not found..

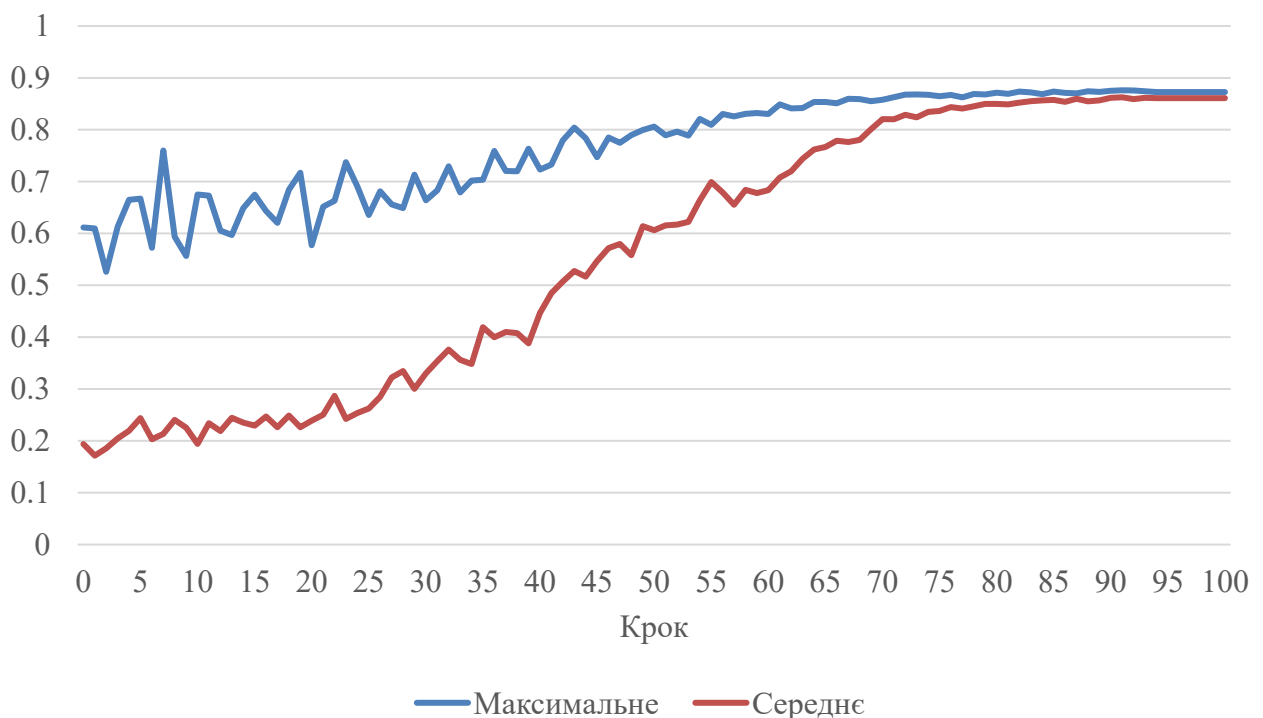


Рис. 4.10. Процес калібрування агентної моделі. Максимальне та середнє значення похибки на кожному кроці.

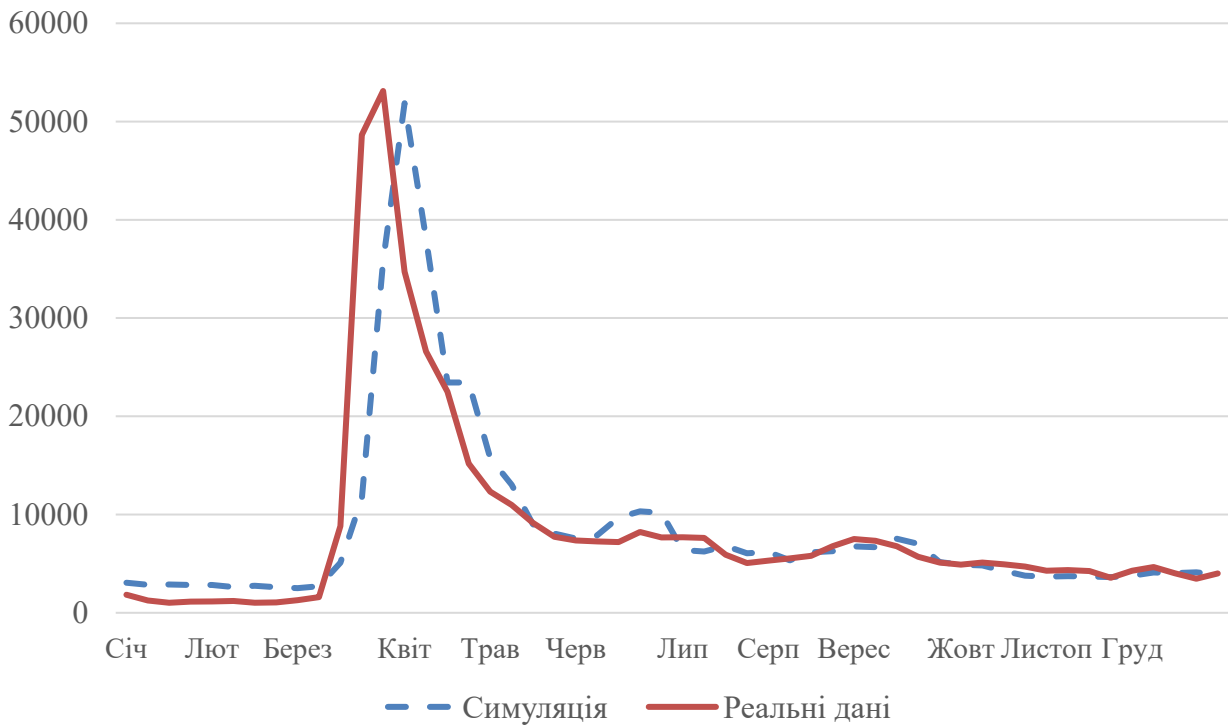


Рис. 4.11. Відповідність між щотижневими симульованими та реальними значеннями кількості нових заявок *SUI* до кінця 2020 року. Симульовані дані отримані як середнє значення результатів 10 симуляції.

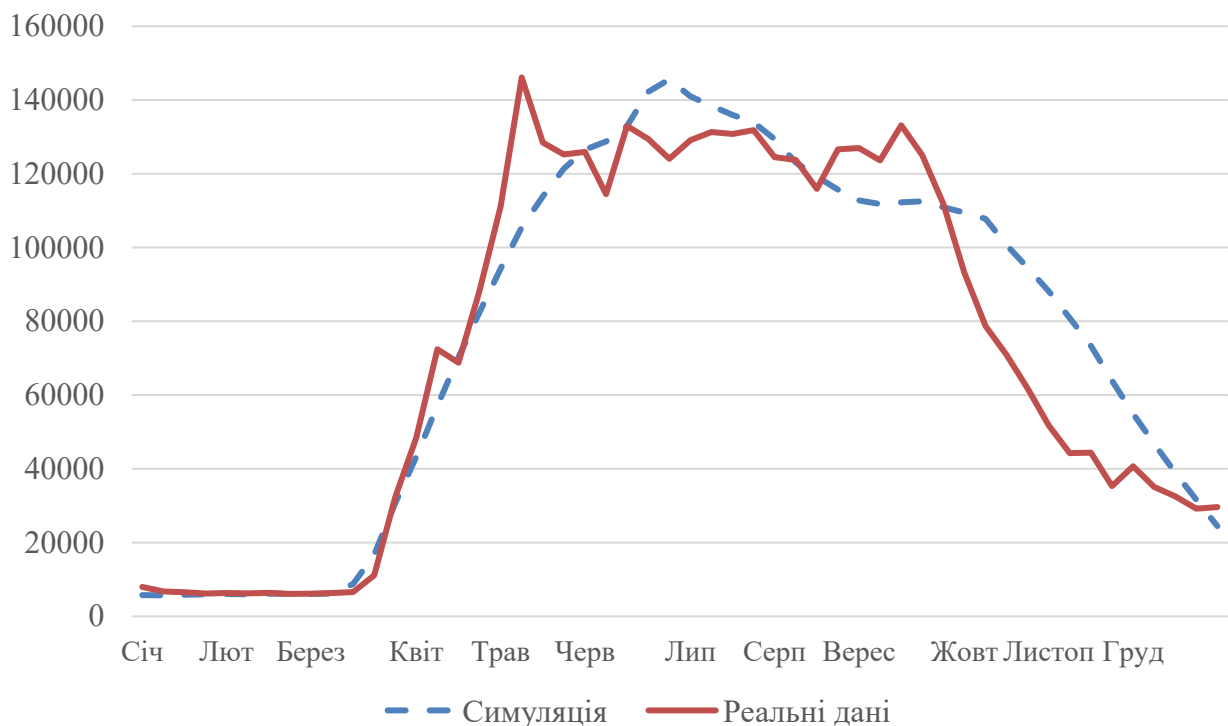


Рис. 4.12. Відповідність між щотижневими симульованими та реальними значеннями кількості виплачених страхових платежів *SUI* до кінця 2020 року. Симульовані дані отримані як середнє значення результатів 10 симуляції.

4.3.3 Прогнозування стану мережі

Після проведення процесу калібрування, агентна модель є готовою до використання для побудови прогнозів. Використовуючи результати моделювання сценаріїв, див. Рис. 4.13 та Рис. 4.14, надаємо відповіді для уряду штату щодо найважливіших показників економічної ситуації:

На скільки значними є поточні економічні збитки від пандемії COVID-19?

Враховуючи особливості економічних втрат як функцію від кількості транзакцій, див Рис. 4.13, для сценаріїв *S5*, *S1*, *S2*, та *S3* економічні збитки становлять 21% від до пандемічних показників на кінець серпня 2020 року, та зростають до 26-30% на кінець 2020 року. Без впровадження урядових програм, цей показник становив би 48% на кінець серпня, та 54% на кінець року.

На скільки урядові програми змогли зменшити ці збитки?

Урядові програми відіграли ключову роль в запобіганні катастрофічних збитків. Без урядових програм, рівень безробіття міг би досягнути 1000000 безробітних. Уряд зміг стримати безробіття, за берегти його рівень в розмірі не більше 400000 безробітних, див Рис. 4.14 (сценарій *S5* в порівнянні з *S1*, *S2* або *S3*).

Найбільш ефективною виявилася програма *PPP*. Очікується, що з вичерпання коштів *PPP*, багато працевлаштованих та повторно працевлаштованих, багато з яких більше не можуть отримувати виплати програми *SUI*, знову опиняться без роботи. Поточна реальна ситуація 2020 року близька до сценаріїв *S1* та *S2* не є достатньою для підтримки поточного рівня зайнятості населення. Тільки сценарій швидкого відновлення *S3* може утримати та знизити рівень безробіття до кінця 2020 року.

Якою повинна бути швидкість відновлення кількості туристів та відвідувачів островів для стабілізації економіки штату?

Під час сценарію повільного відновлення *S2*, кількість відвідувачів до кінця року становитиме 25% від до пандемічних показників, що стабілізує економіку в термінах обсягів економічної активності. Однак, сценарій *S2* не зможе запобігти росту безробіття, і виглядає менш імовірним через наслідки пандемії. Це змушує уряд штату розглядати сценарій *S1* як найбільш реалістичний, та готувати програми які б могли вплинути на його негативні наслідки.

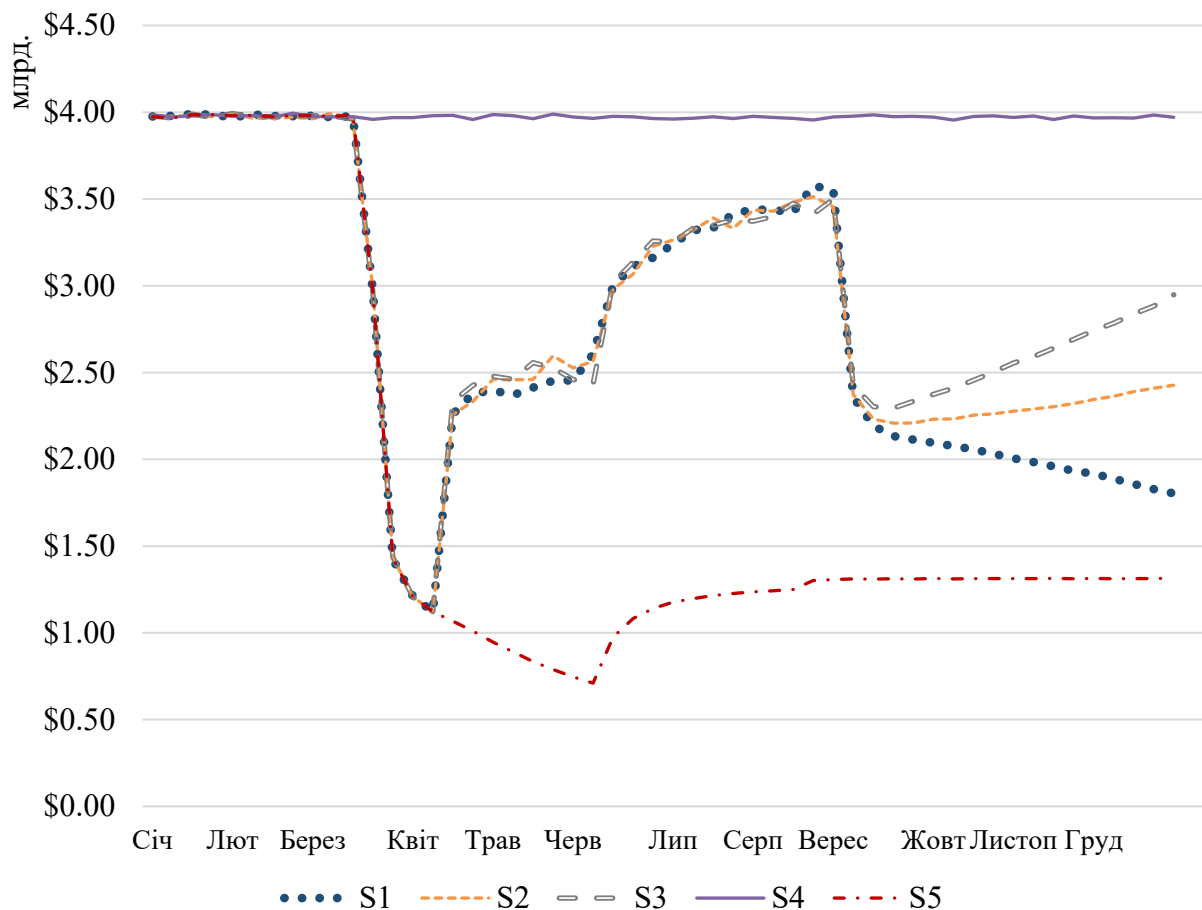


Рис. 4.13. Загальний щотижневий обсяг транзакцій в симульованій економіці штату, не враховуючи зарплати та урядові програми, серед яких транзакції між господарствами та підприємства, туристами та підприємствами а також між підприємствами. Симульовані дані отримані як середнє значення результатів 10 симуляцій.

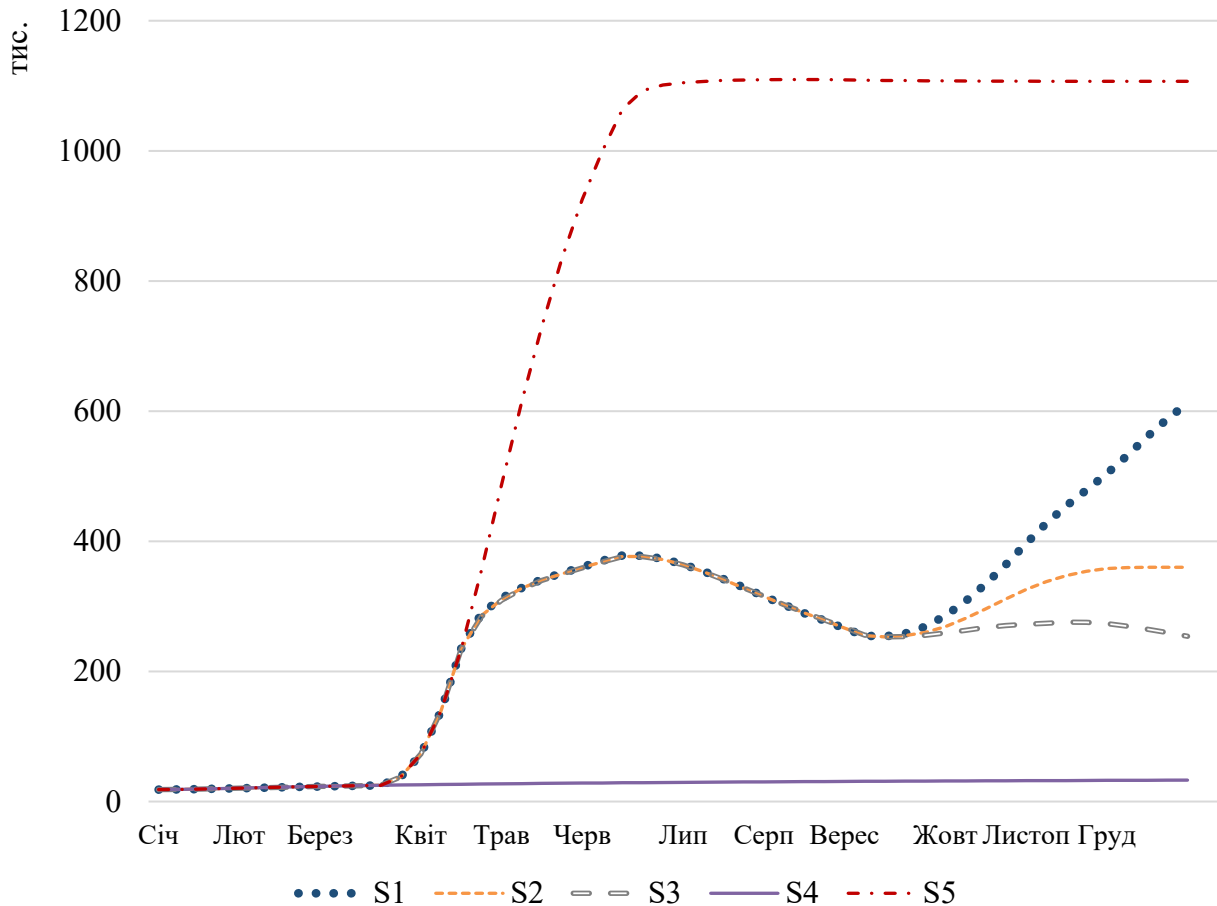


Рис. 4.14. Загальна кількість безробітних в штаті Гаваї впродовж 2020 року. Симульовані дані отримані як середнє значення результатів 10 симуляцій.

4.3.4 Аналіз отриманих результатів

Агентне моделювання з використанням синтетичних популяцій дозволило отримати точні результати короткострокових та довготривалих прогнозів стану економічної мережі штату Гаваї. Такі прогнози дозволили інформувати уряди та осіб, відповідальних за прийняття рішень в кризових умовах. Прогнози з використанням агентного моделювання особливо актуальні в ситуаціях, які характеризуються відмінністю від історичних даних, та можливість збору даних обмежено можлива через кризову ситуацію, у випадку цього дослідження – пандемію *COVID-19* [18].

4.4 ВИСНОВКИ ПО ЧЕТВЕРТОМУ РОЗДІЛІ

У четвертому розділі дисертаційної роботи було проведено збір публічних даних як основу для побудови синтетичної популяції населення штату Гаваї. Для побудови синтетичної популяції було реалізовано алгоритм імітації відпалу з перевіркою результатів на основних демографічних характеристиках населення.

Для моделювання процесів в економічних мережах, було розроблено агенту модель взаємодії населення, в якій кожному агенту було призначено набір правил поведінки. Для перевірки коректності агентної моделі, було реалізовано процес калібрування з використанням еволюційних стратегій.

Розроблено програмний комплекс застосування агентної моделі економічних зв'язків для побудов прогнозів стану економічної мережі в випадку наявності зовнішніх факторів.

ВИСНОВКИ

Дисертаційна робота присвячена дослідженню та формалізації еволюційних процесів розповсюдження інформації в соціальних мережах, на основі побудови ефективних моделей відображення складних мереж та моделюванні взаємодій між сутностями в складних мережах.

Було проведено огляд актуальних досліджень розповсюдження впливу та інформації в соціальних мережах. Детально проаналізовано основні особливості моделювання процесу розповсюдження інформації в контексті наборів даних на яких проводилося дослідження, обраної методології та розв'язання проблеми вибору оптимальних цільових аудиторій для оптимізації впливу. Виходячи з результатів розглянутих досліджень, було встановлено принципові недоліки підходів до дослідження розповсюдження інформації, що характеризуються системною відсутністю реалістичних моделей соціальних мереж на основі публічних даних та застосуванню методології, що не здатна описувати динамічну структуру процесів та взаємодій в соціальних мережах та соціальних мереж загалом.

Було розглянуто актуальні дослідження соціальних зв'язків в економічних мережах, як важливого аспекту соціальної взаємодії. Детально проаналізовано основні особливості моделювання процесів в економічних мережах. Виходячи з результатів розглянутих досліджень, було встановлено схожі недоліки моделювання, в основі яких – відсутність моделей реальних економічних мереж, а також методології, що не враховує поведінковий аспект.

Для моделювання розповсюдження інформації в соціальних мережах та моделювання процесів в економічних мережах, було запропоновано використати методологію аналізу складних мереж з використанням агентного моделювання та синтетичних популяцій.

Було формалізовано процеси розповсюдження інформації в соціумі, які згодом масштабовані як основна складова сучасних соціальних мереж. Було

розглянуто основні механізми та моделі розповсюдження інформації в соціальних мережах.

Було досліджено динамічні фактори впливу на розповсюдження інформації в соціальних мережах, та запропоновано застосування моделювання соціальних мереж, яке базується на дослідженні поведінки користувачів. У третьому розділі дисертаційної роботи було проведено дослідження особливостей різних соціальних мереж, формалізовано їх основні сутності та способи взаємодії. Для ефективного моделювання процесу розповсюдження інформації, було запропоновано модель мережі перетину аудиторій на основі публічних даних. В умовах обмеження соціальними мережами доступу до даних, було показано як модель мережі подібності публікацій та мережі взаємодій дозволяють оцінювати перетин аудиторій. З використанням мережі перетину аудиторій, було досліджено динаміку еволюції спільнот соціальної мережі *Telegram* та процес розповсюдження впливу.

Для розв'язання задачі оптимального поширення впливу було запропоновано модель з використанням агентів, а мережу перетину аудиторій використано для вибору оптимальних цільових аудиторій для агентів. Для ефективного застосування агентів, було проведено моделювання поширення інформації для випадків з різною кількістю агентів.

Було проведено збір публічних даних як основу для побудови синтетичної популяції населення штату Гаваї. Для побудови синтетичної популяції було реалізовано алгоритм імітації відпалу з перевіркою результатів на основних демографічних характеристиках населення.

Для моделювання процесів в економічних мережах, було розроблено агенту модель взаємодії населення, в якій кожному агенту було призначено набір правил поведінки. Для перевірки коректності агентної моделі, було реалізовано процес калібрування з використанням еволюційних стратегій. Розроблено програмний комплекс застосування агентної моделі економічних зв'язків для

побудови прогнозів стану економічної мережі в випадку наявності зовнішніх факторів.

Таким чином, основними результатами дисертаційної роботи, що виносяться на захист, є те, що:

- запропоновано підхід до визначення перетину аудиторій соціальних мереж, застосовано його до розв'язання проблеми побудови мережі перетину аудиторій на публічних даних.
- застосовано мережі перетину аудиторій для ефективного вибору цільових аудиторій.
- вперше запропоновано підхід до розповсюдження інформації на основі агентів з використання оптимального розташування.
- розроблено підхід до побудови синтетичних популяцій на основі публічних даних.
- розроблено агенту модель мережі економічних зв'язків.
- розроблено програмний комплекс застосування агентної моделі для побудови точних прогнозів еволюції соціальних мереж економічних зв'язків.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] M. Newman, *Networks: An Introduction*, Oxford University Press, 2010.
- [2] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, p. 47–97, January 2002.
- [3] D. J. Watts and S. H. Strogatz, "Collective dynamics of ‘small-world’ networks," *Nature*, vol. 393, p. 440–442, 1998.
- [4] P. Erdős and A. Rényi, "On random graphs I," *Publicationes Mathematicae (Debrecen)*, vol. 6, p. 290–297, 1959.
- [5] A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks," *Science*, vol. 286, pp. 509-512, 1999.
- [6] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Networks*, vol. 1, pp. 215-239, 1978.
- [7] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, p. 7821–7826, June 2002.
- [8] M. J. Keeling and K. T. D. Eames, "Networks and epidemic models," *Journal of The Royal Society Interface*, vol. 2, pp. 295-307, 2005.
- [9] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, 1994.
- [10] R. Albert, H. Jeong and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, p. 378–382, 2000.
- [11] R. J. Beckman, K. A. Baggerly and M. D. McKay, "Creating synthetic baseline populations," *Transportation Research Part A: Policy and Practice*, vol. 30, pp. 415-429, 1996.
- [12] J. Barthélemy and P. Toint, "Synthetic Population Generation Without a Sample," *Transportation Science*, vol. 47, pp. 266-279, May 2013.

- [13] X. Ye, K. Konduri, R. Pendyala, B. Sana and P. Waddell, "Methodology to match distributions of both household and person attributes in generation of synthetic populations," January 2009.
- [14] W. E. Deming and F. F. Stephan, "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known," *The Annals of Mathematical Statistics*, vol. 11, p. 427 – 444, 1940.
- [15] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*, vol. 4, Springer, 2006.
- [16] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proceedings of the National Academy of Sciences*, vol. 99, pp. 7280-7287, 2002.
- [17] R. L. Axtell and J. M. Epstein, "Agent-Based Modeling: Understanding Our Creations," *The Bulletin of the Santa Fe Institute*, Vols. 28-32, 1994.
- [18] R. Axtell, "Why Agents? On the Varied Motivations for Agent Computing in the Social Sciences," December 2000.
- [19] N. Gilbert and K. Troitzsch, *Simulation For The Social Scientist*, McGraw-Hill Education, 2005.
- [20] T. C. Schelling, "Dynamic models of segregation," *The Journal of Mathematical Sociology*, vol. 1, pp. 143-186, 1971.
- [21] J. Epstein, "Modelling to contain pandemics," *Nature*, vol. 460, no. 7256, 2009.
- [22] L. Tesfatsion and K. Judd, "Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics," *Iowa State University, Department of Economics, Staff General Research Papers*, January 2006.
- [23] L. Dang, Z. Chen, J. Lee, M.-H. Tsou and X. Ye, "Simulating the spatial diffusion of memes on social media networks," *International Journal of Geographical Information Science*, vol. 33, pp. 1545-1568, 2019.

- [24] P. Sobkowicz, M. Kaschesky and G. Bouchard, "Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web," *Government Information Quarterly*, vol. 29, pp. 470-479, 2012.
- [25] I. Garibay, A. V. Mantzaris, A. Rajabi and C. E. Taylor, "Polarization in social media assists influencers to become more influential: analysis and two inoculation strategies," *Scientific Reports*, vol. 9, p. 18592, 2019.
- [26] F. Cinus, M. Minici, C. Monti and F. Bonchi, "The Effect of People Recommenders on Echo Chambers and Polarization," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 16, pp. 90-101, May 2022.
- [27] T. K. Biswas, A. Abbasi and R. K. Chakraborty, "A two-stage VIKOR assisted multi-operator differential evolution approach for Influence Maximization in social networks," *Expert Systems with Applications*, vol. 192, p. 116342, 2022.
- [28] J. Stein, M. Keuschnigg and A. van de Rijt, "Network segregation and the propagation of misinformation," *Scientific Reports*, vol. 13, p. 917, 2023.
- [29] M. Azzimonti and M. Fernandes, "Social media networks, fake news, and polarization," *European Journal of Political Economy*, vol. 76, p. 102256, 2023.
- [30] R. Mallipeddi, S. Kumar, C. Sriskandarajah and Y. Zhu, "A Framework for Analyzing Influencer Marketing in Social Networks: Selection and Scheduling of Influencers," *SSRN Electronic Journal*, January 2018.
- [31] E. Yanchenko, T. Murata and P. Holme, *Influence maximization on temporal networks: a review*, 2023.
- [32] A. Chin, D. Eckles and J. Ugander, *Evaluating stochastic seeding strategies in networks*, 2020.
- [33] S. Rothut, H. Schulze, J. Hohner and D. Rieger, "Ambassadors of ideology: A conceptualization and computational investigation of far-right influencers, their networking structures, and communication practices," *New Media & Society*, p. 146144482311644, May 2023.

- [34] C. Ling, J. Jiang, J. Wang, M. Thai, L. Xue, J. Song, M. Qiu and L. Zhao, *Deep Graph Representation Learning and Optimization for Influence Maximization*, 2023.
- [35] E. Sahafizadeh and B. T. Ladani, "Soft rumor control in mobile instant messengers," *Physica A: Statistical Mechanics and its Applications*, vol. 609, p. 128359, 2023.
- [36] A. Averza, K. Slhoub and S. Bhattacharyya, "Evaluating the Influence of Twitter Bots via Agent-Based Social Simulation," *IEEE Access*, vol. 10, pp. 129394-129407, 2022.
- [37] F. N. van der Vlist, A. Helmond, M. Burkhardt and T. Seitz, "API Governance: The Case of Facebook's Evolution," *Social Media + Society*, vol. 8, p. 20563051221086228, 2022.
- [38] J. Calma, "Twitter just closed the book on academic research," *The Verge*, May 2023.
- [39] A. Rakha, H. Hettiarachchi, D. Rady, M. Gaber, E. Rakha and M. Abdelsamea, "Predicting the Economic Impact of the COVID-19 Pandemic in the United Kingdom Using Time-Series Mining," *Economies*, vol. 9, p. 137, September 2021.
- [40] X. Tang, Z. Li, X. Hu, Z. Xu and L. Peng, "Self-correcting error-based prediction model for the COVID-19 pandemic and analysis of economic impacts," *Sustainable Cities and Society*, vol. 74, p. 103219, 2021.
- [41] H. Zhang, H. Song, L. Wen and C. Liu, "Forecasting tourism recovery amid COVID-19," *Annals of Tourism Research*, vol. 87, p. 103149, 2021.
- [42] B. Teng, S. Wang, Y. Shi, Y. Sun, W. Wang, W. Hu and C. Shi, "Economic recovery forecasts under impacts of COVID-19," *Economic Modelling*, vol. 110, p. 105821, 2022.

- [43] J. Chen, A. Vullikanti, J. Santos, S. Venkatramanan, S. Hoops, H. Mortveit, B. Lewis, W. You, S. Eubank, M. Marathe, C. Barrett and A. Marathe, *Epidemiological and Economic Impact of COVID-19 in the US*, 2020.
- [44] C. Li and X. Lin, "COVID-19 and trade: Simulated asymmetric loss," *Journal of Asian Economics*, vol. 75, p. 101327, 2021.
- [45] M. Akbarpour, C. Cook, A. Marzuoli, S. Mongey, A. Nagaraj, M. Saccarola, P. Tebaldi, S. Vasserman and H. Yang, "Socioeconomic Network Heterogeneity and Pandemic Policy Response," *SSRN Electronic Journal*, January 2020.
- [46] J. E. Kim, H. Choi, Y. Choi and C. H. Lee, "The economic impact of COVID-19 interventions: A mathematical modeling approach," *Frontiers in Public Health*, vol. 10, 2022.
- [47] E. M. a. S. A. a. Q. M. M. Rogers, "Diffusion of innovations," in *An integrated approach to communication theory and research*, Routledge, 2014, pp. 432-448.
- [48] M. Granovetter, "Threshold models of collective behavior," *American journal of sociology*, vol. 83, pp. 1420-1443, 1978.
- [49] Є.В.Івохін, "Про деякі математичні моделі формалізації соціо-інформаційних потоків," *існик Київського національного університету імені Тараса Шевченка. Серія ФМН.*, no. Вип.2., pp. 70-73, 2017.
- [50] J. M. E. Robert L Axtell, "Agent-based modeling: Understanding our creations," *The Bulletin of the Santa Fe Institute*, vol. 9, no. 2, pp. 28-32, 1994.
- [51] S. Kemp, "datareportal," 11 5 2023. [Online]. Available: <https://datareportal.com/essential-facebook-stats>.
- [52] J. a. M. K. L. Berger, "What makes online content viral?," *Journal of marketing research*, vol. 49, pp. 192-205, 2012.
- [53] e. a. Freberg, "Who are the social media influencers? A study of public perceptions of personality," *Public relations review*, vol. 37, pp. 90-92, 2011.

- [54] "Standard v1.1 | Docs | Twitter Developer Platform," [Online]. Available: <https://developer.twitter.com/en/docs/twitter-api/v1>.
- [55] S. Mukerjee, S. Majó-Vázquez and S. González-Bailón, "Networks of Audience Overlap in the Consumption of Digital News," *Journal of Communication*, vol. 68, pp. 26-50, February 2018.
- [56] Y. a. J. R. a. Z. Z.-H. Zhang, "Understanding bag-of-words model: A statistical framework," *International Journal of Machine Learning and Cybernetics*, vol. 1, pp. 43-52, 2010.
- [57] B. Trstenjak, S. Mikac and D. Donko, "KNN with TF-IDF based Framework for Text Categorization," *Procedia Engineering*, vol. 69, pp. 1356-1364, 2014.
- [58] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, *Attention Is All You Need*, 2023.
- [59] D. Cer, Y. Yang, S.-y. Kong, N. Hua, N. Limtiaco, R. S. John, N. Constant, M. Guajardo-Cespedes, S. Yuan, C. Tar, Y.-H. Sung, B. Strope and R. Kurzweil, *Universal Sentence Encoder*, 2018.
- [60] G. Hinton and S. Roweis, "Stochastic Neighbor Embedding," vol. 15, June 2003.
- [61] S. Greenland, S. Senn, K. Rothman, J. Carlin, C. Poole, S. Goodman and D. Altman, "Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations," *European Journal of Epidemiology*, vol. 31, April 2016.
- [62] А. Ольшанников, "Владимир Соловьев: как стать звездой Telegram меньше, чем за год," *PIA URA.RU*, January 2020.
- [63] A. Roth, "Russia blocks millions of IP addresses in battle against Telegram app," *The Guardian*, April 2018.
- [64] V. D. Blondel, J.-L. Guillaume, R. Lambiotte and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, p. P10008, October 2008.

- [65] S. Ruggles, S. Flood, M. Sobek, D. Brockman, G. Cooper, S. Richards and M. Schouwiler, *IPUMS USA: Version 13.0*, Minneapolis, MN: IPUMS, 2023.
- [66] J. Kim and S. Lee, "A simulated annealing algorithm for the creation of synthetic population in activity-based travel demand model," *KSCE Journal of Civil Engineering*, vol. 20, November 2015.
- [67] A. Justel, D. Peña and R. Zamar, "A multivariate Kolmogorov-Smirnov test of goodness of fit," *Statistics & Probability Letters*, vol. 35, pp. 251-259, October 1997.
- [68] B. Kaminski and P. Szufel, *Julia 1.0 Programming Cookbook*, Birmingham, England: Packt Publishing, 2018.
- [69] "Southampton Geospatial. University of Southampton.," [Online]. Available: <https://www.southampton.ac.uk/research/institutes-centres/southampton-geospatial>.
- [70] S. Luke, C. Cioffi, L. Panait and K. Sullivan, "Mason: A new multi-agent simulation toolkit," *Proceedings of the 2004 SwarmFest Workshop*, January 2004.
- [71] J. Thiele, W. Kurth and V. Grimm, "Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models: A Cookbook Using NetLogo and 'R'," *Journal of Artificial Societies and Social Simulation*, The, vol. 17, June 2014.

ДОДАТОК 1

Табл. 1. Дані транзакцій 20 галузей, семи секторів кінцевого попиту, включаючи відвідувачів, та п'ять секторів кінцевих платежів. Податки на виробництво та імпорт з вирахуванням субсидій. Усі значення наведено в мільйонах доларів США.

Галузь	<i>Professional services</i>	<i>Business services</i>	<i>Visitor expenditures</i>
Agriculture	1.6	2.8	40.6
Mining and construction	16.6	28.2	0
Food processing	1.2	0	101.6
Other manufacturing	97.4	222.6	98.4
Transportation	129.6	80	5365
Information	159.8	179.8	33.6
Utilities	29.4	79	0
Wholesale trade	52.6	53	195.2
Retail trade	82.8	91.8	3742.8
Finance and insurance	64.4	71	0
Real estate and rentals	605	348.2	2606.2
Professional services	352	345.4	223.4
Business services	134.8	338.2	580.6
Educational services	21.6	9.8	254.2
Health services	0.8	0.2	269
Arts and entertainment	0	0	944.8
Accommodation	44.2	15.4	10092.2
Eating and drinking	49	42.4	2896.4
Other services	60.4	67.2	237.6
Government	64	99	151.2

Imports	398	344.8	1970.6
Labor income	4885.2	4614	N/A
TOPILS	274	166	N/A
Other capital costs	914.8	704	N/A
Total Value added	6074	5484	N/A
Output	8438.2	7902.4	N/A
Earnings	4554.4	4130.2	N/A
State taxes	662.8	625	N/A

ДОДАТОК 2

Табл. 2. Незмінні параметри економічної симуляції.

Параметр	Значення
<i>CARESAmount</i>	\$1200
<i>SUIAmount</i>	\$625
<i>PUAAmount</i>	\$600
<i>covidQuarantineStart - covidQuarantineStop (week)</i>	11 - 23
<i>CARESStart - CARESStop (week)</i>	18 - 22
<i>PPPStart – PPPStop (week)</i>	16 - 35
<i>UIRStart (week), UIRDuration (weeks)</i>	18, 13
<i>PUAStart - PUAStop (week)</i>	18 - 31
<i>SUIDuration (weeks)</i>	26
<i>visitorDecreaseStart – visitorDecreaseStop (week)</i>	11 - 34

ДОДАТОК 3

Табл. 3. Калібровані параметри економічної симуляції.

Параметр	Початкове значення	Каліброване значення
<i>hiringSlack</i>	0.4	0.16
<i>firingSlack</i>	0.0	0.34
<i>CARESDemandEffectivness</i>	0.25	0.37
<i>PPPDemandEffectivness</i>	0.1	0.02
<i>PPPProcessingSpeed</i>	0.75	0.15
<i>SUIDemandEffectivness</i>	0.1	0.6
<i>SUIProcessingSpeed</i>	15000	19471
<i>SUIValidClaimThreshold</i>	0.75	0.93
<i>businessDemandMultiplierCoefficient</i>	0.75	0.49
<i>businessInstantDemandMultiplierCoefficient</i>	0.25	0.51
<i>businessWageReduction</i>	0.5	0.76
<i>businessBaseWageReduction</i>	0.9	0.70
<i>businessHiringIncrease</i>	0.25	0.7
<i>householdDemandReduction</i>	0.5	0.8
<i>businessToBusinessDemandReduction</i>	0.75	0.35
<i>householdToBusinessDemandReduction</i>	0.75	0.27
<i>visitorToBusinessDemandReduction</i>	0.5	0.42
<i>rehirePriorityCoefficient</i>	0.025	0.74